N. P. Sudharshana
Lina Mukhopadhyay *Editors*

Task-Based Language Teaching and Assessment

Contemporary Reflections from Across the World

Springer

Task-Based Language Teaching and Assessment

N. P. Sudharshana · Lina Mukhopadhyay
Editors

Task-Based Language Teaching and Assessment

Contemporary Reflections from Across the World

Springer

Editors
N. P. Sudharshana
Department of Humanities and Social Science
Indian Institute of Technology Kanpur
Kanpur, Uttar Pradesh, India

Lina Mukhopadhyay
Department of Training and Development
English and Foreign Languages University
Hyderabad, Telangana, India

ISBN 978-981-16-4225-8 ISBN 978-981-16-4226-5 (eBook)
https://doi.org/10.1007/978-981-16-4226-5

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Using tasks for language teaching emerged as an alternative to structural approaches to language teaching. Task-Based Language Teaching (TBLT), now, is one of the most popular language teaching methods, and numerous coursebooks are being developed across the world based on TBLT. We are also moving towards using tasks for language assessment as well, a domain that has hitherto been dominated by discrete multiple-choice items with some occasional use of performance-based assessments in large-scale tests. While TBLT is now a well-established teaching methodology, Task-Based Language Assessment (TBLA) is still in its nascent stage with respect to its application for classroom-based purposes. At this juncture, we must note that one of the most significant contributions to TBLT came from N. S. Prabhu from Indian contexts through his Bangalore project in the 1980s. TBLT/TBLA continues to be a robust research area attracting the attention of both practising teachers and language researchers working from diverse fields.

This volume aims to bring together multiple perspectives on TBLT and TBLA primarily from the ESL/EFL contexts. This book has two sections: the first one has papers focusing on TBLT, whereas the second one consists of papers on TBLA. It is a judicious mix of theoretical and experimental research papers contributed by SLA researchers, teacher educators and practising teachers from across the world. We sincerely hope that with an interdisciplinary appeal, this book will be a valuable resource for researchers in task-based language teaching and assessment and for teachers with practical suggestions for designing and using tasks for teaching and testing.

The origin of this volume lies in discussion sessions with our graduate students. We are forever thankful to them for insights, comments and perspectives. We would like to thank all our contributors for such a great amount of patience! All the papers in this volume were peer-reviewed, and we are extremely thankful to our reviewers: Profs. Malathy Krishnan, J. Willis, Geetha Durairajan, Jacob Tharu, Maya Pandit, Meera Srinivas and Santosh Mahapatra! This work would not have been possible

without the support and understanding of the family members, friends, the institute authorities and all the well-wishers. Big thanks to all of them!

Kanpur, India — N. P. Sudharshana
Hyderabad, India — Lina Mukhopadhyay

Contents

Editors and Contributors

About the Editors

N. P. Sudharshana is Assistant Professor of English in the Department of Humanities and Social Sciences at IIT Kanpur, India. His main research areas include applying cognitive semantics to teaching and learning English as a second language, materials development, and classroom language testing. He has previously worked with many state governments in India on school curriculum design and implementation projects. He is the co-author of *English for Technical Communication* (2016) and *English for Engineers* (2018) both published by Cambridge University Press, India.

Lina Mukhopadhyay is Professor in the School of English Language Education, The English and Foreign Languages University, Hyderabad, India. Her research interests are second language acquisition, reading development, language assessment and testing, and bi/multilingual education, and she has publications in each of these areas. She is currently involved in a multilingual education project in collaboration with the University of Cambridge and funded by the ESRC-DFID, UK. The five-year-long project is examining the interrelatedness of language of instruction, school skills, and cognition of young learners studying in challenging contexts in government low-aid primary schools in India.

Contributors

Gavin Bui holds a Ph.D. in applied linguistics. He is an associate professor and associate head in the Department of English at the Hang Seng University of Hong Kong.

Rod Ellis is Research Professor in Curtin University, Australia; a longstanding professor at Anaheim University; a visiting professor at Shanghai International

Studies University; and Emeritus Distinguished Professor of the University of Auckland. He is also a fellow of the Royal Society of New Zealand. He has written extensively on second-language acquisition and task-based language teaching. His most recent (co-authored) book is *Task-based Language Teaching: Theory and Practice* (2020) published by Cambridge University Press.

Roger Gilabert is an associate professor and researcher within the Language Acquisition Research Group (GRAL) at the University of Barcelona. His research interests include second-language production and interaction, task and syllabus design, oral and written modes and CAF, TV genres and subtitles and captions, and personalization and adaptivity in serious games for reading skills (currently working on a current Horizon 2020 project).

Chihiro Inoue is Senior Lecturer at the Centre for Research in English Language Learning and Assessment, University of Bedfordshire, specializing in the assessment of L2 speaking and listening, particularly in the task design, test-taker processes and features of learner language. Her recent TBLT-related publications include the effects of planning time with TOEFL lecture listening-into-speaking tasks (*ETS Research Report*, forthcoming) and variables of accuracy and syntactic complexity (*Language Learning Journal*, 2016).

Aleksandra Malicka focuses on second-language acquisition in traditional and online contexts, learning based on pedagogical tasks, curriculum design, personalized learning and role of individual differences in the process of learning a second language. He is a member of the Open University of Catalonia (UOC) TechSLA Lab research group.

Sajit M. Mathews has a Ph. D. in English from the Department of Humanities and Social Sciences at the Indian Institute of Technology Kanpur, India. His research interests are TBLT, TBLA, oral proficiency teaching and assessment and second-language acquisition. He has taught English at the University of Hyderabad and Rajiv Gandhi University of Knowledge Technologies, Kadapa, India. He has published several articles for teachers in magazines and blogs.

Veena Nair is Assistant Professor (English) at Mar Baselios College of Engineering and Technology, Thiruvananthapuram. Her major areas of interest are cognition and second-language writing, second-language acquisition and learning, bilingualism and language across the curriculum.

Fumiyo Nakatsuhara is Reader in Language Assessment at the Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire, UK. She has carried out a number of international testing projects, working with ministries, universities and examination boards. Her publications

include the book *The Discourse of the IELTS Speaking Test* (2018, CUP) and articles in journals such as *Language Testing*, *Language Assessment Quarterly*, *Modern Language Journal* and *System.*

Pankaj Narke is Assistant Professor in the Department of Languages at VIT-AP University, India. With a Ph.D. in English Language Education from EFL University, he specializes in the area of technology in teaching and assessment of ESL.

Vrishali K. Patil is Ph.D. in English Language Education from the English and Foreign Languages University, Hyderabad. Her research areas are cognition, reading assessment, reading comprehension and multilingual education.

N. S. Prabhu is a pioneer in task-based language teaching and is the author of 'Second Language Pedagogy' (OUP, 1987). He is widely known in the ELT profession for his 'Bangalore Project', a five-year experiment in 'procedural syllabus' that has significantly influenced current conceptualizations of 'task' and 'task-based language teaching'. His project also inspired numerous similar interventions in ESL/EFL contexts across the world. He worked at National University of Singapore as Professor of Applied Linguistics and has been visiting faculty at universities in Hong Kong, Japan and Brazil.

Esther Ramani and **Michael Joseph** are retired professors and currently honorary senior research fellows in the School of Languages and Literatures, Rhodes University, South Africa. They were part of a communicational teaching of English to young learners, a project inspired by Prabhu's research. Ramani along with Joseph was the first to launch a bilingual degree, using an indigenous African language along with English at University of Limpopo, South Africa.

Shruti Sircar is Professor in Linguistics and Contemporary English at the English and Foreign Languages University, Hyderabad. Her major areas of interest are spelling development, reading research, second-language acquisition especially in the areas of verb semantics and cognitive linguistics.

Vasim Tamboli has a Ph.D. in English Language Education from the English and Foreign Languages University. He has worked as Research Associate on ESRD-DFID-funded project titled 'Multilingualism and Multiliteracy: Raising learning outcomes in challenging contexts in primary schools across India'. He is interested in language assessment, task complexity and multilingual education.

Baburam Upadhaya is a doctoral student of English at the Indian Institute of Technology Kanpur, India. His research interests include teaching metaphor to ESL learners from a cognitive perspective and task-based language teaching.

Jane Willis is a writer of several prize-winning books and has worked extensively overseas as an English teacher and trainer before moving to Aston University's

Masters in TESOL programmes. Now retired, she advises teachers and gives talks on task-based language teaching. More details and free task-based lessons are available at willis-elt.co.uk.

Rhett Yu holds an MA in applied linguistics from the Chinese University of Hong Kong. He is a lecturer in the Department of English at the Hang Seng University of Hong Kong.

Part I
Task-Based Language Teaching

Chapter 1
Task-Based Language Teaching: A Multifaceted Approach

N. P. Sudharshana and Lina Mukhopadhyay

Abstract This chapter is an introduction to Section A on task-based language teaching. The aim here is to present an overview of task-based language teaching discussing its origin, later developments and the current trends. We note here that TBLT emerged as a reaction to the grammar-based synthetic syllabuses that tend to present target language in a piecemeal fashion as structures, functions or a combination of both. The assumption here is that learners will 'synthesize' these elements and build their second/foreign language. The analytic syllabuses, such as procedural syllabus, process syllabus and TBLT, on the other hand, focus on exposing learners to holistic samples of target language and allowing learners to 'infer' grammar rules by engaging them in meaning negotiations. The chapter notes how 'Procedural' and 'Process' syllabuses predate task-based syllabuses and observe some common theoretical rationale and methodological practices. The chapter further discusses how TBLT has emerged as a multifaceted approach over the years drawing on psycholinguistics, second-language acquisition research, cognitive processing and educational philosophy. The final section briefly introduces the other chapters on TBLT in this section.

Keywords Analytic syllabus · Synthetic syllabus · Form focus · TBLT

Task-Based Language Teaching: Origin and Development

Task-based Language Teaching (popularly known as TBLT) is perhaps the most popular and most widely discussed teaching methodology at present. TBLT emerged as a reaction to the erstwhile grammar-based syllabuses which are variously known

N. P. Sudharshana (✉)
Department of Humanities and Social Sciences, IIT Kanpur, Kanpur, India
e-mail: sudh@iitk.ac.in

L. Mukhopadhyay
Department of Training and Development, EFL University, Hyderabad, India

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_1

as audiolingual approach, structural–oral–situational (SOS) method or simply structural approaches. These are referred to as 'synthetic approaches' (Wilkins, 1976) because in these approaches the target language is divided into units (mostly grammatical structures but also vocabulary items sometimes) and is presented to learners in piecemeal fashion in a linear sequence. Learners' role is to synthesize these parts and build up the whole structure of the language. Here the underlying assumption is that acquisition is 'a process of gradual accumulation of the parts' (Wilkins, 1976, p. 2).

One of the earliest attempts to move away from grammar-based syllabus towards meaning-based was 'Notional syllabus' (Wilkins, 1976). Wilkins proposes that instead of grammatical structures, the basic unit of a syllabus be a set of notions (e.g. time, space), modalities (e.g. probability, possibility) and functions (e.g. persuading, agreeing). Also, instead of the linear model of grammar-based syllabuses, he proposes a 'cyclic model' where at each cycle, previously introduced linguistic forms are represented along with a few new forms. However, the 'notional syllabus' also is based on the fragmentation process, just like the grammar-based syllabuses, only difference being structures are replaced with notions-functions.

In a true 'analytic approach', learners are presented with holistic samples of language and learners' role is to analyse the samples and induce rules for themselves. Here the emphasis is on 'natural and authentic representations of target language communication as possible', and learners are engaged in 'genuinely communicative (or at least, meaningful) target language production' (Long, 2015, p. 20). Here the focus is on 'learners' and 'learning process', whereas grammar-based syllabuses are product-oriented; they focus on the final outcomes.

Procedural Syllabus

One of the earliest true analytic approaches was Prabhu's 'Procedural syllabus' (Prabhu, 1984/2019, 1987). Explaining the basis of 'procedural syllabus', Prabhu (1984/2019) states:

> A precondition for genuine deployment is mind-engagement and the resultant need to communicate - a genuine preoccupation with understanding, thinking out, doing or saying something. The focus here for the course-designer is entirely on what to do in the classroom, not on what (piece of language) to teach… I am calling such syllabuses procedural syllabus. (p. 29)

Prabhu is extremely critical of grammar-based structural methods (e.g. structural–oral–situational method). He opines that such methods are built on the assumption that language acquisition is 'a process of successive input-assimilation' wherein the teacher regulates the sequence of inputs and believes 'what is taught = what is (or ought to be) learnt' (Prabhu, 1984/2019, pp. 26–7). However, learners who were taught using SOS method were 'unable to use (i.e. deploy) the language when necessary outside the classroom' or achieve 'an acceptable level of situational appropriacy

in their language use' (p. 16) though they could 'produce' language in classroom contexts. The SOS method focused heavily on grammatical accuracy, yet learners did not achieve 'an acceptable level of grammatical accuracy in their language use outside the classroom' (p. 16).

'Bangalore Project' (1979–1984), a direct reaction to the failures of structural–oral–situational (SOS) method that was very popular in India at that time, is an experiment in implementing 'Procedural syllabus'. In this approach, there was no pre-selection of language items or an emphasis on grammar practice/production in any stage. Instead, the basis of each lesson was a problem or a task which drew on other subject-matters such as mathematical problems. When learners stated something incorrectly, teachers rephrased it with appropriate forms, but there was no whole-class explanation or practice of any kind. The assumption was that while attempting such tasks, learners 'use' language and eventually 'acquire' it. Here, explicit attention to language is 'entirely incidental' or 'reactive', as a specific response to learners' output. See Prabhu (Chap. 2 this volume) for more details.

Process Syllabus

At around the same time, Breen and his colleagues (see e.g. Breen et al., 1979; Breen & Candlin, 1980, 1984) proposed 'Process syllabus'. Its proponents argued that this is 'communicative' in nature with a primary focus on 'the process of learning a language' rather than 'learning a language'. Breen (1984) argues that language learning is not accumulating 'a repertoire of communication'; rather, it is developing 'a capacity for communication' (p. 51). This capacity includes 'applying, reinterpreting, and adapting the knowledge of rules and conventions during communication by means of underlying skills and abilities' (p. 52).

A process syllabus would include data materials (e.g. various kinds of written and auditory texts), information materials (e.g. dictionary) and process materials (that are guidelines for using the above materials) (Breen et al., 1979). Together, a Process Syllabus addresses 'the overall question: 'Who does what with whom, on what subject-matter, with what resources, when, how, and for what learning purpose(s)?' (Breen, 1984, p. 56).

In the process syllabus, the tasks necessarily have two characteristics, viz. differentiated and problem-posing. The former would allow learners to contribute to task completion in different ways at different times depending on their proficiency levels and other factors. The latter would imply that activities are challenging, open-ended and arouse curiosity among learners (Breen et al., 1979). Regarding sequencing, the process syllabus recommends 'a cyclic approach' where 'learners are continually developing related frameworks or aggregations of knowledge and ability use' instead of a linear sequence that is based on 'accumulating separable blocks of 'static' knowledge or a sequence of ordered skills' (Breen & Candlin, 1980, p. 103). Another important feature is that the process syllabus argues for an extremely active

role for learners with an aim of involving 'learner reconstruction and reinterpretation of subject-matter in an explicit way in the day-to-day work of the classroom; to engage learners themselves in the design of their own classroom syllabus' (Breen, 1984, p. 58). However, the process syllabus failed to take off since it lacked a strong psycholinguistic rationale and was largely educational and philosophical (Long & Crookes, 1992).

Task-Based Language Teaching and Learning

Task-based Language Teaching (TBLT), as a method, draws heavily on procedural and process syllabuses (see Ellis, Chap. 4 this volume for more details on origin and development of TBLT). Over the years, many different perspectives on TBLT have emerged. Though they share many psycholinguistic and philosophical underpinnings, and methodological practices, the contentious area has been whether or not to focus on form and if yes, how. In this regard, here we present three main approaches to TBLT, viz. the psycholinguistic perspective of Long and his associates, the cognitive processing perspective of Skehan and his associates, and the explicit form-focused approach of R. Ellis.

Long, along with his associates (e.g. Long, 1985, 2015; Long & Crookes, 1992; Long & Robinson, 1998), has been one of the foremost exponents of a psycholinguistically based approach to TBLT. Long (1985) proposes a four-stage task-based framework: (i) Identifying learners' needs, (ii) Selecting and sequencing tasks, (iii) Actual classroom transaction and (iv) Language assessment. In this model, designing a task-based syllabus would start with a needs analysis through which tasks typically required of an individual in a particular target domain (occupational, vocational or academic) are identified. Based on these 'target tasks', 'pedagogic tasks'—tasks which learners perform in classroom settings as preparation for target tasks—are developed. A more rational approach is adopted for task sequencing as well: task difficulty is determined by various design and implementation factors such as number of elements to be processed, amount of background knowledge presupposed, displacement in space and time, etc. Methodological principles are similar to those of procedural syllabus such as exposing learners to a large amount of comprehensible input through teacher talk and other authentic sources, a focus on communication and not form, tolerance towards errors committed by learners and absence of explicit error correction. Language assessment is supposed to be criterion-referenced examining whether or not students can perform a specific task to pre-decided criterion. In later conceptualisations, Long and his associates (see e.g. Long, 2015, 2016; Long & Robinson, 1998) while adhering to this three-stage model, argue for bringing in form focus in TBLT. They argue that second-language acquisition research, studies in immersion contexts and psycholinguistic studies on the role of consciousness in language acquisition have highlighted the need for conscious attention to forms particularly those which are infrequent, semantically opaque and/or perceptually non-salient. This form focus, however, is different from that in structural methods;

it has to be incidental, implicit, and purely reactive brought about in interactional settings, when learner–learner or learner–teacher engage in negotiation for meaning. In other words, when learners are engaged in problem-solving and are communicating with peers or the teacher, the more proficient speaker may use a variety of interactional adjustments (such as clarification requests, confirmation checks and comprehension checks). These result in reformulations, repetitions, prompts or paraphrases that direct the attention of learners to specific linguistic forms and eventually acquisition. Over the years, different ways of form focus such as input enhancement, input flooding, processing instruction, implicit form-focused tasks, etc. have been proposed (see Sudharshana, 2020 for a detailed review of such options under form focus).

Skehan (1996, 1998) conceptualizes second-language proficiency as consisting of three dimensions—accuracy, complexity and fluency and underscores the need for focusing on accuracy. He argues that communicative tasks by default prioritize meaning over form, but such an approach may not 'automatically drive interlanguage forward' since it may 'proceduralize strategic solutions' (Skehan, 1996, p. 42) without a focus on accuracy. As a result, there is a need to focus on forms explicitly and direct learners' conscious attention towards accuracy 'without losing the values of tasks as realistic communicative motivators, and as opportunities to trigger acquisitional processes' (p. 42). Another factor to keep in mind is that human attentional resources are limited, and one cannot focus on all three aspects of proficiency at the same time. In order to address this, Skehan (1996) proposes a framework for TBLT that consists of three stages: (i) a pre-task stage that pre-emptively introduces linguistic forms, uses consciousness-raising activities, and/or reduces processing load through pre-task planning, (ii) a during-task stage where through careful choice of tasks and manipulation of implementation variables (e.g. ±surprise elements, ±visual support) accuracy and fluency is balanced, and (iii) a post-task stage that redirects attention to complexity and accuracy through processes of analysis, tests, task repetition and parallel tasks. Skehan (1998) argues that task selection and use are governed by a list of task design factors such as nature of information processing (e.g. two-way vs. one-way tasks), task difficulty (e.g. small number of participants vs. large number) and performance goals (fluency, accuracy or complexity). Skehan and his associates have extensively examined over the years how task performance is affected by such design and implementation factors. See Bui and Yu (Chap. 7 this volume) and Nair and Sircar (Chap. 8 this volume) for a detailed review of research on task complexity and implications for teaching–learning.

Over the years, the very definition of 'task' has changed. In procedural syllabus (Prabhu, 1987) and some versions of TBLT (e.g. Long, 1985; Willis & Willis, 2007), the term 'task' was used to refer to a problem-solving activity that closely resembles a real life task in a particular domain ('situational authenticity') or involves deployment of skills and strategies similar to those in real-life contexts ('interactional authenticity'). As a result, explicit focus on form was eschewed. However, in later versions of TBLT a distinction is made between 'unfocused task' and 'a focused task' (e.g. Ellis, 2003, Nunan, 2004). The former is equivalent to the original conception of the task, whereas the latter is one where learners need to use a

particular structure to complete it. Ellis (2003, 2017) has been a strong advocate of focused tasks in the TBLT framework. For instance, Ellis (2017) argues that focused tasks are necessary because they serve many purposes: (i) some features of target language (which are infrequent, non-salient and/or semantically opaque) may be difficult for learners even at advanced levels, (ii) focused tasks can direct learners' attention towards meaning and use of linguistic features and not just form, (iii) focused tasks allow for negative feedback which as shown by studies is more effective than providing comprehensible input, and (iv) focused tasks are useful in language assessment since the outcomes of the task are clearly specified and learner performance can be judged against them. Focused tasks could be either input-based or output-based (Ellis, 2003, 2017). Input-based tasks such as an interpretation task focus on interpretation or 'understanding' rather than production. For instance, learners read a pair of sentences with object pronouns and reflexive pronouns (e.g. John hurt him vs. John hurt himself) and choose the correct picture. In production-based tasks such as a consciousness-raising task, learners process an input text and hypothesize about usage of a particular linguistic feature. Once the learners form hypotheses, they may be presented with additional data to verify the hypotheses to confirm or discard them. For instance, learners read a descriptive text and find patterns regarding distribution and function of present and past participial adjectives (e.g. The movie was boring vs. John was bored). See Upadhaya and Sudharshana (Chap. 10 this volume) for an illustration of consciousness-raising tasks.

Apart from these three main streams of research on TBLT, there have been attempts to illustrate task-based syllabus design with numerous examples and practical suggestions for practising teachers. Prominent among such frameworks are Nunan (1989, 2004), J. Willis (1996), and D. Willis & J. Willis (2007). Nunan (1989) suggests selecting tasks based on learner needs and goals of the programme, and sequencing tasks based on cognitive and processing demands on learners. For instance, a syllabus unit ideally starts with tasks aiming at processing without any overt response (e.g. listening to a text and ticking boxes), then move to production tasks (e.g. read a text and complete an exercise) and finally interactive tasks (e.g. opinion gap activity). Developing on this, Nunan (2004) proposes a six-stage framework (that brings together both meaning and form focus) for a task-based unit: (i) schema building activities, (ii) controlled practice of the target language elements (vocabulary, structures and functions), (iii) intensive listening practice which could incorporate and extend on the target language elements, (iv) controlled practice of the target language elements in a communicative context, (v) free production tasks such as role plays and finally (vi) communicative tasks such as information gap activities. J. Willis (1996), and D. Willis and J. Willis (2007) proposes a predominantly meaning-focus framework where form focus appears in the post-task stage. See Willis (Chap. 5 this volume) for more details.

Overview of Chapters in This Volume

Chapters in this volume represent different streams of research under TBLT. Many chapters have sample tasks in appendices in addition to theoretical and experimental research details. We hope practising teachers find it useful. Each of nine chapters in this section is summarized below.

Chapter 2 'Second Thoughts on Second Language' by N. S. Prabhu outlines theoretical underpinnings for communicative language teaching approaches in general and TBLT in particular. Prabhu notes that in L1 acquisition, children make sense of language through exposure to it bit by bit and by trial and error. Further, the processes in naturalistic L2 acquisition resemble very closely those in L1 acquisition. Following this, classroom L2 acquisition also needs to emphasize on exposure to holistic samples of language and meaning making and not on discrete structures. He argues that L1 interference is actually a result of LSRW procedure followed in most ESL/EFL contexts where learners are under pressure to speak as early as possible. Though this LSRW procedure in L2 classroom is said to mirror those in L2 naturalistic contexts, there is not much emphasis on input through listening and reading. Regarding teaching grammar, Prabhu notes that in TBLT grammar acquisition is seen as an integral part of the comprehension process very similar to mother tongue acquisition; there is no explicit or concealed approach to teaching grammar. He emphasizes on reading and observes that 'it is like nutrition that reduces the need for medication'. If it does not work or in exceptional circumstances, then one needs medication in the form of 'remedial grammar teaching'.

Chapter 3 'The Undiscovered Vygotsky in Prabhu' by Esther Ramani and Michael Joseph argues that though separated by time and space, there are striking convergences and divergences between Vygotsky and Prabhu regarding thoughts on language acquisition and language teaching. The authors observe resemblance between ZPD and reasoning gap activities, and between the concepts of 'internalization' and 'subconscious acquisition'. In addition, both Vygotsky and Prabhu believed in centrality of mind, consciousness; both argued that comprehension precedes production; and that ideation and motivation are linked. The authors also report on the project African Language Group, at the University of Limpopo in South Africa where Prabhu's and Vygotsky's ideas were applied to train, teach and learn South African indigenous languages. The appendix contains classroom transcripts and sample tasks.

Chapter 4 'Task-based Language Teaching: Early Days, Now and into the Future' by Rod Ellis traces the origin and development of TBLT since the 1980s. Ellis discusses how TBLT grew out of communicative language teaching approaches fed by findings from second-language acquisition research and educational philosophy. Noting that TBLT has multitudinous strands, the chapter attempts to define 'a task', classify various kinds of tasks, discuss various methodological practices prevalent within TBLT framework, identify interface between TBLT and other recent developments in methods and techniques such CBI/CILL and CALL and finally address some common misconceptions about/ criticisms levelled against TBLT.

Chapter 5 'An Evolution of a Framework for TBLT: What Trainers and Teachers Need to Know to Help Learners Succeed in Task-based Learning' by Jane Willis explains in detail the framework for TBLT developed by Willis (1996). The chapter begins by outlining major influences of that time such as dissatisfaction with PPP approach, Prabhu's Bangalore Project, assembling of Collins and Birmingham University International Learners Database (COBUILD) and emergence of lexical syllabus based on it, and publication of a course book based on lexical syllabus. Later, the chapter discusses at length each stage in the three-phase TBLT framework justifying pedagogical practices and illustrating with several examples. The chapter ends with implications for teacher training.

Chapter 6 'From Needs Analysis to Task-based Design: Methodology, Assessment and Programme Evaluation' by Roger Gilabert and Aleksandra Malicka focuses on Needs Analysis (NA) and how it can significantly inform decisions related to methodology, assessment and evaluation. The authors note that there has been a significant amount of research on what NA is and how NA is to be conducted, but less on how it can be utilized in task selection, design, sequencing, teaching and assessment. The chapter begins by explaining NA in detail including definition, methods of conducting it, and potential issues and ways to address them. The authors, subsequently, analyse how information collected through multiple methods such as interviews, observations, linguistic analysis of samples in NA can influence methodological principles such as 'learning by doing', pedagogical practices such as pre-emptive focus on form, language assessment at the levels of interactional, psycholinguistic and cognitive demands and finally external programme evaluation.

Chapter 7 'Differentiating task repetition from task rehearsal' by Gavin Bui and Rhett Yu attempts to distinguish between task repetition and task rehearsal based on their functionalities, as the former leads to implicit learning, whereas the latter leads to explicit learning. This chapter begins with an overview of the perceptions on task repetition and rehearsal from conventional and theoretical angles by citing some milestone research by Ellis (2005), Skehan (1998, 2014), Robinson (2001), Levelt (1989) among others. The authors define types of task repetition in terms of reception condition, repetition intervals and frequency of repetition. While discussing problems with the conventional view on task repetition, the authors argue that the existing studies in this area suffer from insufficient information about learners' awareness of their future performance or predictive validity of their performance. They observe that quite often there is an overlap of perceptions between task repetition and task rehearsal in these studies. Consequently, the authors express the need to distinguish between these two constructs considering the role of awareness as a common factor. Task repetition, although beneficial in language learning, does not necessarily lead to anticipation of future learning. On the contrary, during task rehearsal, learners are conscious about the future task and prepare themselves accordingly. The chapter concludes with pedagogical implications by suggesting that task rehearsal could be used with low proficiency language learners whereas task repetition with more proficient learners.

Chapter 8 'Task Complexity and Language Proficiency: Its Effect on L2 Writing Production' by Veena Nair and Shruti Sircar reports on a study set in Skehan and

Robinson's frameworks investigating how task complexity and learner proficiency impact writing performance of ESL learners. There were two groups of learners, one low proficiency and the other high proficiency and they performed on (+complex) and (−complex) tasks. The written production was assessed on various linguistic measures such as syntactic complexity, syntactic variety, lexical density, lexical variety and accuracy (in terms of error-free T units). The results showed a discernible effect of proficiency on the outcome: proficiency levels positively correlated with scores on linguistic measures. There was an interaction between proficiency levels and task complexity: with an increase in task complexity, there was an increase in scores in the high proficiency group but a decrease in the scores of low proficiency learners. Accuracy scores were, however, not affected. These findings validate Robinson's cognition hypothesis with significant implications for teaching of writing. The chapter argues that conceptual and linguistic requirements need to be kept in mind while selecting tasks for classroom teaching. For instance, one may increase complexity as learners move along the interlanguage continuum since complex tasks provide opportunities for pushed output leading to language acquisition.

Chapter 9 'From Cognitive Grammar to Pedagogic Grammar: Macrostrategies for Designing Form-Focused Tasks' by N. P. Sudharshana focuses on integrating meaning and form focus in communicative tasks. The chapter begins with a brief review of form focus in communicative contexts and then underscores the interface between reference grammar and pedagogical grammar. The author argues that Cognitive grammar (CG) is a better source for pedagogical grammar that in turn forms the basis for form-focused tasks: CG is able to offer meaning explanations for most grammatical phenomena, traditionally thought of as unexplainable. The chapter aims to offer a set of macrostrategies to draw more effectively on CG to design contextualized form-focused tasks. It is argued that such strategies help teachers first arrive at a comprehensive set of pedagogical grammar principles which eventually can be exploited in various tasks with a focus on specific forms. The paper illustrates an application of CG to pedagogical grammar in teaching participial adjectives in ESL contexts.

Chapter 10 'Designing and Using Tasks to Foster Metaphoric Competence among Learners in Indian Contexts' by Baburam Upadhaya and N. P. Sudharshana has two large aims: present a framework, drawing on principles of Conceptual Metaphor and TBLT, for designing tasks to foster the development of metaphoric competence among ESL learners, and present results from a small-scale intervention study on using such tasks in an actual classroom. The chapter begins with an introduction to metaphoric expressions and a brief discussion of the concept 'metaphoric competence'. The authors then present an analysis of a few popular textbooks at secondary level in India and note that activities in textbooks are based mostly on the traditional views which treat metaphoric expressions as largely arbitrary linguistic constructs. Such an approach encourages memory-based learning. In contrast, the framework developed in the chapter that draws on the theory of conceptual metaphor (CM) in cognitive semantics, helps identify meaning motivations for such expressions and aims to use form-focused tasks to raise metaphoric awareness of learners. The chapter offers practical suggestions regarding choice of metaphoric expressions,

selecting and sequencing tasks and pedagogical practices in the actual class based on a small-scale intervention study with a group of intermediate-level ESL learners in India.

References

Breen, M. (1984). Processes in syllabus design. In C. Brumfit (Ed.), *General English syllabus design* (pp. 47–60). Pergamon Press.

Breen, M. P., & Candlin, C. (1980). The essentials of a communicative curriculum in language teaching. *Applied Linguistics, 1*(2), 89–112. https://doi.org/10.1093/applin/I.2.89

Breen, M. P., Candlin, C., & Waters, A. (1979). Communicative materials design: Some basic principles. *RELC Journal, 10*(2), 1–13. https://doi.org/10.1177/003368827901000201

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R. (2005). *Planning and task performance in a second language*. John Benjamins

Ellis, R. (2017). Position paper: Moving task-based language teaching forward. *Language Teaching, 50*(4), 507–526. https://doi.org/10.1017/S0261444817000179

Levelt, W. J. M. (1989). *ACL-MIT Press series in natural-language processing: From intention to articulation.* The MIT Press.

Long, M. (2015). *Second language acquisition and task-based language teaching*. Wiley Blackwell.

Long, M. (2016). In defense of tasks and TBLT: nonissues and real issues. *Annual Review of Applied Linguistics, 36*, 5–33. https://doi.org/10.1017/S0267190515000057

Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77–100). Multilingual Matters.

Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly, 26*(1), 27–56. https://www.jstor.org/stable/3587368

Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). Cambridge University Press.

Nunan, D. (1989). *Designing tasks for the communicative classroom.* Cambridge University Press.

Nunan, D. (2004). *Task based language teaching*. Cambridge University Press.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press.

Prabhu, N. S. (2019). Procedural syllabuses. In G. Durairajan (Ed.), *Perceptions of language pedagogy* (pp. 26–34). Orient Blackswan (Reprinted from Read, T. E. (Ed.). (1984). *Trends in language syllabus design* (pp. 272–280). Singapore University Press/RELC).

Robinson, P. (Ed.) (2001). *Cognition and second language instruction*. CambridgeUniversity Press.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*(1), 38–62.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skehan, P. (Ed.) (2014). *Processing perspectives on task performance.* John Benjamins.

Sudharshana, N. P. (2020). *Teaching grammar in the ESL/ EFL context: Changing perspectives, current trends and challenges ahead* [Unpublished manuscript]. Department of Humanities and Social Sciences, IIT Kanpur.

Wilkins, D. A. (1976). *Notional syllabuses.* Oxford University Press.

Willis, D., & Willis, J. (2007). *Doing task-based teaching.* Oxford University Press.

Willis, J. (1996). *A framework for task-based learning*. Longman.

Chapter 2
Second Thoughts About Second-Language Teaching

N. S. Prabhu

L1 – L2

It was thought then that a child's acquisition of its first language or mother tongue was a very different process from the acquisition of a second language. Mother-tongue acquisition was seen to happen naturally, without any deliberate teaching or effort at learning, unfailingly in all cases. No one fails to learn and, if we ignore the literate skills, all seem to learn equally well. Learning to speak happens as naturally as being able to walk. In sharp contrast, a second or subsequent language needs to be taught and typically leads to varied levels of success between different learners. It therefore looks as if first-language learning is a gift of nature while second-language learning depends on the quality and quantity of teaching. That led to teaching methods such as systematic selection and grading in the teaching of words, structures, functions, etc. in schools.

But there are some contexts where a second language too gets learnt without being taught. This happens when schooling happens in a language other than the mother tongue, when children make friends with those next door who happen to speak a different language, when families migrate to a different language area, and so on. There is then either a pressing need or desire or pleasure in understanding what is heard, and achieving it involves making sense of a different language—slowly, bit by bit, by trial and error. Learning is then driven by the learner's effort, not by the teaching effort. Success in such learning is generally varied, but it is clearly and uniformly higher than that achieved from classroom second-language teaching.

That makes one wonder if mother-tongue acquisition too might not be similar: driven by a need to make sense of a new and bewildering world, and doing so slowly, in very small bits, continuously through the waking hours, over the first year or two. If that is so, might it not be that a second language is better taught in the classroom

N. S. Prabhu (✉)
National University of Singapore, Singapore, Singapore

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_2

by getting students to understand something new by making sense of a new language bit by bit with effort? Teaching then is a process of creating classroom situations in which learners try to make sense by making an effort. Problem-solving, which appeals to young people in the way competitive games do, can be a classroom activity which involves effort at comprehending language in order to understand and solve a problem. This is what is called task-based teaching. The process of learning a second language is not, therefore, different from that of learning the mother tongue, though of course less intensive because of less urgency in making sense, hence slower and perhaps falling short of level of mother-tongue acquisition. Teaching a second language then is best done by offering learners something they want or need to make sense of by comprehending the language in the process.

The Four Skills

It was also thought that a second language was best taught by getting learners to practise the four aspects of its use—listening, speaking, reading and writing—cyclically, in each lesson and with each teaching item. L-S-R-W was an elementary rule in second-language teacher training.

The four skills represent two dimensions: comprehension/production on the one hand and speaking/writing on the other. Let us look a bit more closely at the first. Learners were asked first to comprehend a piece of the new language with a few examples and, soon thereafter, asked to produce it in speech. This contrasts sharply with what happens in leaning the mother tongue. To begin with, the new-born child takes a whole year listening and making sense of what is heard, before it attempts to utter the first word around its first birthday, and takes a couple years more before uttering something like a sentence—and it is leaning all the time, not an hour each weekday! Secondly, it is well known that all adults have a much larger passive vocabulary than active vocabulary, i.e. they understand many more words than they use in their own speech or writing, and that includes professional writers and speakers. Furthermore, when memory begins to fade in old age, it becomes difficult to recall the words one wants before it becomes difficult to recall the meanings of known words. This means that comprehension ability by far precedes production, stays consistently larger than production and lasts longer than production. Does it not indicate that there is just one ability which develops as comprehension, rises to the level of production and in the end recedes back to just comprehension?

When a whole family migrates to a different language area, the adults who have to cope immediately in the new language soon develop a pidgin (mixing the new language with the known one), while the children who are under no such pressure end up speaking the new language more or less like the locals. The reason is not that the adults suffer from ‘L1 interference’, but rather that their comprehension in the new language has not yet reached the level of production and they therefore have to cope by drawing on the language they already know. What is called ‘L1 interference’ is in fact a creation of the L-S-R-W procedure.

Sentence and Text

L-S-R-W is practised mostly with sentence-length examples highlighting the relevant words or grammatical features. This can at most be a preparation for conversation and exchange of information which tend to happen through speaking and listening. The other two of the four skills, reading and writing, can also serve such exchange of information, but they can do far more than that when they constitute texts. Texts represent chunks of knowledge and thought, capable of engaging minds and stimulating thought. A piece of text is not just a sequence of sentences; it is an organized cluster of information, thought and logic, representing the intellectual world.

We said earlier that a second language is learnt much better in the process of understanding something which one needs or wants to understand, than when it is taught and learnt as a language and that problem-solving has an appeal to people, especially young people, as an opportunity to feel pleased with success, especially in peer-group competition. Texts offer a great opportunity to do this at a more advanced level. Comprehension questions demanding close reading, interpretation and inference can be a problem-solving activity in the classroom with multiple benefits. They possess the same attraction as competitions and sports, with the prospect of success, pleasure and a sense of achievement. That can naturally lead to more and more attentive reading and interpretation of pieces of text and interconnecting of different facts to make inferences, in the process bringing about a form of intensive exposure to the language, which I believe accelerates and consolidates acquisition of the language structure. As in other forms of problem-solving, each instance of successful effort increases slightly one's confidence as well as competence in making the next effort at a slightly higher level. Such experience of perceiving, interrelating and inferring facts, thoughts and reasoning in pieces of text is likely to be the beginning of intellectual ability. The famous linguist, Michael Halliday, spoke of a 'pragmatic' function and a 'mathetic' function of language, meaning roughly practical/transactional as against intellectual/ideational. While sentence-based learning can prepare for the former, text-based comprehension can lead to the higher-level mathetic ability.

The Role of Grammar

Six decades ago, it was thought that teaching a second language by teaching the rules of grammar was a waste of time and what was needed, instead, was plenty of practice/repetition of short sentences representing different structures. L-S-R-W were the four modes of practicing each piece of structure. Each structure was in fact a grammatical unit, but that was not to be highlighted or drawn attention to. It was to be absorbed by the learner in the course of being practised in the four modes. This therefore was a form of concealed grammar teaching.

What I have said about task-based language teaching by bringing about an effort at comprehension leaves grammar acquisition to occur as an integral part of the

comprehension process. Grammar is not the target, explicit or concealed, any more than in mother-tongue acquisition.

There is, however, a problem. While mother-tongue acquisition goes on intensively for long enough to reach the full level of production, second-language acquisition can fall short of that level when circumstances oblige the learner to produce. That can result in embarrassing errors in speech or writing, in need of repair. I think that situation calls for and justifies teaching remedial grammar, as a way of filling the gaps in acquisition. Such grammar teaching is like medication for illness while effort at comprehension is like nutrition which reduces the need for medication. One further point can be made. The fewer the errors, the greater the effect of remedial grammar; so it is worthwhile using any scope there way be to raise the level of comprehension further—by cultivating a habit of reading what one enjoys, for example—thus minimizing the errors that call for remedial grammar.

Chapter 3
The Undiscovered Vygotsky in Prabhu

Esther Ramani and Michael Joseph

Abstract The ideas of Prabhu and Vygotsky have endured in our intellectual landscape for the last forty years. Having drawn from both of them in various language programmes in South Africa, we are intrigued by the thought of finding convergences and divergences in these two scholars, so separated in time, context and geography. What further compels us to explore this possibility are Prabhu's brief but beguiling references in his books (1987 and 2019a) to Vygotsky's work. Prabhu's The Learner's Effort in the Language Classroom in his second book helped us examine Vygotsky's thoughts on first- and second-language learning, spontaneous and scientific concepts, oracy and literacy, and his profoundly influential ideas of the Zone of Proximal Development and Mediation. We explore how these ideas relate to Prabhu's work in the Bangalore Project (1979–1984). We also comment briefly on the work of Second-Language Acquisition researchers, who use both Prabhu and Vygotsky's ideas in ways quite removed from their original perceptions. We also offer a piece of classroom data from the African Language Group (ALG) Project, in which we applied a task-based pedagogy (the 'strong version' of Communicative Language Teaching) to Zulu, and also an example of a problem-solving task.

Keywords Vygotsky's Zone of Proximal Development · Strong version of Communicative Language Teaching · Cerebration · Bangalore Project · Subconscious acquisition

Introduction

For several years now, we have drawn upon the work of Lev Vygotsky, to make sense of our own teaching and research at school and university. In language projects that we have initiated in South Africa on the teaching and learning of indigenous South African languages, we have been inspired by the ideas of Prabhu. Also in a dual-medium undergraduate degree we developed at the University of Limpopo in

E. Ramani (✉) · M. Joseph
School of Languages and Literatures, Rhodes University, Makhanda, South Africa

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_3

South Africa, we were influenced by the work of both Vygotsky and Prabhu (Ramani & Joseph, 2010; Joseph & Ramani, 2004) but never thought of exploring the relation between them.

In this paper, we seek out convergences between Vygotsky and Prabhu to see if Vygotsky's ideas could deepen our understanding of Prabhu; and to seek Prabhu's unexplored synergy with Vygotsky. Of course, we also hope to find out how Vygotsky and Prabhu differ from each other. We believe that such a search is an absorbing intellectual journey in its own right.

We are also intrigued by the absence of discussion on the relation between Prabhu and Vygotsky in the Second-Language Acquisition (SLA) and Vygotskian-oriented SLA literature, and that cite Prabhu's work only as historical background.

Lantolf (2000:1–26), and Lantolf and Appel (1994: 1–32), leading Vygotskians in SLA studies, criticize 'mainstream' SLA for ignoring Vygotsky. Lantolf is probably unaware that in his 1987 book, Prabhu sees some convergence between his own ideas and Vygotsky's concept of the ZPD (for an exception, see Demetrion, 1997: 2 whom we refer to in section 4). Our own aim is to explore the mutual influence that Prabhu's SLA and Vygotskian studies can have on each other.

In this paper, we will also refer to some of our own pedagogic experiments with both Prabhu's and Vygotsky's ideas in South Africa, in which we used a task-based approach to design second-language pedagogy for training, teaching and learning South African indigenous languages (see Ramani & Joseph, 1997). We used the principles of a task-based syllabus combined with Vygotsky's notion of activity in a dual-medium undergraduate degree that we initiated in 2003 at the University of Limpopo, in the northern-most province of South Africa. A brief description of these projects is given in the Appendix.

Prabhu's Use of Vygotsky's Concepts of Mediation and the Zone of Proximal Development (ZPD)

The first two references to Vygotsky in Prabhu's first book (1987) appear on page 63, under note 3, the third is on page 65, marked as note 12 and the last on page 66 as note 15. All four appear in the Notes section of Chap. 3, simply titled *Teaching.* It is significant that these three references appear in the chapter on Teaching, which may be seen as what teachers do (as opposed to what learners might do) to bring about learning. It is also significant that Prabhu asserts that 'Vygotsky's view of the development of reasoning in children seems to support the conjecture made here on the strength of classroom experience' (Prabhu, 1987: 63), and he proceeds to give the two quotes below, the first from Vygotsky, the second from Frawley and Lantolf (1985: 20–21) interpreting Vygotsky. It is important to state here that both Vygotsky and Prabhu give central importance to the role of the teacher in bringing about learning. There is little doubt that the knowledge or understanding that a teacher (or expert or a superior peer) possesses is crucial to learner development.

Moving on to the references themselves, Prabhu makes his first explicit reference to Vygotsky and the concept of internalization in note 3 on page 63 of his 1987 book. He cites Vygotsky:

> An operation that initially represents an external activity is reconstructed and begins to occur internally ….An interpersonal process is transformed into an intrapersonal one. Vygotsky (1978: 56–7)

This clearly relates to the internalization of thinking and reasoning processes (though the term 'internalization' may also be used to characterize the internal system of language that is common to theories of subconscious acquisition). In the same note 3, Prabhu also cites Frawley and Lantolf (1985: 20–21) who interpret Vygotsky's perception of this phenomenon:

> All human beings as children are initially integrated into the strategic process of reasoning through social interaction, between the self and a more experienced member of a culture, either an adult or an older peer who is capable of strategic reason ….The transition from inter- to intra-psychological reasoning through mediation, as we said earlier, is a dialogic process, a process in which an adult undertakes to direct a child through a task, and where the child provides feedback to the adult, who then makes the necessary adjustments in the kind of direction offered to the child.

These 'necessary adjustments in the kind of direction offered to the child' is a very apt way of defining mediation.

As is well known, Prabhu arrived at the idea that tasks that demand thinking and reasoning (what he called reasoning-gap activities) create the conditions for the maximal engagement and sustained meaning-making required to acquire a second or additional language. In the last reference to Vygotsky in his first book (note 12 on page 65–66), Prabhu acknowledges that Vygotsky's concept of the Zone of Proximal Development (ZPD) 'seems to lend some support' (p. 66) to the principle behind his construction of 'reasoning-gap' tasks to bring about learner effort based on 'the concept of reasonable challenge' (p. 55). This concept is central to Prabhu's approach and 'implies that learners should be not be able to meet the challenge too easily but *should* be able to meet it with some effort.' (p. 56, italics Prabhu's). It is when learners are faced with a reasonable challenge, as Prabhu defines it, that the stage is set for an interaction between the teacher and the learners that leads to learners' success in meeting the challenge with the teacher's support. Prabhu writes,

> Teacher-class negotiation – in the sense of a sequence of exchanges connecting one point to another on a given line of thought and adjustable at any point as it occurs – was thus identified as a classroom procedure which was both feasible and desirable. Opportunity for such negotiation became an important consideration in selecting classroom activities, and it was recognized that negotiation was most likely to take place – and to prove satisfying – when the demand on thinking made by the activity was just above the level which learners could meet without help (1987: 23–24).

It is in this formulation that Prabhu comes closest to Vygotsky's concept of mediation in the ZPD.

Prabhu's cautious acknowledgement of Vygotsky, expressed in the phrase 'seems to lend some support' (Prabhu, 1987: 66) needs some elaboration. It is not accidental

that all Prabhu's quotations of Vygotsky appear in his chapter on 'Teaching'. We think Prabhu saw greater value in Vygotsky's concept of mediation as a strategy for *teaching* (and in particular, of knowledge) than as a theory of language *learning.* As such, it calls for a closer examination of the differences and similarities between Prabhu and Vygotsky on theories of learning, an area we believe is as yet unexplored. Prabhu was naturally interested in any idea related to teacher-assisted cognition by learners—but only as a condition for the subconscious learning of a grammatical system (the process referred to as 'acquisition'). The difference in goals between Prabhu and Vygotsky is that Prabhu was developing a pedagogy for second language (English, through the Communicational Teaching Project), whereas Vygotsky was concerned with the broader goal of human consciousness. So Prabhu's concern, in the Bangalore Project, was not the internalization of knowledge (content) per se, but in using content for unconscious acquisition of the grammatical system of a new language. However, it does seem to us, that Prabhu's use of a pedagogy that uses reasoning processes, but starting with the learners' own cognitive abilities, might be a good way of providing access to scientific knowledge and more generally, for developing rational ways of thinking. Prabhu perceives such cognitive gains as a useful byproduct of task-based teaching, and one more compatible with the focus on content in formal education, favouring thereby the acceptance of innovations such as a task-based approach (see Prabhu, 1987: 4, 52).

Prabhu and Vygotsky's Views on First and Second ('Foreign') Language

Teaching and Learning Theories

While it is clear from our reading that both Prabhu and Vygotsky share similar views about how first (home/native) languages are learnt, their views about second-language learning are profoundly different. For Vygotsky, L1 is naturally learnt in the pre-school phase, or as Vedeler sums up:

> He (the child) required full consciousness (Vygotsky, 1987: 213), as, for example, when learning a foreign language through deliberate study of grammar and vocabulary, something quite different from the spontaneous learning of a mother tongue. (Vedeler, 2015: 340).

In contrast, Prabhu argued, 'Second language learning involves essentially the same processes as first language learning (a hypothesis inherent in the thinking on the project)' (Prabhu, 1987: 6), and for an 'acquisition' approach based on Palmer, whom he cites as saying:

> in learning a second language, we learn without knowing what we are learning", "the utilization of (the adult learner's conscious and focussed attention (on language) militates against the proper functioning of the natural capacities of assimilation,

and in teaching a second language,

> "we must design forms of work in which the student's attention shall be directed towards the subject matter and away from the form in which it is expressed" (Palmer, 1921: 44, 8, 51, cited in Prabhu, 1987: 6).

Prabhu, independently of Vygotsky, worked out the value of teacher-assisted learner cognition, based on an intuition firmed up by classroom practice (Prabhu, 1987:1). As shown above, his views resonate with Vygotsky's concept of the mediation of reasoning processes. However, Vygotsky's own views on 'foreign' or L2 learning appear to be unknown to Prabhu (and to us, until very recently). It is our impression that Vygotsky's views (on second-language learning) are barely alluded to in the Vygotskian SLA literature, even though they are found in his most important book, *Thinking and speech*, in Chap. 6 (1987: 167–242). We have therefore taken the liberty of quoting Vygotsky's brief account of L2 instruction. Vygotsky's aim is to show:

> that the learning of a scientific concept differs from the learning of an everyday concept in much the same way that foreign language learning in school differs from learning a native language. Second we must show that relationship between development of the two types of concepts are much the same as the relationships between the processes of foreign and native language development. (Vygotsky, 1987: 222)

By 'scientific concepts', Vygotsky meant systematized, generalized knowledge, and by 'everyday concepts', the unorganized concepts learners acquire naturally before schooling begins. These everyday concepts are subsequently transformed by the internalization of scientific concepts and become the new and higher everyday concepts of learners. Vygotsky was clearly using L2 instructed learning (as the more familiar experience for his readers) as an analogy to compare scientific and everyday concepts.

Vygotsky rejects the idea that new knowledge (of the world) is acquired through L2 instruction, with these words:

> Finally, learning a new language does not begin with the acquisition of a new orientation to the object world. It is not a repetition of the development process that occurred in the acquisition of the native language. (Vygotsky, 1987: 180).

Instead of 'new knowledge' (which for Prabhu is the same as 'ideation'), Vygotsky argues, matter-of-factly, for grammatical knowledge of the L2. So, for Vygotsky, L1 acquisition in the pre-school phase is naturally acquired through ideation, but L2 as taught in school requires explicit grammatical knowledge.

Referring to the 'agrammatical movement' (Vygotsky, 1987: 205), which seems to have been initiated during his time, Vygotsky counters it by pointing out that learning a 'foreign language' has strengths that knowing one's native language does not have. One of these is that 'in the foreign language, the learner is able to consciously decline and conjugate': i.e. 'determine gender, case, or grammatical form that he applies correctly in a given phrase'. (ibid.: 221). Here, the conscious learning of 'grammar' and its correct application are a requirement. 'From the outset he has conscious awareness of the proper declensions and grammatical modifications.' (ibid.: 221). Uncritical of the formal teaching of grammar at that time in soviet Russia, Vygotsky

assumes that such explicit teaching of grammar would result in 'its correct application' in discourse. He reiterates this point: 'the development of a foreign language begins with conscious awareness and volitional mastery *of language*, and culminates in free spontaneous speech.' (ibid.: 221) (italics ours).

Learner Effort

In trying to make sense of Vygotsky's support for a studial approach to learning the grammar of a foreign language, we find Prabhu's article 'The Learner's Effort in the Language Classroom' (2019c) useful. This article, in our view, offers the clearest theorizing we have come across so far, of the different kinds of learner effort that could be brought into play in a language classroom. We see its value to be in its comprehensive delineation of (i) how different strategies of *teaching* may lead to particular kinds of learner effort and (ii) how learner efforts can be aligned to underlying theories of learning.

Prabhu makes a distinction between two broad categories of learner effort, i.e. what the learner actually does: imitation and cerebration. Imitation involves 'the learner attempting to do, as accurately as possible, what the teacher does in an exemplary role' (2019c: 146). The audiolingual method, aptly nick-named 'mim-mem', is an example of one form of imitation which Prabhu labels as 'reproduction.' 'Simulation' is the other form of imitation associated with role-playing a dialogue. The older form of simulation is the 'situational method', and its modern reincarnation is the (weak form of) the Communicative approach. Prabhu dismisses these forms of imitation as unproductive for language acquisition.

Vygotsky's views on learner effort in the L2 class are meagre. He does refer to the existing resources that the learner brings to L2 instruction, and that interact with it. These resources are equivalent to the spontaneously acquired L1 and the thinking processes acquired from a learner's pre-school days.

Applying Prabhu's categories of learner effort, it is clear that Vygotsky is not talking about 'reproduction' of language forms or 'simulation' (to use Prabhu's terms). Vygotsky, in fact, in his other writings on scientific concepts clearly rejects 'rote learning' and the memorizing of scientific definitions as useless (see Karpov & Haywood, 1998: 29, 31, for a detailed discussion of this). But Vygotsky, without making this very explicit, does uphold cerebration, as implied by the learner's requirement to apply their resources to grammar instruction. Prabhu takes a strong position against all forms of imitation, but does support cerebration. He makes a distinction within the category of cerebration between what he calls 'construction' and 'deployment'. Construction is 'cerebration focused on language itself…and (deployment is) cerebration focused on a knowledge of the world handled through language'. He points out that 'construction' refers to the learner's effort to, 'understand some part of the language system …and to any effort to put together linguistic expressions on the strength of that understanding.' (Prabhu, 2019c: 152).

From our analysis of Vygotsky's approach to L2, we think that Vygotsky is advocating a form of construction. This is unsurprising as the grammar translation approach and similar focus on rational approaches to grammar were in vogue at Vygotsky's time. In contrast to construction, Prabhu favours 'deployment' which is the use of cerebration around knowledge as object. 'Deployment' as an approach originated in the Bangalore Project and is further theorized in his 2019 book. As Prabhu points out: 'Deployment is specifically the effort called for in problem-solving activities'. (2019c: 152), and we add, around ideational knowledge.

Though Prabhu's 'deployment' differs with 'construction', he defends the value of construction as follows:

> … construction has been a part of language pedagogy for a long time, in such approaches as the Grammar-Translation Method and the Cognitive-Code method (ibid.: 153)
>
> One may in rejecting construction, reject all cerebration, as happened more than once when grammar teaching was rejected as a means of language teaching. Such rejection of all cerebration leaves one with only imitation as desirable learner effort, whether in the form of reproduction or simulation…. If, in rejecting this, one rejects both the cerebral effort and the focus on language, one is left with relatively mindless activity…. (ibid.: 153).

In this, both Vygotsky and Prabhu believe in the centrality of mind, consciousness, cerebration and volition in human learning.

Vygotsky's Concept of 'Internalization' and Prabhu's Concept of 'Subconscious Acquisition'

We return here, to what we began in section 2, namely Prabhu's links with Vygotsky's main theory of consciousness. Vygotsky's concept of internalization is central to his theory of consciousness ('higher mental processes'), and here we examine how it relates to Prabhu's concept of 'subconscious acquisition'.

Since most SLA and TBLT scholars have been concerned with the concept of subconscious acquisition, it is interesting to note criticisms of this concept. Demetrion (1997: 5) for instance is suspicious of Prabhu's view of learning 'subconsciously' as a 'somewhat mystical internalization process unchallenged in other second language acquisition research'. The suspicion of 'mysticism' attributed to Prabhu's views on 'subconscious acquisition' is also echoed by some Vygotskian scholars who prefer to replace (or reinterpret) Vygotsky's concept of internalization with that of overt mastery (an example of this is Rogoff, 1990: 165–66).

Vygotsky on the other hand pointed out: 'Our hypothesis establishes the unity but not the identity of learning processes and internal developmental processes. It presupposes that one is converted into the other.' (Vygotsky, 1978: 90–91). Vygotsky's distinction between external (learning processes) and internal (development processes) does not mean that this distinction is a 'barrier' (as Rogoff, 1990:195–196 sees it). Vygotsky constantly pointed out that internalization operates by different laws from learning and that 'as yet, the barest outline of this

process is known'. (Vygotsky, 1978: 57). His view that 'the external form of activity'…'continues to exist for a long time'…'before definitively turning inwards' (ibid.: 57) is very similar to Prabhu's 'recurrent deployment in comprehension'…'can firm up the internal system.' (Prabhu, 1987: 80).

In this respect, Vygotsky's view of internalization strikingly resembles Prabhu's views on subconscious acquisition, in which he claimed that 'learning-centred methodology is perhaps largely a matter of coping with the unknown in language pedagogy.' (Prabhu 2019b: 57). Despite these convergences, however, there is a fundamental difference between the two. For Prabhu, 'subconscious acquisition', as for all SLA scholars, refers to a biologically innate language capacity. The biological domain has not been Vygotsky's focus. We agree with Lantolf and Appel (1994: 5) and Wertsch (1985: 197) that this is a limitation of Vygotsky. Vygotsky talks about the 'internal' plane being 'created' or formed, whereas Prabhu talks about a subconscious capacity being 'triggered off', and also developed unconsciously, with external dialogue and its interiorization, as a favourable condition.

Apart from the issue of whether language is acquired or learnt, it is worth also comparing Prabhu's views of learning in general with Vygotsky's. As we have already noted, Vygotsky, like Prabhu, believed that teaching and learning are different phenomena, and that learning as a mental phenomenon (including Vygotsky's inner speech and thought) is not yet known, and not caused directly by teaching. But compared to Vygotsky, Prabhu suggests a much more indirect link between teaching and learning. For Prabhu, unconscious learning is a.

> 'happening', 'an accident inherently …subject at most to a degree of probability', similar to 'a thought, idea or image arising in our minds. We cannot make an idea arise, or tell when or whether it will arise; we can only recognize it after it arises. (Prabhu, 2019e: 245–246). (It is also) 'completely inaccessible to us' (ibid.: 245)

As with Prabhu, so also with Vygotsky, learning as a mental phenomenon is a topic for rational speculation, and not open to empirical investigation. Prabhu's ideas about the unpredictability of learning, while a plausible theory, is likely to make teachers feel vulnerable, to a greater extent than Vygotsky's theory. But it is a vulnerability Prabhu confronts the teacher and language specialists with, as he does in his paper *Coping with the Unknown in Language Pedagogy* (2019b). We believe that Vygotsky's and Prabhu's theories of learning can help to lower the anxiety that teachers normally experience when they demand too perfect a match between teaching and learning. A focus on learning would enable teachers to observe and value those episodic moments where genuine learning seems to happen, which in turn could help make their teaching efforts more responsive to learning processes. Vygotsky's concept of mediation was meant to support learning processes.

Comprehension and Production

Despite the differences between Prabhu and Vygotsky on second-language pedagogy, we find that both argue that comprehension precedes production. For Prabhu, a long silent period in which learners make an effort to comprehend 'input' would be followed in time with spontaneous, voluntary production of the target language. He points out the reason for delayed production as follows: 'abstract structures' (that have already been formed as a result of earlier comprehension) 'need to be formed more firmly for deployment in production than they need to be for comprehension.' (Prabhu, 1987: 71).

Vygotsky holds a similar view for first-language and multilingual learning. Commenting on multilingual children, he affirms his position of the necessity of '… understanding and reading or passively using many languages …' but extends this position to ontogenesis:

> The prevalence of passive speech over active is retained over the whole course of childhood. The child learns to understand speech sooner than he learns to communicate. (Vygotsky, 1998: 273).

In establishing the precedence of comprehension over production, Vygotsky goes to the roots of how human consciousness emerges from two separate 'lines': practical thinking from tool making and use; and language for emotive and phatic purposes ('primitive functions' for him). According to Vygotsky, at a certain point in human history, and in the history of the individual human being, the two lines converge giving rise to what he called 'verbal thinking' (a view more fully developed in chapter 4, 'Genetic roots of Thinking and Speech' in *Thinking and Speech*, 1987:101–120). His research showed that approximately at the age of 11 months, the convergence between the two lines happens through the 'pointing gesture,' which we take up in section 6 on Ideation. But in the months preceding the emergence of the pointing gesture and in the year that follows this first preverbal but object-oriented gesture, the child is listening and observing. The silent period is therefore not an inactive period, but one of effort to comprehend adult language in relation to the child's action on objects.

Prabhu, however, extends the role of comprehension to all languages and to the whole of human life (not just during childhood), for as he says,

> As adults, we all comprehend a lot more language in our listening and reading than we produce in our speech and writing, in the mother-tongue as well as in any second language we know. … This means that comprehension not only precedes production in language learning but remains larger in scope than production through one's life and lasts longer than production at the end. (Prabhu, 2019d: 325).

We turn briefly to how these views of Prabhu and Vygotsky could be related to our second-language learning projects (see Ramani & Joseph, 1997: 17). In these projects, we operated with the view that full comprehension of teacher input was not a requirement; learners needed to comprehend just enough to complete the task.

Prabhu believes that beginning learners need much more to struggle to comprehend for a purpose, than to focus on language form of any kind. Deepening comprehension is for him, as for us, preferable to premature production. Prabhu believes that voluntary production will occur at some future date. Geetha Dorairaj's interview with Prabhu (see Dorairaj, ed. 2019: 324–25) reminds us of the shock and intrigue adult learners (in the ALG and the NSL projects, see Appendix) voiced at our faith that production will happen. Dorairaj (2019: 325) asks Prabhu:

> So …we must just wait for students to begin speaking the language, without knowing when it is likely to happen. Isn't this something of a leap of faith, waiting for something without knowing when to expect it?

to which Prabhu replied,

> The leap of faith is therefore much smaller if you see comprehension as an earlier stage of production. (Prabhu, 2019d: 325).

In our own transcript of an adult beginner lesson, we found learners wanted to engage in production well before they could comprehend, as well as to focus on second language form (in this case the spelling of the Zulu word 'igedlela'; see the transcript in the Appendix). The cerebral effort the learners made was largely guess work, to start with, and elimination based on earlier unsuccessful efforts. But our emphasis (as project leaders and learners) was less on the linguistic stretching, and more on the cognitive stretching as well as exposure to the teacher's use of the target language. As we have pointed out earlier, both Prabhu and Vygotsky argue for the strategic processes of reasoning being internalized (in Vygotsky's general theory of consciousness). For us, spelling was only used as a problem to bring about thinking through interaction with the target language. Such interaction is seen as a favourable condition for acquisition.

This is not to say that no production occurred. Greetings, leave takings and other ritualistic uses of the target language were picked up and used enthusiastically. More noticeable was the appropriate use of phrases like 'Is this right?', or 'yes, come' (to the board', 'write', 'read', 'that's right', or 'no' 'try' (all recurrent uses of classroom language related to instruction). They were picked up by learners and used spontaneously, and with much relish. Also, longer stretches of target language use occurred more readily in writing than in speech, as the project progressed.

We also preferred our university students, for whom English was a second language, and who struggled to express themselves in it, to deepen comprehension of academic texts in our three-year dual-medium undergraduate degree (see project 6 in the Appendix). These undergraduate students enabled themselves to comprehend academic texts in English (prescribed or self-selected from the Internet) and in peer groups, used their mother-tongue through self-initiated translation to facilitate comprehension (for a description of such classes, see Hornberger, 2010). Students' academic English at the end of the term surprised us by their originality, as well as their fairly sophisticated discourse. Similar findings emerged from Amritavalli's (2017) research of students in Indian universities, in which she contrasted initial errorful passages produced by learners with their final almost error-free academic discourse.

Ideation in Prabhu and Vygotsky

In this section, we take up Prabhu's concept of 'ideation' (based on Halliday's concept: see Prabhu 2019g: 268) in second-language pedagogy and compare it with Vygotsky's mainstream ideas on 'object-oriented' knowledge (or 'scientific' knowledge) in formal schooling. The convergence of their views on objective knowledge is striking.

It is significant that Prabhu did not start with a prior preference for ideational knowledge but discovered it through classroom experiments in the Bangalore Project. Through trial and error, he found that learners preferred more factual (information-gap and reasoning-gap) tasks to opinion-gap tasks, and that 'a preference for reasoning-gap activity developed gradually' Prabhu (1987: 47). Among the several reasons he gave for discontinuing opinion-gap activities, the one that struck us most was that they made too great a demand on the learners' interlanguage, or from a Vygotskian perspective, pushed learners beyond their ZPD. Information-gap activities made too little (cognitive) demand and would be therefore below their ZPD.

Prabhu also argued that innovations in language pedagogy should not depart too radically from formal educational practice, and that the ideational knowledge he used in his task-based approach was closer to the curricula in formal education in India. We find support for Prabhu's argument in Howatt who showed that Michael West initiated a reading approach to texts in English, (instead of oral English) in 1919 in Bengal (Howatt, 2004: 278–283). Howatt went further to give ideation an international status by pointing out that (written) texts were favoured by educational institutions the world over (ibid.: 347), and finally that 'Palmer's association of spontaneous/studial learning with the contrast between speech and writing' (Howatt, 1984: 241) has far-reaching implications for the language teaching profession.

Interestingly, Vygotsky, like Palmer, links a conscious awareness of the rules of language with the development of writing and progress in literacy. In his own words,

> When a child learns to write, he begins to do with volition what he has done without volition in the domain of oral speech. Thus, both grammar and writing provide the child with the potential of moving to a higher level in speech development (1987: 206).

The source of ideation for Vygotsky is his hypothesis that the separate lines of thinking (that arise from tool use) and that of language (for phatic and emotional purposes) common among apes, converge in human beings to form 'verbal thinking'. In ontogenesis, i.e. the history of the child, this convergence begins to occur in the 'pointing gesture' that emerges at about 11 months, when:

> The child stretches his out his hand towards a distant object and cannot reach it, but his hand remains stretched out toward the object. Here we have a pointing gesture, with the objective meaning of a word.' (But) 'It can effect only the people nearby. ….The mother hands the child the object; for her rather than the child, the unsuccessful grasping movement is converted into a pointing gesture……Only significantly later does the action become a pointing gesture for the self, that is a conscious and deliberate action of the child himself. (Vygotsky, 1998:171–172).

Despite controversies around the pointing gesture in modern research, Vygotsky's account (of the origins of ideation, and of the adult's use of language to mediate the child's own efforts based on practical thinking) remains convincing. Vygotsky's theory is radical because he shows that ideational knowledge is specific to human beings. Vygotsky expanded the scope of ideational knowledge considerably by showing that at different stages in the development of the child, there were different knowledge goals. One of his most far-sighted arguments for ideation was children's make-believe play (also called pretend play, or fantasy play), through which the child around the age of five years 'assimilated' (to use Piaget's term) the real world of human behaviour to its wishes. Vygotskian scholars such as Gonsalez et al. (2005) have further expanded the scope of ideation by including everyday local knowledge as 'funds of knowledge' in the design of university curricula (see Joseph and Ramani, 2011; Joseph et al., 2014 for a curriculum that converged with these authors' view).

Motivation

In Prabhu and Vygotsky, ideation and motivation are always linked. Prabhu and Vygotsky see motivation in all human beings, including learners, as always involving ideation. Prabhu uses the phrase 'the rage to know' (Judson, 1985, cited by Prabhu, 2019g: 265) and 'a natural desire to solve the problem' (Prabhu, 2019d: 315).

As early as 1987, Prabhu identifies learner motivation with ideation, and rational processes of thinking. Prabhu's view of ideation (in contrast to subjective forms of knowledge) was viewed with suspicion on the grounds that it was devoid of emotion, too dull, and treated 'learners as mere reasoning machines.' Prabhu (1987: 520). For a recent similar criticism of Prabhu, see Lambert (2018). Prabhu countered this criticism by saying that he was pleasantly surprised by the learners' enthusiasm. In Prabhu's own words, 'learners' involvement and interest were, in fact, the features most noticed by observers in project classrooms' (Prabhu, 1987: 52). Howatt, too remarked on how struck he was by the energetic response to tasks on the blackboard (Prabhu, 2019f: 333).

Vygotsky's research into early childhood bears out his theory of motivation. Towards the end of his most well-known work *Thinking and Speech*, Vygotsky hints that thinking and speech are not the last word in the analysis of consciousness but that behind both lies motivation. He turned to the work of the seventeenth-century philosopher, Spinoza, whom he had read in his adolescent days. In his research into children's play, he writes, almost cryptically of Spinoza's view that an emotion can only be overcome by a higher emotion, mediated by reason (Vygotsky, 1985/1933: 549). This theory of motivation has been extended pedagogically by Paley (2004) *A Child's Work: The Importance of Fantasy Play*. Currently in Australia, Veresov (2017: 47–70) and March and Fleer (2017: 124) are developing Vygotsky's concept of 'perezhivanie' (a Russian term used by Vygotsky, 2019) as a 'moment of drama.' Vygotsky himself defines the term thus, 'The child's *perezhivanie* (is) how the child is aware of, interprets and affectively relates to a certain event.' (2019b: 71).

Prabhu uses the term 'Eureka' moments in his later writings, when he reminisces on his own teaching on the project,

> ….a smallish boy used to get so excited on thinking up an answer that he compulsively walked to the blackboard and (since he could not reach the blackboard) raised his arms for me to lift him up. I began to think that learning happens in something like "Eureka" moments which are exciting. When later I was in Edinburgh writing Second Language Pedagogy and came across Vygotsky's concepts of learning in small steps, each involving a successful effort, it seemed like an articulation of my feeling and a form of theoretical support to task-based-teaching. (Prabhu, 2019d: 321, 333).

These 'Eureka moments' remind us of Veresov's interpretation of Vygotsky's term 'perezhivanie' that we quoted earlier.

We ourselves were motivated to continue teaching using tasks, especially when we found that even unsuccessful efforts did not dampen learners' enthusiasm, and the eagerness with which their hands went up was a continuous source of encouragement. One of the authors of this paper (Esther) recalls that during her teaching on the Bangalore Project, the learners' excitement contributed very much to her own motivation to give them even more challenging tasks, and her delight when the learners succeeded. As learners in the African language projects, we too experienced elation when we solved a particularly difficult task. This was also the experience of the teacher, Nompumelelo Frans, who we worked with (in the Biliteracy Project, number 7 in the Appendix). Ms Frans was surprised and pleased that a learner from her Grade 3 class approached her after she had solved a reasoning-gap task on her own, and asked if she could be given a 'more difficult' task.

We saw also this in our teaching of Vygotsky's ideas in a university course. An undergraduate student in her third year, researching private speech among African children, found that her four-year-old subject produced private speech in his mother tongue, Sesotho sa Leboa, while engaged in a building block activity. At the end of his play, he screamed with joy, jumping on the chair, saying: 'I have made a house for the cat' and proceeded to dance and sing (fortunately, we have a video of this moment; a transcript of it can be found in Joseph & Ramani, 2011: 300). We could understand, from such experiences, what Vygotsky meant when he said:

> In play the child is always above his average age, above his daily behaviour; in play it is as though he were a head taller than himself (Vygotsky, 1985/1933: 552).

Again, to use our Biliteracy Project, a learner playing the role of a monster came out of this role to correct another player who had missed a vital turn, and through gesture and gaze prompted the player to re-enact the role by including the missing turn. Such re-enacted scenes in a learner-directed drama based on the teacher's reading of a story gripped the audience, who were none other than fellow classmates. Veresov's notion of 'a moment of drama' enabled us to see more clearly and theoretically, Ms Frans' intuition that dramatized versions of stories created in learners the motivation to read (Joseph et al., 2017: 201). Vygotsky wrote more extensively of motivation in relation to the child's imagination as it develops from mediated play in children to self-mediated fantasy in adolescence (Vygotsky, 1998: 151–166). He makes the generalization, 'Not only artistic works are produced with

the help of fantasy, but also all scientific inventions and technical constructions.' (ibid.: 165). This perspective would include Prabhu's concept of motivation related to logical thinking in closed-ended tasks. Howatt best summarizes the strengths of task-based learning and teaching:

> If Corder is right in saying that 'given motivation', it is inevitable that a human being will learn a second language if he is exposed to the "data", the eventual outcome of the Bangalore project should show not only that it can be done, but that it can be done with the simplest means. (Howatt, 1984: 288)

The key phrase here is 'given motivation.' Prabhu discovered that task-based learning created this motivation.

Conclusion

The Bangalore Project continues to inspire us as a model of small-scale implementation, which by its openness, has led to professional dialogue and debate around the strong version of CLT. We believe, like Prabhu, that though an innovation may arise in a very specific setting, there is value in pushing its principles in other contexts. This belief emboldened us when we came from India to South Africa to apply these principles in a new country to an adult group, with ourselves as learners exploring our own beliefs in language acquisition through a task-based approach.

Howatt's historical perspective and his perseverance to support debate between the 'weak' and 'strong' versions of communication have further strengthened our convictions, for in a recent paper, he points out sharply:

> Teachers need to know whether they are ……teaching language for communication at some later time (and this 'weaker version' has remained the pedagogical norm), or whether they are committed to a view, like Prabhu's, which gives communication a much 'stronger' role in language acquisition in the sense that effective communicative experiences are needed for the successful extension of communicative competence — 'communicating to learn' and not just 'learning to communicate. (Howatt & Smith, 2014:91)

We see our experiments in learning Zulu and Sesotho sa Leboa as contributing to the enrichment of African language pedagogy, sadly trapped in obsolete methods. As more non-African language speaking people realize the value of learning African languages as additional languages, there is need for continuous innovation.

For speakers of African languages as L1, the growing realization that their languages can be used for ideation and for university-level academic literacy is also a powerful impetus for developing curricula that are relevant and cognitively challenging. This is what Alexander meant when he called for the 'intellectualisation of African languages' (Alexander, 2005 cited by Kaschula & Maseko, 2017: 20). Our dual-medium undergraduate degree at the University of Limpopo has shown the value of using Prabhu and Vygotsky's ideas in courses that aim to reveal the strong relation between language, cognition and the development of human consciousness.

Finally, like Prabhu, we too have upheld the principle of teacher autonomy as a space for a teacher's sense of plausibility (1987: 104; 2019d: 331–340) to grow.

We have likewise argued for a role for language academics to make their theories available through demonstrations of classroom teaching, leading to dialogue and debate, and possible revision of perceptions of language pedagogy.

Appendix

The projects in which we used Prabhu and Vygotsky's ideas are listed below:

1. **The Communicational Teaching of Adults Project (CTAP)**: Johannesburg, February-November 1995 (voluntary teaching of English to 25–30 predominantly Zulu-speaking adult learners, aged 22–60, mostly domestic and industrial workers, two-hour sessions two evenings a week, totally 120 h of teaching). In this adult literacy project, we replicated NS Prabhu's Communicational Teaching Project, using task-based materials, predominantly reasoning-gap activities, on adult themes such as shopping, bus and class timetables, bank rules, home loan applications and maps. Comics and stories, of interest to adults, provided for the acquisition of English through meaningful engagement with extended texts.
2. **BA Honours in Applied English Language Studies**: Wits University, Johannesburg, 1992–1997: a unit on Prabhu's ideas and the Communicational Teaching Project in the English Methodology module, including demonstrations of task-based teaching.
3. **The African Language Group (ALG) Project**: Johannesburg, October-1995 to mid-1997: training by MJ and ER of three Zulu-speaking teachers in the CTP methodology to teach Zulu to 30 largely white English-speaking monolinguals and English-Afrikaans bilinguals. Almost all of these adult learners were themselves language and literacy teachers, applied linguists, university academics and practitioners in Non-Governmental Organizations (NGOs). About half of them had tried to learn Zulu before but had very little competence in using Zulu and found earlier grammar and conversation-based courses too difficult, impoverished and demotivating.

A weekly one-hour training session for the Zulu teachers was used to plan tasks, view videos of previous lessons, look at the feedback sheets that learners filled in at every lesson, and sort out any difficulties that the teachers had. The actual lessons took place in one two-hour session per week, totalling about 75 lessons in all. Every eighth lesson was a reflective workshop, in which teachers, trainers and learners clarify issues of methodology, suggest new themes for tasks, express disagreement or scepticism. Learners were also keen to have cassette recordings of texts and dialogues, but the teachers always included tasks based on these texts. In the ALG Project, both of us (ER and MJ) were trainers, learners and researchers, giving us new perspectives on what it means to learn a totally new language (Zulu) through a task-based approach as well as training teachers in the methodology of task-based teaching for a language we do not know.

At this time, 25 years later, we do not have access to any videos we made of the lessons, but we did come across a transcript of a lesson taught during the early days of the project. It is not a typical lesson, but illustrates very well, some of the principles of our project,

Some of these will be discussed in the analysis of this transcript given below.

Turn	Speaker/actor	Utterance/action	English translation
1	T	Anthony, thatha igedlela ulibeke kutafula Anthony, thatha igedlela ulibeke kutafula	Anthony, take the kettle and put it on the table (T repeats this instruction)
2	L1	*Anthony takes the kettle and puts it on the table*	
3	T	Ngiyabonga. Igedlela. Carol, bhala igedlela. Igedlela	Thank you. **Igedlela**. Carol. Write **igedlela. Igedlela**
4	L2	*Carol comes to the board.* Please repeat	
5	T	I-ge-dle-la	
6	L2	*Carol writes* 'ikedlela'	
7	T	Ubhale kahle?	Has she written it correctly?
8	LL (chorus)	Yebo!	Yes!
9	T	Yebo?	Yes?
10	L3	Cha!	No!
11	T	Woza. Lungisa, Graham	Come. Correct it, Graham
12	LL (chorus)	Woza!	
13	L3	*Graham comes to the board, erases* 'dl' *and writes* 'tl'	
14	T	Ubhale kahle, uGraham. Peter, ubhale kahle?	Has Graham written it correctly? Peter, has he written it correctly?
15	LL (chorus)	Cha! Cha!	No! No!
16	T	Lungisa, lungisa!	Correct it! Correct it!
17	LL (chorus)	Lungisa, lungisa!	Correct it! Correct it!
18	T	Yebo, lungisa!	Yes, correct it!
19	L4	*Peter comes to the board, erases* 'tl' *and writes* 'hl'	
20	T	Ubhale kahle, uPeter?	Has Peter written it correctly?
21	LL (chorus)	Yebo! Yebo!	Yes! Yes!
22	L4	I think so	

(continued)

(continued)

Turn	Speaker/actor	Utterance/action	English translation
23	T	Awu, iketlela, hayi! *(Laughs)* IsiSotho lesi, la eklasini isikhuluma isiZulu, hayi isiSotho iketlela. Bhala isiZulu, i- ge-dle- la (*Stresses each syllable*) Lungisa, Esther	Oh no, iketlela, no! (*Laughs*) That is Sotho, in this class we are speaking Zulu, not Sotho iketlela. Write in Zulu, i-ge-dle-la (*Stresses each syllable*). Correct it, Esther
24	L5	Esther comes to the board and writes 'ikedlela'	
25	T	uCarol ubhale kahle la (*pointing to the board).* Kuhle, kodwa kunegama elilodwa eliy one elingekho right. Kuhle, kodwa kunegama elilodwa eliy one elingekho right. Ubani? Igedlela, igedlela	Carol has already written that *(pointing to the board).* Okay, except for one letter, it is right Okay, except for one letter, it is right. Who? **Igedlela, igedlela**
26	LL (chorus)	Igedlela!	
27	T	Yebo, wozolungisa, Sarah	Yes, come correct it, Sarah
28	L6	Is this still wrong?	
29	L6	Sarah comes to the board, erases 'k' and writes 'g'	
30	T	Yebo, ubhale kahle manje. Igedlela	Yes, she has written it correctly now. **Igedlela**
31	LL	(not clear)	
32	T	Nge-English igedlela yini?	In English what is **igedlela?**
33	LL (chorus)	Kettle, kettle!	
34	T	Yebo!	Yes!

Analysis of this transcript

We comment here on only some significant aspects of this interaction. All of the learners (including us, ER and MJ) were quite unexpectedly preoccupied with writing in Zulu (both requesting the teachers to write their oral instructions on the board, and ourselves writing these instructions in our notebooks). We realized that this was actually a strategy to facilitate comprehension: when listening to a new language, we have no clue about where words end and new ones begin! By seeing the word boundaries in Zulu, we could work out the meanings of words and their relation to each other! This way of arriving at meaning, by making the teacher's instructions permanent through writing (and something we could look at later at home) seemed to be very successful. The problem with this approach was that as learners we got

very preoccupied with the spellings of words and their pronunciation, but as adult literate learners, we realized that literacy played a crucial role in our comprehension of oral instructions.

The teacher, in this lesson, decided quite spontaneously, to engage the learners in a spelling task but to conduct (as was the principle in the ALG Project) the entire interaction in Zulu. It can be seen from the 34 turns in this interaction, that it was not easy. We learners spelt the word **igedlela** as we heard it, and through a process of trial and error, involving six different learners, finally arrived at the correct spelling.

While it may be tempting to view this as a 'form-focused' interaction, in the sense that the task had to do with working out a sound-symbol correlation, it involved an effort to understand the teacher's instructions and carry them out at the board. Each error made by a learner led to us working out through elimination and inference, what the right spelling could be. This is what made it a problem-solving task.

Other aspects worth commenting on are the spontaneous use of much classroom management talk by the learners in Zulu: yebo (yes), cha (no), woza (come), lungisa (correct), the use of English by L6 at turn 28 and the teacher's unfailing use of Zulu and only Zulu.

In contrast to the simplicity of this early lesson, given below is a task sheet from a lesson a year later. Learners were given a real-life newspaper advertisement in English for the sale of houses (such material being unavailable in Zulu at the time) and were required to respond to written questions in Zulu. Thners oere was NO expectation that the answers had to be in complete sentences, as the aim of the task was for the learners to make an effort to comprehend the questions.

ELKANA *PROPERTIES*
c/o Duvenhage & Okkerneut St., Edleen
Tel: 975-9520/1 Fax: 970-1630
UNDER NEW MANAGEMENT! TRY US. WE ARE THE BEST!

OUTSTANDING VALUE IN EDLEEN **R 210,000** Lounge, family room, dining room, 2 baths, garage, beautiful wooden kitchen, sparkling pool with built-in braai **Contact CORA** **083 271 2074**	**NORKEM PARK TOWN HOUSE** **R 137,000** Walking distance to shops and schools, 3 beds, lounge, 1 bath, open plan kitchen, garage, community pool, good security **Contact CORA** **083 271 2074**	**KEMPTON PARK WEST** **R 125,000** 5 beds, 2 lounges, TV room, dining room, bathroom, 4 carports, brand new pool To find out more about this house, contact **CHRIS** **082 4589323**

Umsebenzi Obhalwayo (Written work) English translations were NOT provided in the original task sheet

i. Zingakhi izindlu ezithengiswayo kule-advert? (How many houses are for sale in this advert)?
ii. Ngezayiphi inkampani lezizindlu ezithengiswayo? (Which company is selling these houses)?

iii. Iyiphi indlu oyibona ishibhe ukudlula ezinye? (Which house is the cheapest)?
iv. Iyiphi indlu ohamba ibanga elincane uma uya esitolo? (Which house is close to the shops)?
v. Uma ufuna ukuthenga iTownHouse ungashayela bani ucingo? (if you want to buy a townhouse who can you call)?

4. **The Northern Sotho Language Project (NSLP)**: University of Limpopo, January–November 1999. When we moved to the University of Limpopo, we replicated the ALG Project using Northern Sotho (officially now called Sesotho sa Leboa), a dominant language of the Northern (now Limpopo) Province. Again we trained three Masters students to teach Northern Sotho to a group of 15 academic staff, mostly white, but also a couple of academics from other African countries. Many of the ALG materials were translated from Zulu to Northern Sotho, and once again, a task-based approach was used.
5. **BA Honours in English Studies**: University of Limpopo, 1998–2002, a unit on Prabhu's ideas and the Communicational Teaching Project in the English Methods module, including demonstrations of task-based teaching.
6. **BA in Contemporary English and Multilingual Studies (BA CEMS)**: University of Limpopo, 2003 to date. This was the first time we taught Vygotsky's ideas in a third-year module on *Language and Cognition* in a three-year dual-medium undergraduate degree (taught through both English and Northern Sotho).
7. **Biliteracy research project at school level**: CM Vellem School, Joza, Makhanda (Grahamstown) 2014–2019. This was a collaborative, community engagement project in which we worked closely with a primary school teacher developing task-based materials in both English and Xhosa (the dominant language of the Eastern Cape). This primary school teacher-cum-researcher, Nompumelelo Frans, went on to complete her MA studies, drawing on both Vygotsky's and Prabhu's work, at Rhodes University, under our supervision.

References

Alexander, N. (2005). *The intellectualisation of African languages: The African Academy of Languages and implementation of the Language Plan of Action for Africa.* PRAESA, South Africa.

Amritavalli, R. (2017). Second language acquisition. *Language and Language Teaching, 6*(1, 11), 59–66.

Demetrion, G. (1997). Communicative competence and second language teaching: Lessons learned from the Bangalore Project: A review of N.S. Prabhu (1987). *Second language pedagogy*. Oxford University Press. http://en.coplan.ca/library/research/george/prabhu/prabhu.pdf

Dorairaj, G. (2019). Rethinking language pedagogy: An interview. In G. Durairaj (Ed.), *Perceptions of language pedagogy* (pp. 314–330). Orient Blackswan.

Frawley, W., & Lantolf, J. (1985). Second language discourse: A Vygotskian perspective. *Applied Linguistics, 6*(1), 19–44.

Gonzales, N., Moll, L. C., & Amanti, C. (2005). *Funds of knowledge: Theorising practices in households, communities and classrooms.* Lawrence Erlbaum Associates.

Hornberger, N. (2010). Language and education: A Limpopo Lens. In N. H. Hornberger & S. L. McKay (Eds.), *Sociolinguistics and Language Education* (pp. 549–564). Multilingual Matters.
Howatt, A. P. R. (1984). *A history of English language teaching*. Oxford University Press.
Howatt, A. P. R. (2004). *A history of English language teaching* (2nd ed. with H. G. Widdowson). Oxford University Press.
Howatt, A. P. R., & Smith, R. (2014). The history of teaching English as a foreign language, from a British and European perspective. *Language & History, 57*(1), 75–95. https://doi.org/10.1179/1759753614Z.00000000028.
Joseph, M., Frans, N., & Ramani, E. (2017). Teacher development: Promoting a biliteracy approach to epistemic access. In R. Kaschula, P. Maseko, & H. E. Wolff (Eds.), *Multilingualism and intercultural communication: A South African perspective* (pp. 194–210). Wits University Press.
Joseph, M., & Ramani, E. (2011). Researching one's way into Vygotsky. *Education as Change, 15*(2), 287–302.
Joseph, M., Ramani, E., Tlowane, M., & Mashatole, A. (2014). Masekitlana re-membered: A performance-based ethnography of South African black children's pretend play. *South Africa Journal of Childhood Education, 4*(1), 17–41.
Joseph & Ramani (2004) Academic Excellence through Language Equity: A new Bilingual BA Degree (in English & Sotho sa Leboa). In H. Griesel (Ed.), *Curriculum Responsiveness: Case Studies in higher education.* SAUVCA: Pretoria, South Africa.
Judson, H. F. (1985). The rage to know. In J. Fahnestock & M. Secor (Eds.), *Readings in argument* (pp. 65–76). Random House.
Karpov, Y. V., & Haywood, H. C. (1998). Two ways to elaborate Vygotsky's concept of mediation. *American Psychologist, 53*(1), 27–36.
Kaschula, R., & Maseko, P. (2017). Researching the intellectualisation of African languages, multilingualism and education. In R. Kaschula, P. Maseko, & H. E. Wolff (Eds.), *Multilingualism and intercultural communication: A South African perspective* (pp. 19–33). Wits University Press.
Lambert, C. (2018). Educational perspectives on task-based language teaching. In R. Ellis, P. Skehan, S. Li, N. Shintani, & C. Lambert (Eds.), *Theory and practice of task-based language teaching: Multiple perspectives.* Cambridge University Press.
Lantolf, J. P. (2000). In Introducing sociocultural theory. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 1–26). Oxford University Press.
Lantolf, J. P., & Appel, G. (1994). Theoretical framework: An introduction to Vygotskian approaches to second language research. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 1–32). Ablex Publishing.
March, S., & Fleer, M. (2017). The role of imagination and anticipation in children's emotional development. In M. Fleer, F. G. Rey, & N. Veresov (Eds.), *Perezhivanie, emotions and subjectivity: Advancing Vygotsky's legacy* (pp. 105–128). Springer.
Paley, V. G. (2004). *A child's work: The importance of fantasy play*. The University of Chicago Press.
Palmer, H. E. (1921). *The oral method of teaching languages*. Heffer.
Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press.
Prabhu, N. S. (2019a). In G. Dorairaj (Ed.), *Perceptions of language pedagogy*. Orient Blackswan.
Prabhu, N. S. (2019b). Coping with the unknown in language pedagogy. In G. Dorairaj (Ed.), *Perceptions of language pedagogy* (pp. 45–58). Orient Blackswan.
Prabhu. N. S. (2019c). The learner's effort in the language classroom. In G. Dorairaj (Ed.), *Perceptions of language pedagogy* (pp. 145–158). Orient Blackswan.
Prabhu, N. S. (2019d). Rethinking language pedagogy: An interview. In G. Dorairaj (Ed.), *Perceptions of language pedagogy* (pp. 314–330). Orient Blackswan.
Prabhu, N. S. (2019e). Teaching is at most hoping for the best. In G. Dorairaj (Ed.), *Perceptions of language pedagogy* (pp. 241–252). Orient Blackswan.
Prabhu, N. S. (2019f). Plausibility. In G. Dorairaj (Ed.), *Perceptions of language pedagogy* (pp. 331–340). Orient Blackswan.

Prabhu, N. S. (2019g). Ideation and ideology. In G. Dorairaj (Ed.), *Perceptions of language pedagogy* (pp. 265–278). Orient Blackswan.

Ramani, E., & Joseph, M. (1997). Making monolinguals multilingual. *AD Issues, 5*(1), 15–17.

Ramani, E., & Joseph, M. (2010). Promoting academic competence in two languages: A case study of a bilingual BA degree (in English and Sesotho sa Leboa) at the University of Limpopo. *EFL Journal, 1*(2), 33–59.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social contexts*. Oxford University Press.

Vedeler, D. (2015). Intentionality during infancy and early childhood, development of. In J. D. Wright (Ed.), *International encyclopaedia of the social and behavioural sciences* (2nd ed., Vol. 12, pp. 340–346). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.203079-1

Veresov, N. (2017). The concept of Perezhivanie in cultural-historical theory: Content and contexts. In M. Fleer, F. G. Rey, & N. Veresov (Eds.), *Perezhivanie, emotions and subjectivity: Advancing Vygotsky's legacy* (pp. 47–70). Springer.

Vygotsky, L. S. (1978). In M. Cole, V. J. Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: The development of higher psychological process*. Harvard University Press.

Vygotsky, L. S. (1985/1933). Play and its role in the mental development of the child. In J. S. Bruner, A. Jolly, & K. Sylva (Eds.), *Play: Its role in development and evolution*. Penguin Books.

Vygotsky, L. S. (1987). Problems of general psychology: Including the volume thinking and speech. In R. W. Rieber, & A. Carton (Eds.), *The collected works of Vygotsky* (Vol. 1) (N. Minick, Trans.). Plenum Press.

Vygotsky, L. S. (1997). The history of the development of higher mental functions. In R. W. Rieber (Ed.), *The collected works of L.S. Vygotsky* (Vol. 4). (M. J. Hall, Trans.). Springer.

Vygotsky, L. S. (1998). Child psychology. In R. W. Rieber (Ed.), *The collected works of L.S. Vygotsky* (Vol. 5) (M. J. Hall, Trans.). Springer.

Vygotsky, L. S. (2019). L. S. Vygotsky, Volume 1. Foundations of pedology. In M. Fleer, E. Kravtsova, & N. Veresov (Series Eds.), F. González Rey (Founding Ed.), *Perspectives in cultural-historical research* (D. Kellogg, & N. Veresov, Trans. with notes and lecture outlines). https://doi.org/10.1007/978-981-15-0528-7

Wertsch, J. (1985). *Vygotsky and the social formation of mind*. Harvard University Press.

Chapter 4
Task-Based Language Teaching: Early Days, Now and into the Future

Rod Ellis

Abstract Task-based language teaching (TBLT) is now the approach mandated by a number of educational authorities in Asia. In this chapter, I will show how task-based language teaching (TBLT) grew out of communicative language teaching, drawing on both second-language acquisition research and theories of education. I will trace its development from its early days, pointing to the multiple influences that have helped to shape its evolution. I will address key issues such as how to define 'task', how tasks have been classified, how they can be sequenced into a syllabus, and how a complete lesson can be built around a task. I will also briefly consider the use of tasks in computer-mediated language teaching and in assessment. I will emphasize that TBLT is multifarious and conclude with a set of questions that can guide the ongoing development of TBLT.

Keywords Task-based language teaching · Early development · Later developments · Addressing problems

Introduction

Task-based language teaching (TBLT) is now firmly established as an approach that is supported by both second-language acquisition research and by principles of sound education (Long, 2015). It provides a basis for designing and implementing specific-purpose language programmes for adults and is also the recommended approach for teaching foreign languages to young learners in the state education systems of different parts of the world. It has been subject to extensive research investigating the design of tasks and their implementation (see Ellis et al., 2019) and of studies evaluating the effectiveness of complete task-based language programmes. There is now ample evidence that it is effective in developing learners' ability to communicate

R. Ellis (✉)
Curtin University, Perth, Australia
e-mail: rod.ellis@curtin.edu.au

Shanghai International Studies University, Shanghai, China

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_4

in a second language (L2) and that, with appropriate training teachers are able to implement it successfully. Nevertheless, doubts regarding its theoretical and practical viability have been expressed, especially in foreign language contexts where learners have no or little opportunity to use the L2 outside the classroom (Littlewood, 2014; Swan, 2005) while evaluation studies have pointed to problems in implementing TBLT, especially in state education systems.

I have two purposes in writing this article. The first is to examine briefly how TBLT has emerged as the preferred approach for teaching foreign languages by taking a historical perspective. The second is to identify key issues that remain and suggest ways in which they can be addressed. It has become clear to me that there is a clear difference between specific-purpose and general-purpose TBLT as the design of such programmes involves very different principles and procedures. My focus will be on general TBLT at the primary and secondary levels of state education systems. A quite different article is needed to address specific-purpose TBLT for adult learners.

A Brief History of Task-Based Language Teaching

Prior to the advent of communicative language teaching (CLT) in the 1980s, the dominant approach to teaching foreign languages was the structural one, where a language is broken down into phonological, lexical and grammatical bits that are presented and practised sequentially one at a time. While, it would be a mistake to say that task-like activities had no place in this approach, the focus was clearly on accurate production and the underlying assumption was that communicate ability would be achieved only when learners could synthesize the bits into wholes. CLT arose out of the belief that such synthesis often failed to occur and that the ability to communicate needed addressing more in other ways. Two quite different ways of achieving this were proposed—both claiming to be 'communicative'.

The first was the notional/functional approach (Wilkins, 1976) where functions such as requesting and apologizing and notions such as possibility and location replaced linguistic units in a syllabus. Such an approach claimed to be 'analytic' rather than 'synthetic' and, on the surface, appeared a radical alternative to the structural syllabus. However, an inspection of the teaching materials based on the notional/functional approach (e.g. Abbs & Freebairn, 1982) reveals that the 'new' approach was in fact quite similar to the 'old'. It still involved presenting and practising the language required to realize the notions and functions. Thus, while a notional/ functional syllabus was quite distinct from a structural one, the methodology used to implement it remained essentially the same—drills and situational exercises predominated. The problem of synthesizing the bits (in this case the linguistic exponents of the notions and functions) into communicative routines remained unsolved.

The second approach was to adapt the structural approach by adding a 'fluency' component. The adaptation took two forms. One was to provide a separate 'fluency' module to a structural curriculum. That is, in addition to the standard structural approach aimed at 'accuracy', there were separate fluency activities involving tasks

(i.e. activities that required learners to communicate as best they could using whatever linguistic resources they possessed). Books such as Klippel's (1985) *Keep Talking: Communicative Activities for Language Teaching* and Porter Ladousse's (1983) *Speaking Personally: Quizzes and Questionnaire for Fluency Practice* provided teachers with an array of ideas and activities for developing learners' communicative ability. Textbooks such as Aston's (1983) *Interact* contained activities aimed at increasing 'fluency' rather than 'knowledge'. Brumfit (1984) in his *Communicative Methodology in Language Teaching* articulated the rationale for including a fluency module alongside an accuracy module in a language curriculum. The place of tasks in a language curriculum was firmly established.

There was, however, a reluctance to abandon the structural syllabus so the second way of incorporating 'fluency' was as an add-on to the more traditional, drill-like activities in an accuracy-based curriculum. This gave rise to the presentation-practice-production (PPP) model, with the final P consisting of a task intended to provide opportunity for learners to communicate freely using the target feature that had been presented and practised in the previous two Ps. Anderson (2016) traces the emergence of PPP to the mid-1970s, when situational language teaching began to evolve into a more communicative approach. In *Teaching Oral English*, Byrne (1976) made the case for freer use of language on the grounds it as necessary to ensure that learners could 'use the language for themselves' (p. 80). PPP gained in popularity in the 1980s—see, for example, Harmer's (1983) *The Practice of language Teaching*. Howatt (1984) characterized it as a 'weak version' of the communicative approach. PPP has proved highly durable in ELT course books and in short initial teacher training courses such as the Cambridge CELTA and Trinity CertTESOL. Anderson suggests that this was because of its compatibility with a grammatical syllabus.

An important variant of PPP appeared in the 1980s. In the second edition of his book, Byrne (1986) suggested that the order of the three stages could be flexible, with the free production stage preceding the presentation and practice stages. Byrne drew on Johnson's (1982) 'deep-end strategy', where 'the student is placed in a situation where he (sic) may need to use language not yet taught' and thereby activate 'the ability to search for circumlocutions when the appropriate language item is not known' (p. 193). This necessitated communicative tasks where the learner's use of language was judged not in terms of whether it was grammatically correct but in terms of whether the communicative outcome of the task was achieved. Harmer (1997) in *How to Teach English* proposed Engage-Study-Activate (ESA) as an alternative to PPP, where 'engage' involves task-like activities based on stories, music and discussions designed to capture the students' interest, 'study' is a combination of the present and practice stages of PPP, and 'activate' involved free production. In all of these proposals, however, the focus of a lesson remained on a pre-selected target feature with the aim of enabling the learner to use this feature accurately and freely in communication. The tasks needed for this were of the 'focused' kind (Ellis, 2003)—that is, they were designed to elicit the natural use of the target feature. This is what second-language acquisition researchers objected to and that led to proposals for TBLT.

Second-Language Acquisition Research

The second-language acquisition research that started in the 1960s and 1970s fed into the emergence of TBLT. Cross-sectional studies of learners acquiring an L2 naturalistically (e.g. Dulay & Burt, 1973) provided evidence that there was an acquisition order for a common set of English grammatical morphemes that was common to all learners irrespective of their first languages (L1) or their age. A very similar order was found in classroom learners, suggesting that instruction did not have a major impact on the developmental route learners followed. Longitudinal studies (e.g. Cancino et al., 1978) showed that learners passed through a series of stages involving 'transitional constructions' on route to acquiring the target form. Progress was gradual and often very slow and at any one stage of development considerable variability was evident in those constructions available to the learner at that point in his/her development. Furthermore, it was clear that the initial stages of acquisition were lexical in nature. Learners relied on stringing words together that they had acquired and on formulaic chunks (e.g. *I don't know; What's this?*) which might seem to be grammatical but are learned as wholes (Wong-Fillmore, 1979). As they set about acquiring grammatical features, they do not do so in a linear fashion but rather work on several structures concurrently. This research led to the claim that there was a 'natural route' for mastering the grammar of a language and that there is a 'built-in syllabus' for learning it (Corder, 1967), which was independent of the external teaching syllabus.

Drawing on this research, Krashen (1985) argued that true proficiency in an L2 depends on 'acquisition', defined as 'the subconscious process identical in all important ways to the process children utilize in acquiring their first language', and not on 'learning', defined as 'the conscious process that results in 'knowing about' language' (p. 1). *The Natural Approach* (Krashen & Terrell, 1983) constituted an attempt to apply Krashen's ideas about how languages were 'acquired' to pedagogic practice. It emphasized task-like activities that focus learners' primary attention on meaning and cater to 'acquisition'. TBLT is based on the same principle.

Early TBLT Proposals

'Tasks' figured in both early CLT and The Natural Approach but in neither were they conceived of as the units around which a complete language course could be built. It was not until the mid to late 1980s that the first proposals for a task-based approach appeared. These early proposals (Breen, 1989; Candlin, 1987; Long, 1985) were largely programmatic in nature. They focused on the rationale for a task-based syllabus and outlined how to design and evaluate a task-based curriculum. Prabhu (1987) provided the first complete account of a task-based course [1], while Nunan (1989) gave practical advice about how to design tasks.

Rationale for TBLT

From the start, therefore, there were multiple inputs into the rationale for TBLT. Drawing on research in SLA, Long (1985) argued that 'there is no reason to assume that presenting the target language as a series of discrete linguistic or sociolinguistic teaching points is the best, or even *a* way to get learners to synthesize the parts into a coherent whole' (p. 79). He saw an approach based on tasks as providing an 'integrated solution to both syllabus and methodological issues' (p. 89). Candlin (1987) also critiqued traditional approaches but from an educational standpoint. He argued that an approach based on tasks would enable learners 'to become more aware of their own personalities and social roles' (p. 17), foster self-realization and self-fulfilment, and enhance their self-confidence. Along with Breen (1989), he emphasized the importance of teachers and students jointly negotiating the content of a course and argued that tasks provided the best means for achieving this. Prabhu's (1987) starting point was dissatisfaction with the Structural–Oral–Situational Method, which draws on a structural syllabus and was dominant in his teaching context in India at that time. He argued that 'the development of competence in a second language requires not systematization of language input or maximation of planned practice, but rather the creation of conditions in which learners engage in an effort to cope with communication' (p. 1) and that this could be best achieved by having students perform tasks. Drawing on all these sources, Nunan (1989) offered teachers a practical introduction to the design and use of tasks. He claimed that basing teaching on tasks avoided the traditional distinction between syllabus and methodology. He suggested that a structural syllabus was still needed but as a checklist that teachers could consult rather than as a directive about what they should teach.

Defining 'Task'

Central to the development of TBLT is a clear understanding of what a 'task' is. The early proposals for task-based teaching recognized this and offered definitions of a 'task', but these varied in a number of ways. Breen's (1989) definition was the most encompassing. A task is 'a structured plan for the provision of opportunities for the refinement of knowledge and capabilities entailed in a new language and its use during communication'—in effect any type of instructional activity. Other definitions emphasized that a task requires learners to focus on meaning rather than form (Nunan, 1989), that is should constitute a communicative activity in its own right and that the language for performing a task is negotiated as the task is performed. Long (1985) proposed that a task should bear some resemblance to a task that people perform in real life. He defined tasks as 'the hundred and one things people do in everyday life, at work, at play and in between' (p. 89). According to this definition, many of the 'tasks' described teacher resource books (e.g. Klippel, 1985) were not in fact tasks

Table 4.1 A typology of tasks types (Prabhu, 1987; 46–7)

Type of task	Definition
Information gap	This type involves 'a transfer of given information from one person to another—or from one form to another, or from one place to another'
Reasoning gap	This type involves 'deriving some new information from given information through the processes of inference, deduction, practical reasoning, or a perception of relationships or patterns'
Opinion gap	This type involves 'identifying and articulating a personal preference, feeling, or attitude in response to a given situation'

as they were not based on target tasks. Clearly, at this stage in the development of TBLT, there was no consensus as to what a task consisted of.

Classifying Tasks

As well as attempts at defining tasks, commentators proposed ways of classifying them. Candlin, however, felt that it is not possible to 'offer anything other than implicit suggestions that tasks might be catalogued under several distinct types' (p. 14) and that as a result 'a typology is bound to be fuzzy-edged and at most a managerial convenience' (p. 15). Nunan presented several task typologies drawn from different sources, the most useful of which was Prabhu's (see Table 4.1), which is based on how the information in a task is handled by the participants.

Grading and Sequencing Tasks

The construction of a task-based syllabus requires clear criteria for determining the difficulty of tasks. At this stage in the development of TBLT, however, such criteria were programmatic. Lists of factors likely to influence task difficulty appeared—for example, the linguistic complexity of the input in a task, the extent to which the task requires reference to present or past/future events, the intellectual challenge posed, and the learners' familiarity with the topic of the task. But there were no suggestions for how one factor should be balanced against others with the result that, as Prabhu found, the grading and sequencing tasks in the Communicational Teaching Project was a matter of intuition than precise measurement.

Subsequent Developments

The foundation for TBLT was laid, but several issues remained to be settled and TBLT was not yet established as a coherent approach to teaching a language. It was not until the 2nd edition of Richards and Rogers' *Approaches and Methods in Language* in 2001 that TBLT got a mention and, then, not as a distinct approach but as a variant of communicative language teaching. The sections that follow outline the nature of the developments that took place.

Broadening the Rationale for TBLT

Early account of TBLT paid little attention to broader educational principles, framing the rationale solely in terms of second-language acquisition research. Subsequently, however, advocates of the approach looked for support from general theories of education. Samuda and Bygate (2008) pointed to Dewey (1938), who rejected instructional approaches that focused on the mastery of ready-made products and emphasized the importance of learning that connects with experience of the real world. TBLT reflects the emphasis that Bruner (1960) placed on positioning the learner as a 'practitioner' rather than as a 'student'. Long (2015) devotes a whole chapter to the educational underpinnings of TBLT, claiming that TBLT is compatible with the requirements of a sound education—*education integrale*, freedom, emancipation, learner-centredness, egalitarian teacher–student relationships, participatory democracy and cooperation. In other words, TBLT was not only good for language learning but for the education of the whole person.

Defining 'Task'

Defining a task has continued to problematic, and definitions have proliferated over the years. Bygate et al. (2001) offered an all-purpose definition:

> A task is an activity which requires learners to use language, with emphasis on meaning, to attain an objective.

and then went on to suggest how it can be modified to suit the different purposes that tasks are used for. Van den Branden (2006b) reviewed 17 different definitions which they classified in terms of whether they referred to language learning goals or to educational activity. Such proliferation of definitions is not helpful, however. The problem arises because of the failure to distinguish task-as-workplan (i.e. what appears in a syllabus or in a set of teaching materials) and task-as-process (i.e. what emerges when the task is performed). Ellis (2003, 2009) argued that a task cannot be defined in terms of process as this is unpredictable and therefore can only be defined

as a workplan. He also argued that any definition should distinguish a 'task' from an 'exercise' and to this end proposed a definition based on four key criteria:

1. The primary focus is on meaning.
2. There is some kind of gap.
3. Learners have to rely mainly on their own linguistic (and non-linguistic) resources (i.e. they are not given the language they need to perform the task).
4. There is a clearly defined communicative outcome.

He noted, however, that some workplans satisfy some but not all four criteria. In other words, workplans can vary in the extent to they are true tasks of just task-like.

Task Types

Willis (1996) distinguished six types of tasks based on the cognitive operations they involve—listing, ordering and sequencing, comparing, problem-solving, sharing personal experiences and creative. Task-based researchers, however, have sought to distinguish task types in terms of the communicative and cognitive processes involved in performing them as these affect the way language is used. This has given rise to a widely accepted set of task types:

- Information-gap versus opinion-gap tasks
- One way versus two way tasks
- Convergent versus divergent tasks
- Closed versus open tasks.

Tasks differ in several other ways. They can be input-based (i.e. involve listening or reading) or output-based (i.e. involve speaking or writing). They can be 'real-world' (i.e. based on tasks that occur in everyday life such as booking a plane ticket) or 'pedagogic' (i.e. involve artificial activities such as Spot the Difference where learners have to find the differences between two pictures). They can be monologic or dialogic. They can involve different rhetorical modes (e.g. instructions, description, narrative, argument). There is no all-encompassing typology of tasks. This is not surprising given that tasks reflect the range of uses to which language can be put.

One distinction requires special attention, not least because there is controversy surrounding which type is compatible with TBLT. Tasks can also be unfocused or focused (Ellis, 2003). An unfocused task is intended to elicit general samples of language. A focused task must still satisfy the general criteria for a task but aims at orientating learners to the use of a particular linguistic feature—usually a grammatical structure. Advocates of TBLT differ in whether they think tasks should be entirely unfocused or a mixture of unfocused and focused. I will come back to this issue later in this chapter.

Task Selection

Task selection is also an aspect of TBLT that is controversial. Long (1985, 2015) argues that selection should be based on a needs analysis to identify the target (i.e. real-world) tasks that a specific group of learners need to be able to perform to function effectively in the social domains important to them. The obvious advantage of such an approach is that it ensures the relevance of a task-based course for learners with identifiable specific needs. It is less clear, however, how such an approach to task selection suits the needs of general-purpose learners such as those found in the state education systems of foreign language learners. Cameron (2001), for example argued that for young foreign language learners a needs-based syllabus is not feasible. Arguably, task selection for such learners should be based on identifying topics that are of interest to the learners and that involve familiar content. A good example of this approach can be found in Estaire and Zanon (1994), who classified topic areas in in terms of how close or remote they are to the lives of the learners—the students themselves, their homes, their school, the world around them and fantasy and imagination. Ideally, though, learners need to be consulted to identify suitable topics for tasks.

Task Complexity

The general principle for constructing any syllabus is that there should be a progression from simple (easy) activities to more complex (difficult) activities. Applying this principle to TBLT requires being able to identify what makes a task simple or complex. Early TBLT proposals recognized this need and proposed a set of criteria for evaluating the complexity of a task. Subsequent researchers have built on this by developing a theoretical framework of task complexity and by conducting studies to investigate whether manipulating specific task variables impacts in ways predicted by the theory.

Robinson's (2003) Cognition Hypothesis distinguishes resource-directing and resource-dispersing task variables that he claimed influence the difficulty of specific tasks. Resource-directing variables, such as whether there is contextual support, whether the task involves just a few or many elements, and whether reasoning is required, determine the cognitive complexity of a task and result in language use that is more accurate and complex. Resource-dispersing variables, such as whether learners have the opportunity to plan before they perform a task and whether they have prior knowledge of the topic of the task, reduce the processing burden involved in performing a task and encourage fluency. The difficulty of a specific task is determined by the variables it encompasses. Research, however, has not unequivocally supported the claims of Robinson's theory. Jackson and Suethanpronkul (2013) reviewed nine studies. They found that resource-directing variables did lead to more accurate language use but not to more complex.

A problem with Robinson's theory is that, unlike Skehan's (1996) framework for the implementation of task-based language teaching, it does not clearly distinguish variables inherent in the design of a task workplan (e.g. contextual support) from variables relating to the implementation of a task (e.g. planning). This failure may have arisen because of the widely accepted view in the early proposals for TBLT that 'task' is a unit that integrates *what* learners will learn (i.e. syllabus) and *how* they learn (i.e. methodology). In fact, the 'what' and the 'how' are as distinct in TBLT as in any other approach; any task can be implemented in a variety of ways. Task complexity is best considered in terms of the design features of workplans. Later in this chapter, I will address what this means for grading and sequencing tasks in task-based syllabuses.

Methodological Issues

The early proposals had little to say about how a task should be implemented and, Prabhu excepted, even less about how to plan a task-based lesson. Subsequently, however, greater attention has been paid to lesson planning and methodological issues in TBLT.

In Prabhu's Communicational Language Project, a task-based lesson consisted of a pre-task, which served as a preparation for a main task of the same kind. The pre-task was performed in a whole-class context while the main task was completed by the students working individually. Willis (1996) proposed a very different framework for a task-based lesson, one that prioritized learner–learner interaction. In her task-based learning framework, there was a pre-task stage to introduce the task and possibly pre-teach useful vocabulary, a task cycle stage consisting of the performance of the task, planning a report, and then presenting the report to the class, and finally a language focus stage. Later, in Ellis (2003), simplified Willis' framework into three stages—a pre-task stage, a main-task stage and a post-task stage with various options available in each stage. However, not every lesson has to involve all three stages; the only stage that is essential is the main-task stage.

A methodological issue of considerable importance—and one where there is again controversy—concerns how teachers handle the main-task phase. Willis (1996) advised teachers to 'stand back and let the learners get on with the task on their own' (p. 54) and argued they should resist the temptation to provide language support or correct learners' production while learners perform a task. She proposed that a concern for accuracy be delayed until after the task had been completed (i.e. to the planning and report stages of the task cycle). Long (1991, 2015), however, has consistently argued that there was a need to draw learners' attention to form *during* the performance of a task. He coined the term 'focus on form' to refer to a teaching strategy that 'overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication' (p. 45–46). One of the main ways of achieving this is through corrective feedback, but teachers and students also sometimes pre-empt a linguistic gap (see Ellis et al.,

2002) in the attempt to prevent a problem occurring. The recognition that task-based teaching does not necessitate an exclusive focus on meaning but also allows for (indeed requires in the opinion Long and other commentators) attention to form during the performance of a task constitutes one of the major developments in TBLT. It should also be noted that 'form' in this sense does not apply exclusively to grammar but also the phonological, lexical and pragmalinguistic forms.

According to Willis (1996), the point of the pre-task stage of a lesson is 'is not to teach large amounts of new language and certainly not to teach one particular grammatical structure' (p. 43) although she sees merit in helping learners with the vocabulary needed to perform the task. Other opportunities for introducing a focus on form in the pre-task phase are available. For example, giving learners the opportunity to plan before they perform a task will help them to both conceptualize what they want to communicate and to formulate the language they need to express it. There are also opportunities to focus on linguistic accuracy in the planning and reporting stages of the task cycle when the teacher can act as an adviser helping the learners to shape their meanings and to improve their choice of wordings, both grammatical and lexical.

However, it is the post-task stage—what Willis calls the language analysis stage—that offers the clearest opportunities for form-focused work including that of a more traditional kind. Willis and Willis (2007) suggested that when the task cycle stage is complete, the teacher is free to isolate specific linguistic forms for study, drawing on forms that learners used inaccurately when they performed the task or identifying language they did not use, but which would have made performing the task easier. In other words, the selection of the linguistic forms for attention in the final stage of the lesson or for later on is not pre-determined but based on the linguistic features that learners experienced actual difficulty with when they performed a task.

Other Developments

My focus so far has been on how TBLT evolved out of communicative language teaching into a distinct approach with its own theoretical basis, design principles and methodological procedures. To complete this account of its development, I will now briefly address three further aspects—the role of tasks in content-based language teaching, task-based assessment and technology-mediated TBLT.

Tasks in Content-Based and Content Integrated Language Learning

Content-Based Instruction (CBI) and Content Integrated Language Learning (CLIL) share with TBLT the assumption that a language is best learned when learners are

primarily focused on using language (Lyster, 2007). In CBI and CLIL, learners learn language through the process of mastering the content of (typically) academic subjects (e.g. history, science, mathematics). This can involve subject-relevant tasks where the task content is derived from the syllabus for an academic subject. However, CBI/CLIL do not rely exclusively on tasks; direct teaching of the language relevant for a particular subject area can also occur. Ortega (2015) suggests that TBLT focuses more on adult, college-level learners and CLIL on young, school-level learners, but this has more to do with contexts in which the two approaches have been developed than with any fundamental difference. There is in fact growing recognition that TBLT is highly relevant for foreign language contexts and for young children.

Technologically Mediated TBLT

Developments in CALL have mirrored those in language pedagogy in general. There was a structural/ behaviourist phase that gave way to a communicative phase and finally to a more integrative stage with the 'centrality of task-based authentic learning moving increasingly into the foreground' (Thomas & Reinders, 2010; 6). Technology-mediated TBLT has a number of advantages. It affords multimodal opportunities (aural, written and visual) for presenting complex workplans and for performing them both synchronously and asynchronously. For example, Appel and Gilabert (2002) described a task where learners had to plan a route and budget for a one-night trip and which involved email exchanges, the use of web pages and synchronous communication. Another advantage of technology is that it allows for presenting the input for a task sequentially in steps, which is not easy to achieve in the face-to-face classroom. Technology also affords the opportunity to interlock tasks into broader projects (Ortega, 2009), which again is arguably easier to organize with the assistance of technology. The increasing interest in technology-mediated TBLT is reflected in the growing literature on the subject (e.g. Gonzalez-Lloret & Ortega, 2015; Thomas & Reinders, 2015), the appearance of online TBLT courses (e.g. Duran & Ramaut, 2006), and research investigating technologically mediated tasks (González-Lloret & Ortega, 2014).

Task-Based Assessment

The emergence of TBLT has inevitably led to proposals for task-based assessment. In fact, though, as Bachman (2002) pointed out, the use of tasks for assessment purposes figured in language testing long before TBLT appeared on the scene. What was new was not the use of tasks to assess general language proficiency but for assessing whether learners were able to perform specific target tasks. When assessing general proficiency, the assessor makes a judgement of the learner's performance of a task based on a rating scale that specifies the different abilities being assessed and

the level achieved. Popular tests such as TOEFL and IELTS assess proficiency in this way. In task-based language assessment, however, assessment is based on task accomplishment:

> Task-based assessment does not simply utilize the real-world task as a means for eliciting particular components of the language system, which are then measured or evaluated; instead the construct of interest is performance of the task itself (Long & Norris, 2000: 600).

Long and Norris argued that assessment tasks, like teaching tasks, should be derived from target tasks. However, assessing in terms of task accomplishment does not necessitate the use of real-life tasks. Pedagogic tasks, especially if they are closed tasks (e.g. a Spot the Difference task that requires learners to identify a specific number of differences), can also be assessed by examining how successful learners are in achieving the outcome. However, assessment is not limited to examining the task outcome; there is also the possibility of assessing how learners perform a task. Van Gorp and Deygers (2014) based their assessment of a reading task designed for primary school students on a set of questions: (1) was the students' reading of the task-based material goal oriented, (2) could they find the information they were looking for, (3) was the teacher able to identify and address any problems the students experienced and (4) did the students demonstrate self-reliance, positive attitudes and reflective ability. Obtaining information about the performance of a task is especially important if the purpose of the assessment is formative rather than summative. Finally, there is also a strong case for student self-assessing how they performed a task. Butler (2017) showed that young children's self-assessment of their performance of a task correlated well with a more objective assessment. Tasks are not only a way of assessing learners' language proficiency or functional language ability but also serve as a source of information—for both teachers and learners—about whether the tasks were performed effectively and for reflecting on them.

Where We Have Got to?

I have traced the development of TBLT from its origins in the 1980s. It emerged as a strong form of communicative language teaching, supported by research in SLA. It offered a radically different alternative to mainstream language teaching by taking 'task' rather than 'language' as the starting point for the design of a course and by proposing that L2 learning can only be facilitated, not directed. Early proposals addressed a number of key issues—the definition of a task, the different types of tasks and the factors that determine task complexity. Subsequent developments from the 1990s onwards built on this foundation but also addressed how to construct complete lessons around tasks by distinguishing pre-task, main-task and post-task stages. There was also a stronger focus on the implementation of tasks with 'focus on form' seen as a key component of TBLT. Concurrent with these developments was a growing interest in technologically mediated TBLT and task-based assessment.

TBLT is now well established. Starting in 2005, there has been a biennial conference devoted to TBLT and recently a specialist journal called *Tasks* has appeared. A number of countries have officially mandated the use of TBLT. There have also been countless small-scale implementations of TBLT in contexts where teachers are free to choose their own approach (see, e.g., Leaver & Willis, 2004; Edwards & Willis, 2005). TBLT has progressed well beyond theory into actual practice.

It would be wrong to assume, however, that we have now reached a point where TBLT is a well-defined approach. Some advocates (Long and Norris, in particular) view TBLT as an approach that involves needs analysis to identify the real-life tasks that should figure in a teaching syllabus and in assessment. Such an approach is obviously required for specific-purpose courses. It is, however, not suited to school-level learners in foreign language learning contexts as such learners will usually not have any need to use the target language outside the classroom. An approach based on pedagogic tasks involving interesting and familiar topics is better suited to such contexts. There are also different views about TBLT methodology, with some advocates (Willis, for example) arguing against any attempt to focus learners' attention on form while a task is being performed and others (Long and Ellis) viewing a focus on form as essential component of task performance but with differences in how best to achieve this. TBLT is certainly not a 'method'. It is not monolithic but rather an approach with different versions all of which share the fundamental conviction that language learning is more successful when learners have to struggle to communicate in order to achieve the task outcomes.

There is a need to establish the parameters that define the boundaries of variation in TBLT. For example, is there a role for explicit language instruction in TBLT and if so what is it? There is also a need to address a number of outstanding issues relating to both the design and methodology of task-based courses. For example, how can TBLT work for complete beginners? In the next section of this chapter, I consider where TBLT needs to go by addressing these and other issues.

Where Are We Going?

Because TBLT constitutes a radically different approach to mainstream language teaching, it constitutes an innovation for many teachers. Evaluation studies (e.g. Beretta, 1990; Carless, 2004; Hu, 2013) point to problems arising when teachers attempt to introduce it:

- Teachers may not have a clear misunderstanding about what a 'task' is.
- Teachers are often uncertain how to use TBLT with beginner-level students.
- In some cases, teachers with limited target language proficiency experience difficulty in implementing TBLT.
- Students overuse their L1 when performing tasks.
- Teachers have difficulty adjusting tasks to the students' level of proficiency.
- TBLT is difficult to implement in large classes.

- There is a lack of task-based teaching resources and limited time for teachers to develop their own task-based resources.
- Teachers are uncertain about how to handle grammar in TBLT.
- In many instructional contexts, teachers need to prepare students for formal examinations.
- Teachers may lack of training in TBLT.

There have also been a number of published critiques of TBLT. Swan (2005), for example, disputes the theoretical basis of TBLT and argues that it is only suited to second-language contexts such as the USA. Clearly, if TBLT is to thrive, it is necessary to examine the issues raised by the evaluation studies and critics such as Swan. Below I discuss a number of these issues and suggest ways of addressing them.

Misunderstanding of the Nature of a Task

A fundamental requirement for the successful introduction of TBLT is that teachers have a clear understanding of what a task is. Studies have shown this is often lacking. Carless (2004), for example, found the elementary school teachers in Hong Kong had very vague notions of a task and were unable to provide a clear definition. Erlam (2016) found that the tasks she asked teachers to design in an in-service programme did not fully match up to the definition of a 'task' that she had given them. In particular, the teachers tended to design tasks that involved the prior presentation of language reflecting their pre-existing ideas of what an instructional activity should do. To address this problem, teachers need a clear definition of a task—such as the one suggested earlier in this chapter—and the opportunity to evaluate a range of different instructional activities in terms of whether they are 'tasks' or 'exercises'. Another way would be for teachers to try to modify exercises to convert them into tasks.

Teachers' Lack of Certainty About TBLT for Beginner-Level Learners

This problem arises because of a common misunderstanding about TBLT, namely that it involves only speaking tasks. Beginner-level learners—with no or little knowledge of the L2—cannot be expected to perform speaking tasks. However, they can undertake input-based tasks that involve listening to input from the teacher and demonstrating their understanding by performing an action. Such tasks do not require production, but they do not prohibit it and learners are free to try to produce the language they are exposed to if they want to. Ellis (2020) provides detailed guidance

about how to design and implement input-based tasks for beginner-level learners. Slattery and Willis (2001) also provide examples of tasks for beginners.

Input-based tasks are very compatible with what we know about the early stages of L2 acquisition. Learning begins with vocabulary and short phrases—not with grammar, which comes later. It is initially receptive and occurs incidentally when learners hear and understand what is said to them. It requires input that is comprehensible and multiple exposures to the same words and phrases. Shintani's (2016) study was based on this understanding of early L2 acquisition. She used some simple input-based tasks to teacher Japanese children who were complete beginners. In one task, the children had to listen to directions from the teachers (e.g. *Find the squirrel and take it to the zoo*), select the correct picture card from an array of cards to show they had understood, and deposit the card in a pocket of the outline of a zoo pinned to the wall. She repeated the task nine times. Initially, the children could only guess but eventually they were able to understand the teachers' directions. Interestingly, each repetition of the task resulted in quite different activity. Initially, the children resorted to their L1 to seek clarification of what they had to do but gradually they started trying to use English.

Teachers' Limited L2 Oral Proficiency

Teachers' lack of or lack of confidence in their L2 oral proficiency can result in them resisting the introduction of TBLT and relying on traditional approaches. This is, however, a not a problem specific to TBLT but for any approach aimed at developing learners' communicative competence. Input-based tasks can help with this problem too if the input needed for the performance of a task is scripted so that teachers are given the language they need. Input-based tasks involving reading can also help teachers with limited L2 proficiency as they can check in advance that they understand the language.

When it comes to output-based tasks, teachers need to understand that errors are inevitable and they do not need to worry about them. Students will benefit from seeing their teacher engage in the same struggle to communicate as they themselves are experiencing. In fact, teachers can improve their own ability to communicate by performing tasks with their students! Finally, teachers should understand that TBLT does not completely prohibit their use of the L1 and that there are times when it will be appropriate to use it—for example, in helping students to understand a key word needed to perform a task.

Students' Overuse of the L1

TBLT does not ban students' using the L1. Beginner-level learners will need to use it to ensure they understand what is required of them when they perform a task. Even

more advanced learners can make effective use of the L1 in dealing with linguistic problems that arise. Clearly, though, TBLT will not achieve its goals if students avoid using the L2 on a major scale. For TBLT to work, students need to be prepared take risks in using the L2 and to develop skill in using circumlocution strategies when they lack the linguistic resources needed to express their meaning. The problem of overuse arises when students are asked to perform a task that is too difficult for them. For example, as noted above, beginner-level learners cannot be expected to perform production tasks such as Spot the Difference if they lack the basic vocabulary needed to describe their picture. Again, a solution is to start with input-based tasks, where the linguistic input can be carefully controlled and where there are contextual aids such as pictures or actions to help understanding. Asking learners to speak before they are ready will inevitably lead to inappropriate used of the L1.

Difficulty in Selecting Tasks Suited to the Students' Level

Several educators have commented on the problem that teachers faced in selecting tasks at the right level of difficulty for their pupils. Van den Branden (2006c), for example, writing about the introduction of TBLT in schools in Flanders (Belgium), found that even after training teachers were not always able to identify tasks at the right level. This problem is not surprising given that one area of TBLT that remains relatively underdeveloped concerns criteria for grading tasks (see earlier discussion of this).

The long-term solution may be research that systematically investigates how specific task variables impact on the complexity of a task. A good example of such a study is Sasayama (2016). She investigated the complexity of four narrative tasks that differed in terms of the number of characters involved in the stories—one of the variables (Robinson, 2003) claimed affects task complexity. She reported that clear differences in learners' performance of the tasks were only evident between the least and the most complex task, suggesting that fine gradations of task complexity do not impact on the actual difficulty of tasks. Studies such as this can certainly help to develop a better understanding of task complexity but, to my mind, they are very unlikely to result in the kind of predictive model that will enable teachers to evaluate the complexity of a task scientifically. This is because tasks are conglomerates of variables and we know very little about they interact to determine difficulty. The solution is that suggested by Willis and Willis (2007): teachers should consult a list of variables that can affect task complexity but treat these as helping to sharpen intuitions about the suitability of a particular task for a particular group of learners. Ultimately, that will have to be decided based on experience.

Teachers also need to recognize that the difficulty of a task is not just dependent on its design features but on how it is implemented. In fact, how a task is implemented is probably more crucial for ensuring a good fit between the task and students (Skehan, 2016). Making a task is manageable by students, then, will involve careful attention to task-preparation variables such as pre-task planning, providing a model of the

performance of the task, brainstorming ideas relevant to the topic of the task and (more controversially) pre-teaching the language needed to perform the task. See Ellis (2019a) on task preparedness.

Implementing TBLT in Large Classes is Difficult

The management of large classes is a major problem irrespective of the teaching approach but it is perhaps greater in TBLT where the teacher often has to share control with students and so cannot direct proceedings. One solution is small group work, which has a major place in TBLT, but this can be difficult to organize, especially in classrooms where desks are arranged in traditional rows and the groups cannot be easily monitored by the teacher. It is possible, however, to conduct tasks in lockstep with the whole class. In the case of input-based tasks, this is required but it is also possible with output-based tasks if the information is split between the students and the teacher. For example, in a Spot the Difference task, the teacher could hold one picture and the students the other. In Prabhu's (1987) Communicational Language Teaching Project, the teacher first performed a task with the whole class before asking the students to perform a similar task by themselves, an approach that Prabhu adopted partly because he doubted the value of groups but also because of its suitability for large classes. Closed tasks may be easier to handle in a large class, because they make it easier to see if students have accomplished a task successfully and to provide feedback on how well they performed it.

Lack of Task-Based Resources

One of the factors that van den Branden (2006c) identified as important for the effective introduction of TBLT is ensuring that teachers have the task-based materials they need. But these are often lacking. Erlam (2015) interviewed teachers who had completed an in-service training course to investigate to what extent they were able to implement TBLT. She reported that one of the main difficulties the teachers mentioned was the lack of suitable materials. In fact, there are very few truly task-based courses that teachers can just take off the shelf and use in their classrooms so often teachers need to prepare their own tasks. Many modern textbooks, however, even though they are not task-based, do contain tasks and teachers can extract these tasks and plan task-based lessons around them. That is, instead of starting with the exercise-type activities that typically precede a task in these course books, they can adopt Johnson's (1982) deep-end strategy and start with the task and then use of the exercises in the post-task stage of the lesson if the students demonstrate a need for linguistically focused practice. Perhaps, the problem of the lack of resources will lessen as TBLT becomes more mainstream and publishers respond by publishing task-based courses.

Grammar in Task-Based Teaching

Teachers tend to be concerned about grammar, but grammar is not the only aspect of language that is important. As I have already noted, the early stages of language learning are lexical. Help in building vocabulary, including formulaic chunks, is very important in TBLT and achievable in a number of ways—in the pre-task stage of a lesson through direct teaching, through focus on form as learners perform a task, in the planning and reporting stages in Willis' task cycle, and in the post-task stage.

One of the criticisms levelled at TBLT is the lack of grammar. According to Sheen (2003), there is 'no grammar syllabus' in TBLT and its advocates 'generally offer little more than a brief list of suggestions regarding the selection and presentation of new language'. Swan (2005) claimed that TBLT 'outlaws' the grammar syllabus. It is true that there is no grammar syllabus in TBLT, but it is wrong to claim that there is no grammar. Grammar has a place in both the main-task stage—in the incidental attention to grammatical features in 'focus on form'—and, more clearly, in the post-task stage when explicit grammar instruction and traditional practice exercises have a place to address any grammatical problems evident in the performance of a task. Focused tasks—not welcomed by all advocates of TBLT—have a role to play here as they create a communicative context for the use of a specific grammatical feature and enable teachers to observe whether learners are able to use it correctly.

The critics' objection to TBLT really concerns their rejection of explicit grammar instruction. All advocates of TBLT advise against the a priori teaching of grammar as a way of preparing students to perform a task. Both Swan and Sheen favour presentation–practice–production where there is a grammatical syllabus, and each structure is presented and practised. Tasks serve as the means for providing communicative practice of the target structure. In effect, the critics are advocating a task-supported approach. A key question is the relative effectiveness of task-based and task-supported approaches, and there is a clear need to investigate this. However, there are problems in designing the kind of method-comparison study needed, and it not surprisingly that there have been few such studies. One of the best studies to date is Shintani (2016). She provided clear evidence of the superiority of task-based instruction over task-supported for very young, beginner-level learners of English.

Perhaps, though, it is unnecessary to insist that one approach is superior to the other. It is more a question of the instructional context and the stage of development of the learners. Long (2015), an ardent supporter of task-based instruction, acknowledges that in some instructional contexts, where teachers are used to a more traditional type of teaching, it might be an idea to smooth the way to the full introduction of TBLT by starting with task-supported teaching. In Ellis (2019b), I advanced the case for a modular syllabus, where there are separate task-based and task-supported components. Delinking the task-based and the task-supported components overcomes the danger of learners focusing primarily on form when performing tasks in the task-based component. I proposed that a complete language programme should be initially task-based, but the task-supported component could kick in when

learners had achieved functional communicative ability as a way of addressing continuing grammatical problems. I suggested, however, that the task-supported component should be not be based on a grammatical syllabus but on a checklist of potentially problematic grammatical structures—an idea floated by Nunan (1989)—that teachers could consult when deciding if there was a need to teach a specific structure. Evidence for this would come from observation of learners' performing tasks in the task-based component of the syllabus.

Preparing Students for Formal Examinations

In many instructional contexts, teachers have to prepare students for formal examinations where the focus is on grammar and the ability to use the target language correctly. The question arises, then, as whether TBLT is feasible in such contexts as teachers will naturally feel that they have to teach to the examination. It is worth pointing out, however, that TBLT can foster grammatical development. Beretta and Davies' (1985) evaluation of Prabhu's Communicational Language Teaching Project found that learners in the TBLT group outperformed the learners that received grammar-focused instruction in a contextualized grammar test although not in a discrete point test. Shintani (2016) reported that TBLT resulted in better acquisition of plural-s (a structure problematic for Japanese learners of English) than PPP. The input-based tasks in Shintani's study were of the focused kind that made if functionally necessary for the young learners to pay attention to whether a noun was singular or plural. Focused tasks, then, constitute a way in which teachers can prepare students for a formal examination without resorting to explicit grammar instruction.

It likely, however, that teachers will still feel the need to address grammar directly. A practical solution to this problem would be for teachers to adopt the kind of modular approach I suggested above—splitting the available teaching time between explicit instruction and TBLT. In the long run, however, this problem may disappear if task-based assessment replaces traditional forms of assessment.

Teachers May Lack Training in TBLT

This is one the 'real issues' that Long (2016) identified if TBLT was to move forward. Teachers are unlikely to be successful in implementing TBLT unless they have a clear understanding of its theoretical bases and practical know-how about how to implement it. There is some evidence that TBLT figures in undergraduate courses for trainee language teachers (e.g. Ogilvie & Dunn, 2010) and TBLT certainly figures in in-service courses for practising teachers (e.g. Erlam, 2015) and in courses that are part of an MA in TESOL/Applied Linguistics programme. However, many short teacher

preparation courses (e.g. the Certificate in English Language Teaching to Adults—CELTA) pay scant attention to TBLT, instead emphasizing traditional approaches for teaching phonology, lexis and grammar and the four language skills.

Ellis (2020) proposed a framework for developing a TBLT training course by identifying key factors relating to:

1. Course content (e.g. distinguishing a 'task' and an 'exercise'; ensuring that teachers understand that tasks can be input-based as well as output-based; examining examples of task-based materials)
2. Methodology of TBLT (e.g. utilizing training activities that require the same kind of experiential, discovery learning as TBLT; asking trainees to develop and try out their own tasks)
3. Uptake of TBLT (e.g. where possible providing training within teachers' own schools; ensuring there is adequate and ongoing support for implementing TBLT in the teachers' own schools).

Van den Branden (2006c) provides an account of a training programme that introduced TBLT into elementary and secondary schools in Flanders that illustrates these principles in action.

Conclusion

The history of language teaching is replete with approaches that fall in and out of fashion. The 1950s was the decade of Audiolingualism and the 1960s Situational Grammar Teaching. In the 1970s, humanistic approaches such as Suggestopedia and Community Language Learning became popular. From the 1980s, communicative language teaching took over, developing eventually into task-based language teaching. Each of these approaches came with its own theory of learning and language, design principles and sets of techniques and procedures. Each approach claimed to foster the development of the ability to communicate. Should we expect that TBLT will first flourish and then flounder as some of these other approaches have? TBLT differs from previous approaches not just because of its strong theoretical base (in SLA and education) but because of the support it has received from a growing research body of research. However, the problems and issues I discussed in the previous section suggest that there are major obstacles in its way. These problems will need to be addressed if TBLT is to flourish—in particular, there will need to be well-designed training programmes. If I were to make a prediction, it would be that TBLT will find its place not as an alternative to structure-based approaches but alongside them. Perhaps, then, the challenge ahead is how best to construct a balanced curriculum that incorporates both a task-based module and a structure-based one.

References

Abbs, B., & Freebairn, I. (1982). *Opening strategies*. Longman.

Bachman, L. (2002). Some reflections on task-based language performance assessment. *Language Testing, 19*, 453–476.

Beretta, A. (1990). Implementation of the Bangalore Project. *Applied Linguistics, 11*, 321–337.

Beretta, A., & Davies, A. (1985). Evaluation of the Bangalore Project. *ELT Journal, 39*, 121–127.

Breen, M. (1989). The evaluation cycle for language learning tasks. In R. K. Johnson (Ed.), *The second language curriculum* (pp. 187–206). Cambridge University Press.

Bruner, J. S. (1960). *The process of education*. Harvard University Press.

Butler, Y. (2017). Young learners' rationales for self-assessing their performance with and without tasks. Unpublished paper in 7th International Conference on Task-based Language Teaching, Barcelona.

Cameron, L. (2001). *Teaching languages to young children*. Cambridge University Press.

Cancino, H., Rosansky, E., & Schumann, J. (1978). The acquisition of English negatives and interrogatives by native Spanish speakers. In E. Hatch (Ed.), *Second language acquisition* (pp. 207–230). Newbury House.

Candlin, C. (1987). Towards task-based language learning. In C. Candlin, & D. Murphy (Eds.), *Language learning tasks* (pp. 5–22). Prentice Hall International.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*, 161–169.

Dewey, J. (1938). *Experience and education*. Collier Books.

Dulay, H., & Burt, M. (1973). Should we teach children syntax? *Language Learning, 23*, 245–258.

Duran, G., & Ramaut, G. (2006). In Van den Branden (Ed.), *Tasks for absolute beginners and beyond: Developing and sequencing tasks at basic proficiency levels* (pp. 47–75). John Benjamins.

Edwards, C., & Willis, J. (Eds.). (2005). *Teachers exploring tasks in English language teaching*. Basingstoke.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics, 19*, 221–246.

Ellis, R. (2019). Task preparedness. In Z. Wen & M. Ahmadian (Eds.), *Researching L2 task performance and pedagogy: In honour of Peter Skehan* (pp. 15–38). John Bnejamins.

Ellis, R. (2019). Towards a modular curriculum for using tasks. *Language Teaching Research, 23*, 454–475.

Ellis, R. (2020). Teacher preparation for task-based language teaching (pp. 99-119). In C. Lambert, & R. Oliver (Eds.), *Using tasks in diverse contexts*. Multilingual Matters.

Ellis, R., Basturkmen, H., & Loewen, S. (2002). Doing focus on form. *System, 30*, 419–432.

Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2019). *Task-based language teaching: Theory and practice*. Cambridge University Press.

Estaire, S., & Zanon, J. (1994). *Planning classwork: A task based approach*. Heinemann.

González-Lloret, M., & Ortega, L. (Eds.). (2014). *Technology-mediated TBLT*. John Benjamins.

Johnson, K. (1982). *Communicative syllabus design and methodology*. Pergamon.

Leaver, B., & Willis, J. (2004). *Task-based instruction in foreign language education*. Georgetown University Press.

Littlewood, W. (2014). Communication-oriented teaching: Where are we now? Where do we go from here? *Language Teaching, 47*, 249–362.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. Longman.

Krashen, S., & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom*. Pergamon.

Long, M. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77–100). Multilingual Matters.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, R., & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspectives* (pp. 39–52). John Benjamins.

Long, M. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.

Long, M. H. (2016). In defence of tasks and TBLT: Nonissues and real issues. *Annual Review of Applied Linguistics, 36*, 5–33.

Long, M., & Norris, J. (2000). Task-based teaching and assessment. In M. Byram (Ed.), *Encyclopedia of language teaching* (pp. 597–603). Routledge.

Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach.* John Benjamins.

Nunan, D. (1989). *Designing tasks for the communicative classroom.* Cambridge University Press.

Ortega, L. (2009). Interaction and attention to form in L2 text-based computer-mediated communication. In A. Mackey, & C. Polio (Eds.), *Multiple perspectives on interaction in SLA: Research in honor of Susan M. Gass* (pp. 226–253). Routledge.

Ortega, L. (2015). Researching CLIL and TBLT interfaces. *System, 54*, 103–109.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press.

Richards, J., & Rogers, T. (2014; 2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge University Press.

Robinson, P. (2003). The cognition hypothesis, task design, and adult second language learning. *Second Language Studies, 21*(2), 45–105.

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. Palgrave MacMillan.

Sasayama, S. (2016). Is a 'complex' task really complex? Validating the assumption of task complexity. *The Modern Language Journal, 100*(1), 231–254.

Sheen, R. (2003). Focus on form: A myth in the making. *English Language Teaching Journal, 57*, 225–233.

Skehan, P. (1996). A framework for the implementation of task based instruction. *Applied Linguistics, 17*(1), 38–62.

Skehan, P. (2016). Tasks versus conditions: Two perspectives on task research and their implications for language pedagogy. *Annual Review of Applied Linguistics, 36*, 34–49.

Slattery, M., & Willis, J. (2001). *English for primary school teachers: A handbook of activities and classroom language*. Oxford University Press.

Swan, M. (2005). Legislating by hypothesis: The case of task-based instruction. *Applied Linguistics, 26*, 376–401.

Thomas, M., & Reinders, H. (Eds.). (2010). *Task-based learning and teaching with technology*. Continuum.

Thomas, M., & Reinders, H. (Eds.). (2015). *Contemporary task-based teaching in Asia.* Continuum.

Van den Branden, K. (2006a). (Ed). *Task-based language education: From theory to practice.* Cambridge University Press Cambridge University Press.

Van den Branden, K. (2006b). *Introduction: Task-based language teaching in a nutshell.* In Van den Branden (2006a) (pp. 1–16).

Van den Branden, K. (2006c). *Training teachers: Task-based as well?* In Van den Branden (2006a) (pp. 2017–248).

Van Gorp, K., & Deygers, B. (2014). Task-based language assessment. In A. Kunnan (Ed.), *The companion to language assessment* (Vol. 2, pp. 271–287). Wiley Blackwell.

Willis, J. (1996). *A framework for task-based learning*. Longman.

Willis, D., & Willis, J. (2007). *Doing task-based teaching.* Oxford University Press.

Wilkins, D. (1976). *Notional syllabuses*. Oxford University Press.

Wong-Fillmore, L. (1979). Individual differences in second language acquisition. In C. Fillmore, D. Kempler, & B. Wang (Eds.), *Individual differences in language ability and language behaviour* (pp. 203–228). Academic Press.

Chapter 5
An Evolution of a Framework for TBLT: What Trainers and Teachers Need to Know to Help Learners Succeed in Task-Based Learning

Jane Willis

Abstract This chapter starts by outlining some major influences on the development of a task-based methodology as described in A Framework for Task-Based Learning, Willis (1996). It addresses common teacher questions like '*How do I know if it's a task?*' and '*When and how should I teach grammar?*' It introduces the TBLT framework, making explicit the rationale and principles behind each stage in the task cycle, the central part of the framework. It describes the changing roles of the teacher as they handle each stage and subsequent form-focused activities. It summarizes crucial conditions for learning as identified by Second-Language Acquisition (SLA) researchers and describes how phases of the framework can generate them. The focus then switches to the trainer. It reports first on experimental task-based training sessions for novice teachers, and secondly on a recent investigation into Task-Based Language Teaching (TBLT) sessions on short training courses. It explores three current challenges and offers possible solutions to two of them: ways to tackle the lack of training in TBLT, and the marriage of a lexical syllabus with a task-based syllabus. However, the unwillingness of publishers to produce task-based course books remains a problem. Is it largely the latter which impedes the wider uptake of TBLT in the teaching profession today?

Keywords Task-based methodology · Task-based learning · Task-based framework · Task-based language teaching · Task-based teacher training · Task cycle · Task goals · Focus on form · Form focus · Consciousness-raising activities · Meaning focus · Lexical syllabus · Language syllabus · Grammatical competence · Grammar teaching and TBL · Word frequency · CELTA courses · Teacher training courses

Overview This chapter gives a personal account of the development, over 30 years, of a framework designed to help teachers introduce task-based language teaching (TBLT) in their classrooms. Most teachers were following a form-focused approach,

J. Willis (✉)
Freelance Language Teaching Consultant, Kendal, UK
e-mail: jane@willis-elt.co.uk

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_5

with lessons beginning with the presentation and practice of a new language structure to be used by learners in a free production activity. However, TLBT is a meaning-focused approach and starts with a goal-oriented communication activity (a task) where learners exchange real meanings in order to achieve the outcome of the task. Long (2014, p. 6) explains: 'Instead of studying the new language as object in order to use it to communicate at some later date, *students learn language through doing pedagogic tasks.*' The vital difference is that while doing tasks, learners are not speaking simply to practise a recently taught structure or function, but are communicating their ideas or opinions in real time, choosing whichever language forms best express their own meanings. In other words, they are really meaning what they are trying to say.

This chapter differs from other papers in this volume in that it is not a research study. My overall aim is to provide a descriptive and analytical account of one way of implementing TBLT and its salient features. I attempt to build, through discussion and examples, a strong argument for its use both in teaching and training contexts. It is written with language teachers, novice trainees and trainers in mind because, despite 40 years of experimentation with task-based teaching, there has been very little coverage of TBLT on preliminary teacher training courses (Walker, 2020) and many misunderstandings and myths about TBLT still prevail. The decision to tackle this challenge rests with TBLT practitioners like yourselves.

Early Stages of Task-Based Language Teaching

My late husband Dave Willis and I began to experiment with task-based teaching in the early 1980s. We were frustrated by the limitations of methodologies that prioritized form-focused language teaching, even when supplemented with 'skills' lessons. The Presentation, Practice, Production (PPP) model, which focused on accurate use of the grammatical structure taught at the Presentation stage, giving learners a rigidly controlled and impoverished experience of language, was in common use at the time. Very few students who finished their English courses were able to use their English to communicate adequately with others. In many countries, adults with 5 years or more of school English realized that they now needed to be able to speak English and were joining classes at beginner and elementary levels.

Prabhu, the instigator of the Bangalore Communicational Teaching Project, was feeling equally frustrated. The school system in India had brought in the 'Situational, Oral, Structural' approach which focused narrowly on grammar and was getting poor results. Prabhu's main aim was to find out if learners could acquire grammatical competence simply by interacting in meaning-focused activities with their teachers, with no explicit teaching of grammar at all (Prabhu, 1987). His project ran from 1979 to 1984 in state schools, 6 secondary and 2 primary, with classes of between 30 and 60 children from disadvantaged backgrounds. For Prabhu's team, '*an activity which required learners to arrive at an outcome from given information through*

some process of thought, and which allowed teachers to control and regulate that process, was regarded as a task.' (Prabhu, 1987, p. 24).

After a year of experimentation, Prabhu's framework consisted of three parts:

PRE-TASK

Teacher rehearses a task, in English, involving the whole class, using the blackboard

(step by step teacher—class negotiation—questions and answers)

Teacher sets a parallel task giving oral or written instructions.

TASK

Students attempt the parallel task as individuals.

They answer a set of written comprehension questions or correct true/false statements relating to the outcome.

MARKING

Teacher does a quick marking of students' work, (usually overnight)

Assesses on the basis of content, not language form.

Tasks were based on topics covered at school, e.g. geometry diagrams, and real life, e.g. train time-tables. Each topic lasted a week (5 lessons) with tasks gaining complexity. The focus throughout the lessons was on pragmatic meaning-focused interaction rather than grammatical forms. Teachers would recast or rephrase learner contributions, and sometimes correct written errors but not practise or drill the corrected item. Prabhu did not use pairwork in the project because he did not want the class to be exposed to sub-standard English. However, after the project, he admitted to changing his mind on this.

For the end of project evaluation (externally validated), experimental classes were paired with control classes in each school, and although there was no significant difference in the results of the grammar and dictation tests, the other tests revealed that structure acquired without focus on form was more readily deployable than structure learnt with focus on form. Students from the project were able to use their English to communicate with other English speakers outside school far better than the control groups. For more discussion see Prabhu and Durairajan (2019).

We were encouraged by the success of Prabhu's project and also inspired by experiments with tasks and projects in primary and secondary schools in Spain around the same time, reported in Ribe (1997). We felt supported by recent research findings in the field of Second-Language Acquisition as well as experiments with tasks done under research conditions which were reported by Crookes and Gass (1993) Ellis (2003) and Skehan (1996, 1998) and Skehan and Foster (1997) among others.

So a small team of us teaching in Singapore began trialing tasks with learner groups from different backgrounds, including a group of Japanese beginners. Like Prabhu's tasks, the focus throughout was on meaning and getting meaning across. We tried various ways of using tasks, simplifying them and staging them and we listened to learners' feedback. We incorporated pair and group work giving more opportunities for learners to use the language themselves in different socio-linguistic contexts (Labov, 1972). And in a major change from Prabhu's project, we introduced a final phase, after the task, with an explicit focus on form as was, by then, being

recommended by SLA researchers and at the request of learners themselves. A challenge then arose. Which forms should we be focussing on? Much has been written on the integration of language form and tasks (Ellis 2003: 237) but how to combine a systematic and accountable language syllabus with a task syllabus? Which words, phrases and patterns would be the most useful for learners to study?

In the meantime, Collins and Birmingham University under Sinclair (1991) were undertaking the first ever computational linguistics project. They had assembled a large electronic corpus of millions of words of written and spoken English which included a wide range of books, newspapers, conversations and other sources of spoken English to compile the International Learners database (ILD) (hence the project name—COBUILD—Co for Collins, BU for Birmingham University plus ILD). Hitherto, dictionaries and grammars had been derived almost entirely from written text. Their aim was to uncover the most frequent meanings and uses of words in both written and spoken English in order to create a dictionary of 'real' English, with typical examples taken from the corpus.

The findings on word frequency were interesting. Over half the words in that corpus appeared only once, but the 2500 most frequent words were very frequent, accounting for over 80% of all English text (Willis, 1990; Sinclair & Renouf, 1998). The possibility of a lexical syllabus—derived from a checklist of the most frequent words with their most central and typical patterns—became viable. And since the words associated with a structural syllabus are among the most frequent words, a lexical syllabus would cover all 'traditional' grammar, plus many new patterns hitherto un-taught, in addition to the thousands of fixed and semi-fixed phrases frequently used in spoken English and naturally occurring discourse.

While experimenting with tasks in Singapore, we were asked by the COBUILD project to write a series of textbooks incorporating a lexical syllabus drawing on the findings of the COBUILD project. Since it is words and phrases that carry the basic meanings of the language and since TBLT prioritizes meaning over form, the logical vehicle for a lexical syllabus was a task-based methodology with a parallel syllabus of tasks and topics.

We designed tasks on a wide variety of topics selected by classroom teachers and vetted by the publisher. The 3-level courses highlighted the most frequent words, phrases and patterns as they occurred naturally in texts and task recordings. The 200 most frequent words are mainly 'grammar words' like *of, and*, *would, as, you* and verbs like *be, make, get* but also nouns like *time*, *people, way*, *thing, point,* all of which have many different patterns and uses, for example *One way of* + verb + *ing… The thing is/the point is…* The various uses of these very frequent words were treated in depth and revisited throughout the course when suitable contexts arose. Thus, the lexical syllabus ensured a thorough coverage of words, their uses and patterns, which was gradually built up using good examples in the written and spoken texts selected for each unit (Willis, 1990, 2003).

The Collins COBUILD English Course appeared in 1998 and gained excellent results in a small number of institutions worldwide where task-based learning was understood and supported, producing learners who were confident communicators after their first year (Lopez, 2004). However, it failed to become mainstream, partly

because so few teaching establishments trusted the task-based approach and many misapplied it, trying to teach more traditional grammar up front. Finally, a publisher take-over sealed its fate. A fellow course book writer sympathized, saying 'The world is not yet ready for a task-based approach'. They were right.

There was obviously a need for a basic accessible classroom guide to task-based language teaching, so I began to write about the framework we had built up over the years. Beginning with an exploration of teachers' beliefs about language learning, a brief rationale then highlighted the principles that underpin TBLT. I was able to pilot each chapter with newly graduated students of English in a Spanish university who were already doing some teaching. Finally, in 1996 'A Framework for Task-Based Learning', a practical 'how-to' handbook, was published.

The Spread of TBLT

Other books on practical aspects of TBLT were also appearing (Nunan, 2004; Ribe & Vidal, 1993) and gradually a number of teachers worldwide took up the challenge of task-based teaching. Some used tasks in their project work and some adapted the framework to suit their specific learners. Some applied it to the teaching of other languages, including Basque. Quite a few contacted us with questions and concerns.

Meanwhile more research was being undertaken on tasks and areas applicable to TBLT as reported by Ellis (Chap. 4 this volume). Computational linguistics was giving the ELT profession further insights into the close relationships between lexis and grammar, e.g. Pattern Grammar (Hunston & Frances, 2000).

Ten years on, and inspired by the research into aspects of TBLT that our own post-grad distance-learning students were doing in classrooms in different parts of the world (Edwards & Willis, 2005), we felt there was a need to explore and report on the expanding practice of TBLT and gather together the growing body of professional expertise. We called upon teachers worldwide to send us their successful tasks, their ideas for implementing them and advice for teachers new to TBLT. A total of 34 teachers responded, representing teaching communities in 12 different countries, and their experiences were incorporated into 'Doing Task-based Teaching' (Willis & Willis, 2007). We also began to develop a website for teachers interested in task-based teaching which meant we were in regular contact with teachers wanting to know more about TBLT and how to handle tasks in the classroom.

The remainder of this chapter will tackle some of the questions that teachers have asked over the years, going into more detail on the stages of the framework and suggesting solutions to some of the problems listed by Ellis (Chap. 4 this volume.) It begins with the problem of identifying what a task really is.

What is a Task? Criteria Framed as Questions

Originally (in Willis, 1996 pp. 23–24) we defined tasks as

> 'activities where the target language (i.e. English or Spanish) is used by the learner for a communicative purpose (goal) in order to achieve a real outcome'. '… learners are free to choose whatever language forms they wish in order to convey what they mean in order to fulfill, as well as they can, the task goals.'

But we found that this had been too rigid a definition to help teachers to recognize tasks, to adapt activities in their course books and to design and evaluate their own materials. Comments from teachers like '*I am still not really sure what a task really is*' were common. We ourselves had gained extensive experience of designing tasks that work (and many didn't!) through recording fluent speakers of different ages and backgrounds performing large numbers of tasks for the COBUILD course.

So, taking into account the many different definitions of tasks in the literature (e.g. Bachman & Palmer, 1996; Nunan, 1989; Skehan, 1998) and various interpretations reported by successful TBL teachers (Edwards & Willis, 2005), we found that, rather than attempting to define 'task', a more helpful procedure was to characterize successful tasks by listing criteria framed as questions.

The more confidently we can answer ***yes*** to each of these questions the more task-like the activity:

1. Does the activity engage learners' interest?
2. Is there a primary focus on meaning?
3. Are learners allowed free use of language?
4. Is there a specified goal with an outcome that can be shared?
5. Is success judged in terms of outcome? i.e. is task completion a priority?
6. Does the activity relate in some way to real-world activities?

These criteria do not form a watertight definition of the term 'task', but they can act as 'guidelines for the design and evaluation of activities which are task-like in that they involve real language use.' (adapted from Willis & Willis, 2007, p. 13).

The advantage of a list of criteria is that teachers can identify more precisely in which way a text-book activity needs 'tweaking' in order to make it more task-like and engaging. For example, Cox (2005, pp. 171–2) made the task goals more specific so learners would know when they had completed it successfully, e.g. He changed the original instructions: *Discuss which places tourists to your country should visit* to *List the three most interesting cities or places in your country and why people should visit them.* And in instructions for a task based on a problem page letter, rather than simply *What advice would you give this person?* he added, *Discuss your ideas then agree on the two best suggestions.*

Generally, we found that the more specific task instructions are, the richer the resultant interaction. Even story-telling activities in pairs or groups, e.g. '*Your most embarrassing experience*' can often benefit from a final instruction, like *Finally, decide whose experience was the most embarrassing* (Essig, 2005) as this will generate more interaction leading to a more satisfying completion point. So, the

instruction '*Tell your partner about some of your childhood memories*' (which in class can just 'fizzle out') might be better as '*Think of 2 childhood memories and then share them with your partner. Can you find 3 things that your memories had in common? Or 2 major differences?*' Knowing that they have completed a task successfully means that learners feel a sense of achievement which increases overall motivation.

The most important thing for teachers to realize is that tasks give learners opportunities to communicate for real and to mean what they say. They will be using language to exchange meanings for a real purpose. They need to know they will be free to use whatever words or language forms best express their meanings. The games they play, the problems they solve, the experiences they share in class may not always be things they will do in real life, but their use of language, because it is purposeful and real, will replicate features of language use outside the classroom.

The set of questions above can also help teachers to recognize tasks in course books and find activities they can adapt to make them more task-like. But, as Skehan (2003) confirmed—it also helps to exemplify what tasks are **not**. Tasks—as defined above—do not include activities which involve learners saying things just to practice language items, such as 'Describe the picture using some of the phrases below' or 'Ask your partner if he likes the foods below using the forms *'Do you like' 'Yes I do,' 'No I don't'*.' Equally untask-like are role-plays which are tightly controlled or semi-scripted, where there is no outcome or purpose other than practising pre-specified language or functions. These might make perfectly good practice activities but they would not count as tasks—the learner will be focusing on using particular language forms, not primarily on exchanging meanings to achieve an outcome.

A problem here is that tasks in course books often come at the end of a unit, after the Presentation and Practice of new language items, so the learners' mindset is still on the practice of new language item. But the final Production stage is supposed to offer learners a chance to use the new language more freely. However, a communication task in the position of the third P (as this is often called) can give rise to several problems. Firstly, some teachers do not attempt the Production activity at all for fear they will lose control of the class and that chaos will ensue; another common reason is that they have run out of time (Bilborough, 2019). Secondly, some learners will perceive this stage as further practice and overuse the target form creating very unnatural language. Thirdly, some learners happily freed from linguistic control, successfully achieve the goal of the task, without even using the new language items presented earlier, often much to the chagrin of the teacher. But this surely leads us to the conclusion—why not start the units with the task?

Since many teachers come to TBLT from a form-focused tradition like PPP, it is worth taking time on training courses to answer this common question.

How Does TBLT Differ from PPP and Other Lesson Frameworks?

The basic task-based framework differs from a PPP cycle because—from the start—the primary focus throughout the task cycle is on meaning rather than form. An explicit focus on specific language forms comes at the end of the cycle rather than at the start. The communication task itself is central to the framework, rather than a target language structure. After the teacher has introduced the topic highlighting useful words and phrases and clarified the task goal, learners carry out the communication task, using language they have learnt from previous lessons and other sources. Most TBLT teachers would ask learners to report back about how they did the task and compare findings. At some point, learners might listen to recordings of other people doing the same task, or read something related to the theme of the task, again relating this to their own experience of doing the task. Only after that is their attention directed towards useful grammatical and lexical features that occur naturally in the recordings they have heard or in the texts they have read during the task cycle.

In other words, learners begin with a holistic experience of language in use. They end with an analytical look at specific features of language form. By that point, the learners will have worked with that language and processed it for meaning. It is then that the focus turns to the surface forms that have carried those meanings.

One of the main problems we have in the classroom is providing a context for grammar teaching. Novice trainees find this very hard, but with TBL the context is already established by the task.

The basic framework can be summarized simply thus:

PRE-TASK

Introduction to topic and task

TASK CYCLE

Task → Planning → Report

FOCUS ON FORM

Analysis and practice of relevant grammar patterns and lexical phrases

No new teaching techniques are needed for a task-based approach, but it does demand a different weighting and sequencing. It also needs teachers to develop a mind-set that sanctions and encourages learners' attempts to put to use the language they have already learnt, even if they get things wrong. Woodward (1996) talks about achieving change gradually through informed debate and stresses the need for a paradigm shift. Hobbs (2010) suggests practical steps to encourage hesitant teachers to try TBLT.

Why a Three-Stage Task Cycle with a Report Stage?

It would of course be possible to help learners achieve fluency by filling our language lessons with nothing but tasks. Learners, talking to each other in pairs and small groups, would get plenty of opportunities to interact, to express themselves freely and gain confidence in using the target language. But, as Skehan (1996) has stressed, there could be a danger in this. Learners tend to gain fluency at the expense of accuracy. How can we prevent learners from evolving a type of classroom 'pidgin' or from fossilizing early? Skehan suggests that learners need to be kept on their toes, that they need a constant linguistic challenge. It is this that helps to drive their language development forward. This challenge is what a public report stage provides.

Here the studies carried out by Labov (1972) are relevant. Labov collected samples of people talking in a range of social settings. He found that people who made common use of vernacular forms in some settings would change to a more prestige version when speaking in a more formal setting. In the same way, even in our mother tongue, we are always aware of when we need to be on our best linguistic behaviour. If we have to speak in public, or present a case in a business meeting, we often plan beforehand what we want to say. We tend to speak more carefully and use different types of words; we may even change our accent. The same applies when writing. We will dash off a quick, personal email to a friend, but take far more care when drafting a letter to be published in a newspaper or a report to be made public. Willis and Willis (2010) expand on this.

The variety of language we use, then, depends on the circumstances of communication. We can summarize these as follows and relate them to classroom language use: the Task stage being private use, and the Report stage being more public. These are two ends of a natural cline.

Private use in pairs or small groups	**Public use** talking to whole class or writing
Spontaneous	Planned
Exploratory	Rehearsed
Ephemeral	Permanent (written or recorded)
Focus on fluency and getting meanings across somehow	Focus on fluency, accuracy, clarity and organization, as befits a public presentation
Correction rarely requested or acted upon	Correction and advice welcomed and incorporated

In the classroom, learners will need teacher support to bridge their linguistic gaps between appropriate private and public use. When they know they will have to report their task results to a wider audience, they will want to plan what to say, choose how to phrase it in appropriate English and avoid mistakes. They will also realize they should avoid using their first language and need to find ways to do this. Hence, the deliberate inclusion of an intermediate Planning stage, where the teacher takes into account what the learners want to say and helps them improve their language. The planning stage is where learners are really open to learning new things and getting

things right, and where teaching is most helpful. It is worth allowing ample time for learners to plan their reports and to encourage them to try out new ways of expressing themselves, so this stage is likely to last longer that the task itself.

Once learners know that there will be a 'public' report stage in English after the task, they will be more motivated to tackle the task more seriously, and use the Planning time more efficiently. Over time, it seems to reduce the amount of the first language being spoken at the task stage (Willis, 2009).

What Are the Teachers' Roles During a Task-Based Framework?

The overall framework aims to recreate in the classroom essential conditions for language acquisition and learning (see next section). The teachers' roles will vary according to the specific aim of each stage in the Framework. The Pre-task phase and the Task cycle proposed below take advantage of the sociolinguistic norms described above and ensure a smooth and natural transition from private to more public interaction. The roles are detailed below, using abbreviations T and Ss for Teachers and Students.

Pre-task

Introduction to Topic and Task

- T helps Ss to understand the theme and objectives of the task, e.g. brainstorming ideas with the class, using pictures, mime or personal experience to introduce the topic,
- T can highlight useful words and phrases and teach words learners ask for but would not pre-teach new structures. Topic lexis is vital for Ss to get their meaning across somehow; their grammar can be fine-tuned or taught later.
- T may get the class to do a pre-task, e.g. a prediction task based on words and phrases from the text, or an odd-word-out game with topic vocabulary.
- Ss can listen to a recording of a parallel task being done or T can rehearse the task with a good student (so long as this does not give away the solution to the problem). Or, if the task is based on a text, Ss read a short section to get an idea of what it will be about.
- T clarifies the goal of the task, what form the outcome will take, and checks Ss understand exactly what they have to do.
- Ss can be given 3 or 4 minute preparation time to think how to do the task and use their dictionaries if necessary.

- T decides how best to group Ss and sets time limit for the task. T might even begin the task with the class then get them to finish it in pairs.

This initial phase gives useful exposure to language being used by the teacher. This helps students to recall relevant words and phrases and to recognize new ones that will help them get their meaning across. Individual preparation time helps them to think of the kinds of things they can say, look up words they need and seems to result in richer language use at the task stage (Foster & Skehan, 1996).

The Task Cycle

Task

T's role is to encourage Ss to use whatever English they can recall to express themselves and say what they need to say in order to complete the task.

- T walks round and monitors, making sure all learners are actually doing the task, and encouraging everyone's attempts at communication in the target language, in a supportive and positive way. It's important for learners to feel free to experiment with language and take risks.
- If Ss get stuck, T helps them to formulate what they want to say, but will not intervene to correct errors of form unless asked or if they realize the meaning is not clear. T keeps the focus on the meanings Ss are trying to express.
- The emphasis is on promoting spontaneous, exploratory talk and confidence-building, within the privacy of the small group.
- T acts as time-keeper and stops the class once a majority of groups have finished.

Planning

T's overall role is to push learners' language development forward and challenge them to work out better ways to express their meanings.

- Ss draft and rehearse what they want to say or write during the more public Report stage.
- T goes round to advise students on language, suggesting phrases and helping Ss to polish and correct their language.
- If the reports are in writing, Ts can encourage peer-editing and use of dictionaries.
- The emphasis is on clarity, organization and accuracy, as appropriate for a public presentation.
- Individual students often take this chance to ask questions about specific language items.

Report

T's overall role is that of chairperson, letting learners/pairs speak in turn:

- T gives a clear purpose for others to listen. The class may need to take notes.
- T asks some pairs to report briefly to the whole class so everyone can compare findings, or begin a survey.
- T acts as time-keeper. Sometimes only two or three groups report in full; others comment and add extra points.
- T comments positively on the content of each report, but gives no overt public correction at this stage unless the meaning is not clear ('*OK, so what you mean is…?*'). Ts can note down problem language items that can be focused in a later lesson.

This component gives learners practice in public, prestige use of language and increases other students' exposure to spoken or written language. Even if some groups do not get to present their report orally, they will have learnt a lot at the Planning stage, and the teacher can call upon them to report next time. Or T can ask them to write them up either for a wall display or to share using social media or upload to a class newspaper.

Post-task Listening

- T allows Ss to listen (several times if needed) to a recording of fluent speakers doing the same task and asks them to compare this to the ways in which they did the task themselves.
- T leads a class discussion on the similarities and differences and gets Ss' feedback on the task itself, and possible follow-up activities.

This component gives additional exposure to topic-related interaction and increases students' experience of language in use in a familiar context.

Focus on Form

Analysis

- T sets some language-focused tasks, exploring the texts students have read or the transcripts of the recordings they have heard. These are often called 'consciousness-raising activities' (Willis, 2003; Willis & Willis, 2006). Examples include:
  - Find words and phrases related to the title or topic of the text. Circle them.
  - Read the transcript, find words ending in *s* or '*s* and say what the *s* means.

  - Underline all the verbs in the simple past form. Say which refer to past time and which do not.
  - Underline and classify the questions in the transcript.
  - Choose three phrases you think might be useful for others to know.
  - Find seven phrases that are typical of spontaneous spoken interaction but would not occur in writing.

- T starts Ss off, then Ss continue, often in pairs.
- T goes round to help; Ss can ask individual questions.
- T then reviews the analysis with the whole class, possibly writing relevant language up on the board in list form; Ss may make notes. Class shares the useful phrases and practices them.

The aim is to encourage students to explore language for themselves, to develop an awareness of aspects of syntax, collocation and lexis, to help them systematize what they have observed about certain features of language, to clarify concepts and to notice new things.

Practice

- T conducts practice activities as needed, based on the language analysis work already on the board, or using examples from the text or transcript. Practice activities can include:
  - choral repetition of the phrases identified and classified, focusing on pronunciation and intonation
  - memory challenge games based on partially erased examples or using lists already on blackboard for progressive deletion
  - sentence completion (set by one team for another)
  - matching the past tense verbs (jumbled) with the subjects or objects they had in the text
  - dictionary reference work on new words from text or transcript
- T may also draw attention to typical learners' errors noticed during the task cycle and get Ss to practise alternative ways of expressing those meanings accurately.
- It is unlikely that learners will gain instant command of any of these features. The aim is to get Ss to take note of salient features, so that they will recognize them when they meet them again in other texts and recordings, and later use them themselves.

See Appendices 1 and 2 for some examples of Consciousness-Raising activities based on a task transcript.

Optional Follow-up

At the end of the whole framework, students could:

- Repeat the same or a similar oral task but with different partners.
- Go back through the task materials and write down in their language notebooks useful words, phrases and patterns that they have noticed.
- Discuss how they felt about the task and the task cycle and what they might like to do next or some time later (and/or note this down in their diaries for their teacher to read later).

Some Myths About TBLT

With the spread of TBLT, a number of myths became prevalent. For example,

TBLT only works for oral skills,
You have to do pair or group work,
You can't use it with real beginners,
TBLT discourages explicit focus on grammar or language form.

Having read this far, you will be in a better position to recognize that these are in fact myths. However, on training courses, we found that giving a set of examples of tasks and task sequences suitable for different levels could help teachers develop their understanding of what a task is and what TBLT is about. Which of these examples below help to counter which myths?

The first example here can be done with the whole class as a teacher-led task. It would be suitable for most real beginners starting to learn English (instructions could be given in L1), since most of the words will be familiar.

1. **International Words**
   hotel football taxi disco jeans sandwich tennis music hamburger video
   internet goal museum Pepsi dollar basketball radio computer
   All these are English words. How many ways can you classify them? (Teacher brainstorms with class, recasting learner suggestions in English.) How many other English words do you know? Add them to your categories. Finally arrange your words in alphabetical order for each category and practice pronouncing them in English.

2. **Early Schooldays**
   Think of a teacher you remember well. Write down three things you remember about him or her. Then, in pairs, tell each other about them. Try to find two things your teachers had in common.
   Next, write a short paragraph about your teacher for others to read and display it on your classroom wall. Read about other people's teachers and find out if the majority were remembered for positive or negative reasons. List the reasons and classify them.
   Finally listen to a recording about a teacher that X remembered, and identify any similarities with other people's teachers.

3. **A News Story**
   Look at the headlines and the first three lines. Write down five questions that you think will be answered in the full story. With a partner, compare your lists and pick the 5 questions you are sure about. Exchange this new list with another pair. Then read the whole story, and see how many of those questions were actually answered in the text. Who found answers to the largest number of questions?
   Finally read the text again and underline all the verbs and verb phrases that refer to past time and find ways to classify them. Then, as a memory challenge, take turns to read out a verb and see if the class can remember who/what the subject was (exact wording) without looking at the text.

All three of these examples can be used for different topics and can be set up and adapted in different ways.

As well as countering a number of prevalent myths about TBLT, this set of tasks can be evaluated according to the list of criteria given in Section "What is a task? Criteria framed as questions" above—a useful activity on a course for novice teachers. It also shows that tasks can have a variety of starting points: a text (written or spoken), or learners' own input—their personal experience or their knowledge of the world.

How Does This TBL Framework Fulfil the Conditions for Learning Derived from SLA Research Findings?

Although the learning styles of individuals may differ, according to second-language acquisition (SLA) researchers there is overwhelming evidence that, in order for anyone to learn a language effectively in a classroom, there are four key conditions

to be met. In this section, I will briefly outline these conditions and relate them to components in the task framework.

Four Key Conditions for Language Learning

1. **Exposure to a rich but comprehensible input of real language**, i.e., the kind of language that learners will be needing or wishing to understand and use themselves
   Exposure and input come from teacher talk (especially during the pre-task phase and when reviewing language analysis), from students listening to each other, and from reading the texts or listening to recordings of others doing the task. This input is not confined to sentence-level examples, but consists of real, often spontaneous, language use.
2. **Opportunities for real use of language**—Chances for learners to experiment and test hypotheses, to mean what they say and express what they mean in a variety of circumstances.
   During the Task cycle, the Task stage gives students opportunities to use language to express what they want to say, to gain practice in turn-taking, controlling the interaction, interacting spontaneously in pairs. The Report stage then offers them the challenge of drafting and perfecting their report and presenting it to a wider audience. The Planning stage, before the Report stage, gives students the confidence and support that they need to revise and rehearse before they actually perform in public.
3. **Motivation** to process the exposure for meaning—to 'grapple with meaning' as Prabhu put it, and also motivation to use the language to speak and write appropriately.
   The goals of the task provide the main motivation to engage in a TBL lesson. Students generally want to achieve the task outcomes which involve them in working towards a specific goal, such as solving a problem. Success in completing the task is in itself a motivating factor. Then, because they have done, or will do, the task themselves, they are keen to listen to a related recording and read the transcript or a related text.
4. **Focus on language form** Although it is quite possible for people to acquire a new language without instruction, research has shown that, in order to prevent fossilization and to push language development forward, we need to challenge learners to strive for individual improvement. Learners also need chances to reflect on language and to try to systematize what they know.

Within the task cycle, the Planning stage, where they prepare to 'go public' for the Report, encourages learners to correct and improve their own language, and to try out new words and phrases, striving for accuracy as well as fluency.

A more concentrated focus on largely pre-selected language forms happens in the Form-Focused phase after the task cycle. These can be pre-planned because they

consist of language features occurring in the texts and task transcripts used during the task cycle. Analysis activities cast students into the role of 'text investigators'; during the consciousness-raising activities they are free to work as individuals at their own pace; free to make their own discoveries which they will be able to apply at some later time, when they are ready to, and when the need arises.

They are not being forced to work in lock-step, or concentrate on one single structure pre-selected by the book or the teacher, as in a PPP approach. They may of course practise pronunciation of useful language items, and consolidate new language.

Handling Grammar

As Ellis reported in Chap. 4 (this volume), teachers are uncertain about how to handle grammar in TBLT. A large proportion of the TBL Framework prioritizes meaning, but as we have seen, there are points in the task framework where the possibilities for language focus might occur. However a useful distinction between Language Focus and Form Focus is clarified and summarized in the diagram below, adapted from Willis and Willis (2007 p. 133) (Fig. 5.1).

Once students and teachers are used to a task-based approach and become aware of the learning opportunities it offers, they develop both as learners and language users, achieving greater fluency and confidence. It is, however, vitally important that both learners and teachers alike understand the principles behind the approach, and the rationale behind each component of the framework.

How Flexible Can This Framework Be?

Lesson Planning

The TBL 'framework' is not necessarily synonymous with 'lesson'. With a task that would benefit from independent learner preparation (e.g. web research, or reading a text), the pre-task phase could be started at the end of a previous lesson, with learners continuing to prepare at home. In the same way, the finalizing or rehearsing of the Report itself could be continued for homework, or even written or audio-recorded by the learner at home and then presented in the following lesson. This way of splitting up the framework and encouraging learners to record themselves is especially useful when teaching online.

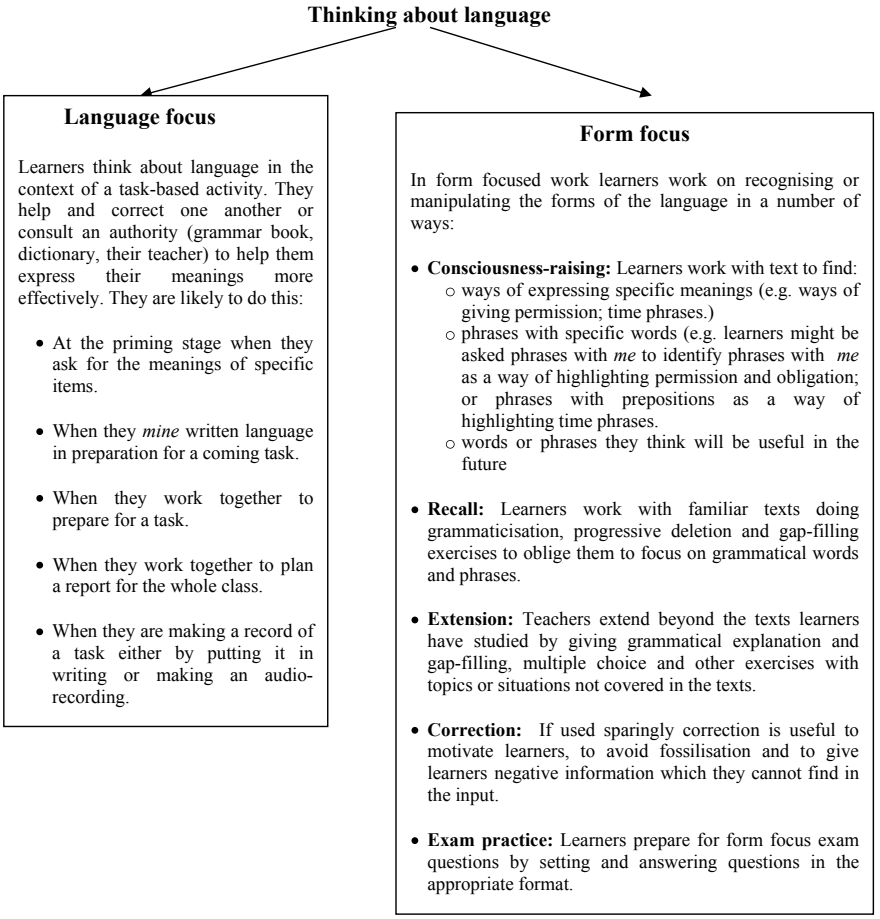


Fig. 5.1 Language focus and form focus

Report Stage

Some lessons—especially those based on reading texts or listening—may involve a sequence of two or three mini-task cycles, each task supplying a different reading goal, with only a brief report after each mini-task. Some tasks will not need a formal reporting phase at all, because the subsequent task grows directly out of the first. Other tasks, such as story-telling, where each pair or group has something quite different, may naturally produce a lengthy reporting phase. With problem-solving tasks, it is sufficient to hear only the groups that can offer different solutions. After the report, a vote can be taken for the best story or solution.

Adaptations to Meet Student Needs

Depending on the needs and backgrounds of students, the components of the framework can be weighted differently. Students who already speak quite fluently, such as those working in an English speaking country, may need a greater emphasis on accuracy and analysis work, i.e. less time on task and more time for planning and formal reporting, and more tasks requiring written outcomes. Recording their reports on audio or video would give them a greater motivation to achieve clarity and accuracy. Conversely, students from a grammar-oriented background, used to writing and reading but unused to using their English, may need a diet of speaking tasks, initially with no reporting stage, or with the teacher reporting their results, to give them confidence in speaking and a chance to develop their fluency.

With beginners, the actual task itself may be a 'listen and do' activity, requiring only recognition of meaning, with no learner speaking, and with the teacher summing up at a Report stage.

With ESP/LSP students who require a reading-only knowledge, the tasks set would be based on a text in the target language and could be discussed and reported in L1.

Group and Pair Work

There can also be flexibility in the way students are grouped. With a task-based approach, students of different levels can work together more easily, the weaker ones can learn from the others and gain confidence from the support of the small group. Sometimes, though, shy students feel less intimidated and contribute more if asked to work together.

More ideas and lesson plans based on adaptations of this framework as well as teachers' own answers to common questions can be found in Willis and Willis (2007) and at www.willis-elt.co.uk.

Is Task-Based Teaching Suitable for Novice Trainees?

I want now to go back to 1992, when I led a workshop on TBL for teachers and trainers in Izmir, Turkey.

As a result of the workshop, one of the trainers agreed to experiment by introducing TBL on her Certificate Course for novice trainees (Willis, 2016). All sessions were video-recorded on a single camera in a back corner of the room. After first experiencing PPP and teaching a PPP lesson, the novice teachers read about the TBL framework as described earlier in this chapter. Trainees were asked to identify similarities and differences between the two approaches. This is what they said

The trainees' initial comments, written after reading about TBL:

Similarities

- 'TBL is like a sort of PPP upside down—the steps are there but in a different order.'
- 'Most of the teaching techniques are similar.'
- 'There is attention to both accuracy and fluency.'
- 'They both include a focus on language form and meaning.'

Differences

- 'TBL doesn't teach isolated chunks of language then attempt to put them back into the "whole body" of language.'
- 'TBL starts with the "whole body" of language—language comes out of what learners know/can do/want to do, and out of the task.'
- 'The skills are really integrated and include what other methods call "micro-skills" as well as "the four skills".'
- 'There is a real need to communicate and to listen.'
- 'TBL distinguishes between private communication (fluency-based) and public communication (accuracy with fluency together).'

Trainees then experienced some TBL cycles for themselves—one based on comparing pictures, one based on a short newspaper story. Finally, for Teaching Practice they planned and shared the teaching of their own task-based lessons with their TP class of basic elementary adult students.

Trainees' subsequent comments (after teaching a 'Spot the difference task using a TBL framework)

- 'Tasks are intrinsically interesting.'
- 'You can do a lot with this approach.'
- 'You never know what the students have in their heads—it's amazing what comes out, and you find they have a lot they can build on.'
- 'Recordings of spontaneous spoken English—of people doing the same task—a radical departure from most course books because they are genuinely authentic; easy to understand because of natural repetition; students are motivated to listen because they have just done the same task and they want to compare how they did it.'
- 'The language-focused work was a bit too analytical for beginners—a focus on key phrases might have been better for them.'
- 'You're using what learners know and can do; much more learner-centred in a genuine way.'
- 'The students really talked a lot … and were very responsive and involved throughout' (despite trainees' initial doubts.)

The trainer said:

- 'It went better than a PPP lesson would have gone at this stage.' (Perhaps because TBL rests on natural communicative behaviour, rather than the tightly controlled teaching behaviour of PPP.)
- 'Trainees found the language analysis component hardest and were not always sure when to correct and feed in new language—but this was their first time.'
- 'In spite of the hitches, which seemed largely due to trainees having to "unlearn" PPP, the lesson flowed and the students loved it.'
- 'The question now is—can we do without PPP altogether?'

I will come back to this question at the end of this chapter. In subsequent discussion, the trainer and I agreed that untrained teachers find the TBL cycle a fairly easy one to follow, since it progresses naturally from private to public use of language, with a planning stage in between. And there is plenty of language data related to the task (reading text, task recordings) to be explored and used for form-focused study, so there is no need to think of examples to fit specific structures or patterns, or invent contexts for isolated language items. Neither do novice teachers need to try to teach complicated rules of sentence grammar that they may not feel confident with themselves.

We also agreed that one of the biggest challenges for the teacher or tutor used to a teacher-led PPP approach is to have the confidence to stand back and to let learners do the actual task on their own (as they will have to in real life). We all know that teaching does not necessarily result in learning, but sometimes it's hard to stop teaching and let learners learn.

I shall now jump forward 28 years to 2020.

The Continuing Need for More Training

As Ellis in Chap. 4 reported, one problem that impedes the wider uptake of TBLT in the language teaching profession today is the lack of training. This view is corroborated by East (2020) after a comparative study of two groups of newly trained teachers in New Zealand in 2014 and 2018:

> It is apparent that the challenges reside in upskilling and supporting practising teachers, who need to become a stronger focus of dedicated and long-term professional development opportunities that will introduce them to the innovation, and sustain them in its implementation.

East here was talking about lack of In-service training, but there seems to be a similar problem on short courses. In May 2020, in preparation for a one day conference led by Neil Walker at the University of Central Lancashire (UK) for trainers on Certificate courses, we devised a pre-conference questionnaire to explore trainers' attitudes to and use of TBLT. Over 60 trainers responded.

A majority (83%) were positive about TBLT, and many used tasks in their own classes. Some typical comments follow:

- …it fits students' needs more, in terms of global improvement of language that is also practical.
- It avoids a 'one-size-fits-all' approach …
- … it's learning by doing (experiential learning) and it's very motivating for students (satisfaction in achieving an outcome).
- It makes use of a variety of skills and language practice and exposes the learners to natural language use and real-life language experiences.
- It promotes more natural language acquisition.

However, when we asked how many sessions on TBLT they held on their certificate course, the following figures came to light:

- 17% No sessions on TBLT
- 44% 1 session
- 25% 2 sessions
- 13% 3 or more.

Nearly, all courses gave priority to a default lesson framework that began with a focus on form, such as PPP. TBLT, if treated at all, was simply one of several 'alternative' approaches. We know from experience that this kind of change involves a major paradigm shift, and that one or even two sessions will be insufficient to make a real difference. Of the 90 trainers attending the conference only six said they worked on a course that was largely task-based in approach, though 24 said they wished that they did so.

So why was there this discrepancy? Walker (personal communication) having reviewed the questionnaire data commented thus:

> It seems that trainers tend to quite like TBL – or most of them. The biggest gap between TBL and their default choice was for the statement: "I would use TBL more if the course books used in teaching practice used it"…
>
> This points to the chicken and egg problem of publishers not commissioning TBL-based books because teachers don't want them, but not training up new teachers in a TBL framework because course books aren't geared towards TBL.
>
> The other issue that jumped out was to do with emergent language. Respondents mostly said that trainees struggle with emergent language for both TBL and their default framework. You could argue this is why they choose frameworks that minimise the chances of having to deal with emergent language.

So how can we help? In the short term, there are two immediate ways in which we could help novice teachers on short courses. In the context of teaching practice, we could help them to identify, adapt and 'taskify' activities in the texts books they do have. The other is to offer regular language-focused tasks, exploring the language in the texts and recordings used in their TP lessons, so they gain more insights into how English works and more confidence in handling emergent language and 'difficult' questions.

However, it would almost certainly not be sufficient just to cover these topics in one-off sessions slotted into an already full course plan; this would be equivalent to an additive PPP model of language teaching which is not conducive to effecting a paradigm shift. A more holistic approach would be beneficial.

As suggested by Ellis earlier in this volume, and following the example of Van de Branden (2006), the ideal would be to adopt a task-based approach to the whole training course with task-based 'input sessions' for topics covered on Certificate courses. Thus, trainees will experience TBLT for themselves, and once the framework and the rationale are made explicit, they can go on to observe and analyse task-based language lessons, to identify task-like activities in text books and to try out their own tasks. In addition, the data used for task-based input sessions could also be used for regular form-focused language awareness work. This is sometimes known as 'loop input'—where the medium is the message (Willis, 2020). Thus, both trainers and teachers will gain deeper experience of TBL frameworks in action.

Current Challenges and Some Possible Solutions

To go back now to one of the questions posed in Turkey in 1992—could we do without PPP altogether? Many entire schools and colleges have succeeded in making the switch, e.g. Lopez (2004) and Moser (reported in Willis & Willis, 2007), and many learners of all ages have benefited. Less successful are the contexts where teachers have not truly grasped the rationale for TBL, where teachers and students adhere to the security of a teacher-led presentation of grammar and pattern practice before the task. Many still believe that learners '*must be taught the correct forms first, otherwise how can they do the task?*' regardless of the facts that natural interactions do not repeatedly use the same patterns and that acquiring the grammar of structure and orientation takes far longer than learning new words and phrases. An understanding of the implications of relevant SLA research findings (e.g. Lightbown & Spada, 2006) would also help, but many short training courses seem not to include any language learning theory.

East (2020) sums up the current situation arguing that

> TBLT remains in practice a contested endeavour. Positive experimental research findings are not necessarily finding their way into classrooms, and practitioner-focused studies play a crucial role in adding to our knowledge of what works and what does not work in authentic contexts.

Long-standing adherence to the structural syllabus is another challenge. Most practitioners agree there is a need to be accountable as far as linguistic coverage is concerned, but a structural syllabus, especially when delivered as an additive 'item-by-item' approach is not borne out by learning theories and does not fit easily with a task-based syllabus. More importantly, a structural approach is essentially a grammar of written English, and pays scant attention to word grammar, collocation and the

thousands of partially assembled phrases and formulaic expressions that abound in spoken language that make it easier for learners to interact spontaneously in English.

The Cobuild project has shown that a lexical syllabus can be easier to integrate naturally with a TBL approach. A syllabus of topics and tasks (selected with learners' needs in mind), supplemented with related texts and task recordings, can form what Willis (2003), and Willis and Willis (2007) call a 'pedagogic corpus'. Examples of high-frequency words in their typical uses and patterns will naturally abound in this data. And these patterns can be the focus of subsequent consciousness-raising activities in a Form-Focused phase. Useful topic vocabulary can be highlighted as it occurs in typical phrases. The inclusion of these highest frequency words, their common uses and patterns would more than adequately cover the patterns traditionally included in a typical structural syllabus. It would in fact provide a far richer diet of natural language abounding in useful lexical phrases that learners (in their role as text investigators) love to identify.

A checklist of the most frequent 300 or so words with their meanings and patterns can help materials writers ensure linguistic coverage of the language learners will meet in real life. This process of syllabus design is covered in more detail in Willis and Willis (2007, pp. 187–198) and summarized in this figure (Fig. 5.2).

But of course, as reported in Ellis (Chap. 4 this volume) and identified in the initial Walker and Willis (2020) investigation, the biggest problem still remains the lack of task-based text books. This seems mainly because of the unwillingness of publishers to take on task-based courses; they are perceived as too risky. A few task-based books have appeared recently for example Anderson and McCutcheon (2019) and Harris and Leeming (2018), but the profession as a whole, from teachers upwards, needs to push harder for change! Let us hope that the chapters in this book will help to increase the understanding and uptake of TBLT.

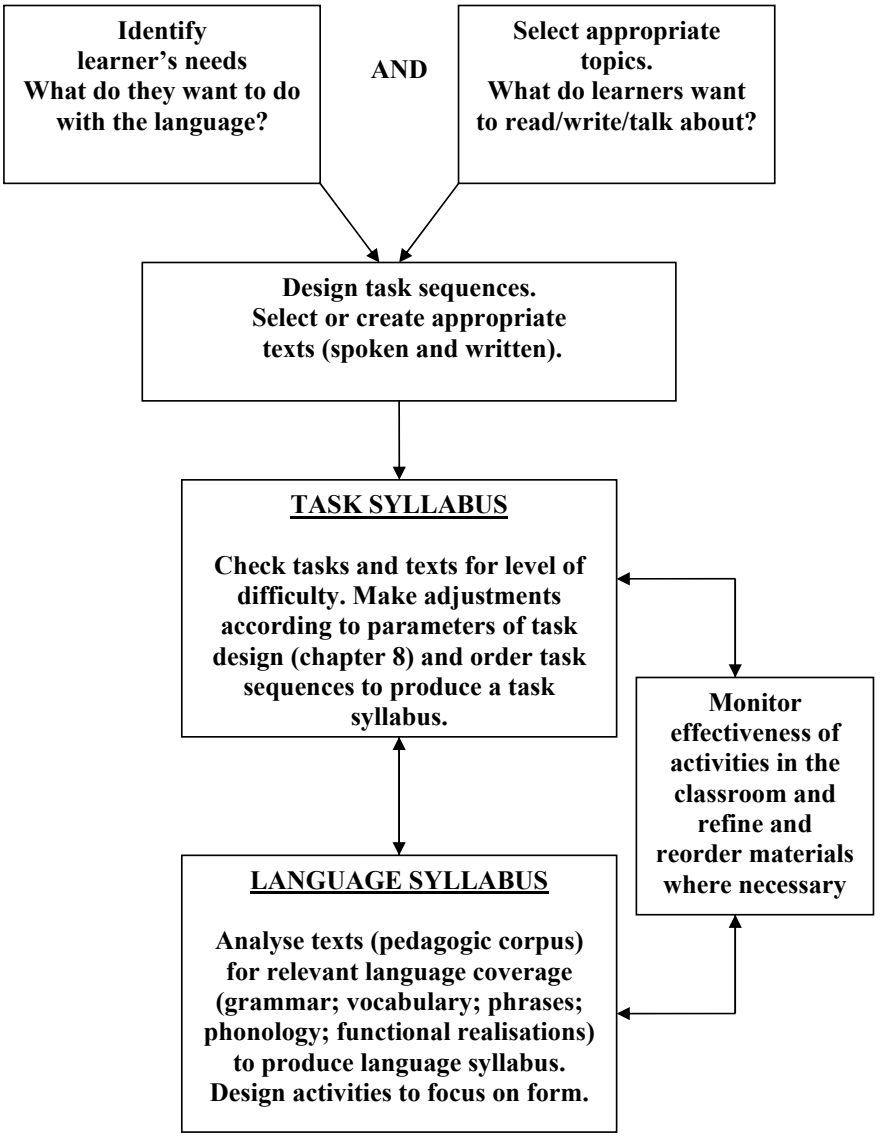


Fig. 5.2 Syllabus design procedures

Appendix 1: A Sample Form Focus Activity—Elementary

This activity is taken from an elementary course (Willis & Willis, 1988: 15) and shows that Consciousness-raising activities can be used from the very early stages.

The activity follows on from two task cycles on the topic of 'families': a class family survey (more men/boys than women/girls in your family?) and a task where learners in pairs have been asked to draw the family tree of their partner, and then see how many names they can remember. The same tasks were done by two pairs of fluent speakers of English and audio-recorded. Learners listened to the recordings before they did the tasks themselves. At the end of the task cycles, they studied the transcripts of the language used by the fluent speakers. The recording transcripts were then supplemented by a short written text about Danny and Jenny's offices which learners had read earlier. This particular form-focused exercise asks learners to investigate why words in English end in *s* or '*s*. But this transcript could also be used to highlight features of spontaneous spoken English.

Form Focus - Words ending in *s or 's*

Look at the transcripts below of David and Bridget talking about their families (Sects. 26b and 19). How many words are there ending in s or 's?

Does the *s* or *'s* always mean the same?
Some words always end in *s*, for example, hi*s*. What about this one?
I've got one brother and he's got two daughters
Put the words ending in *s* or 's into 4 or 5 categories

Bridget's family

DF: If we look at, erm, your mother Sheila. Has she got any brothers and sisters?
BG: Yes, she's got one sister
DF: No brothers?
BG: No
DF: Okay. What about your father?
BG: He's got three sisters
DF: Oh, and no brothers?
BG: No

David's family

BG: Now it's my turn. Your father's called John? and your mother's called Pat?
DF: That's right
BG: and your brother's married—to … Jane?
DF: Jane. Good
BG: Jane. And they've got two daughters called … Emma and Sarah.

Now look at the text in Sect. 24. Find thirteen more words that end in s and put them into categories

Danny and Jenny

Read about Danny and Jenny. Say which picture is Danny's office and which is Jenny's office

Danny lives in London. He's self-employed. He's got a studio office in Holborn in Central London, where he works with his brother.
They have a design agency. 'We do leaflets, brochures, … that sort of thing. So we are both self-employed, both me and my brother.'
They also have a 'rep', somebody who goes out and finds more work for them.
Jenny is also self-employed. She's an editor and writer. She works for a lot of different companies.
Jenny has a flat in North London and she works from home: 'I've got a sort of office in my flat…

Commentary

The aim of these exercises is to draw learners' attention to the wide variety of meanings of the final *s* in English, so that they listen for it and learn to recognize its various meanings and uses. The final *s* is the targeted feature isolated for study here, but they could also be asked to find 5 or 6 phrases with the word *got,* classify them and practise saying them out loud.

While reading through the data, learners may also pick up insights into the nature of spontaneous spoken English, for example the role of words like *Yes, No, Okay, Oh.* They might notice verb-less questions like *No brothers?* other useful phrases like *What about your …?*, *Now it's my turn* and *that sort of thing,* a useful phrase if you want to be vague or if you can't think of what else to say. They could also be asked to choose 3 phrases they like, to share with the class.

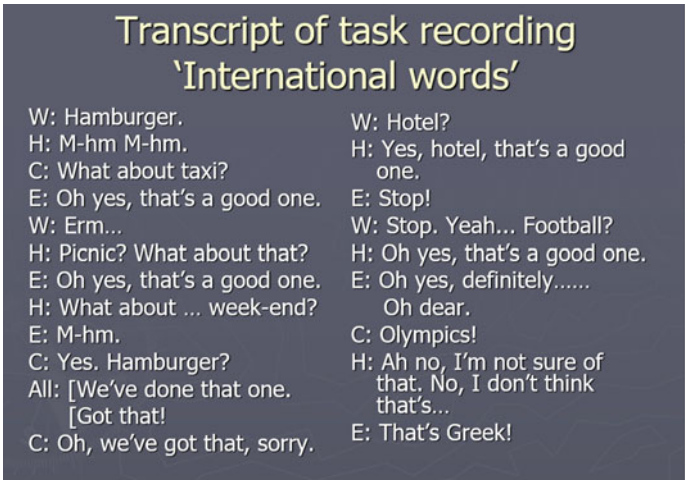
Note that by the time they study these features of spoken language being used in a communicative situation by fluent speakers, learners will already have carried out tasks in similar communicative situations and tried to express similar meanings themselves. Also note that the data for this exercise is drawn from texts that learners have already heard and processed for meaning, at the Pre-task stage, so they are familiar with the contexts and will be more able to focus attention on specific features of the language itself. Thus, the learner has taken on the role of researcher, investigating the targeted features (words ending in *s,* the word *got*) as they naturally occur in the linguistic data provided by the task recordings and short texts.

Appendix 2: A Sample Form-Focused Activity for Beginner Learners

International English Words This activity is taken from a beginners' course (Willis, 1990). Four fluent speakers were recorded doing this task: Together write a list of English words that are known internationally. (Time limit: 2 min). In the classroom, the teacher starts the first lesson brainstorming with the class some words they might already know, like *football, goal* and writes them on the board. They then listen to the recording to identify the international words listed by the fluent speakers.

Form Focus

Listen to the recording of the task again and raise your hand when you hear an international word.
Listen again and identify 6 questions. Then read the transcript and underline the questions. Choose three to practise saying to your partner.
Find three ways of meaning Yes and two ways of meaning No.
Can you find 10 phrases with the word *that* and classify them?



References

Anderson, N., & McCutcheon, N. (2019). DELTA Publishing.
Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford University Press.
Bilborough, N. (2019). Personal communication, reporting on The Hands Up Project. https://handsupproject.org/
Cox, D. (2005). Can we predict language items for open tasks? In C. Edwards & J. Willis (Eds.), *Teachers exploring tasks in ELT* (pp. 171–185). Palgrave MacMillan.
Crookes, G., & Gass, S. (Eds.). (1993). *Tasks in a pedagogical context: Integrating theory and practice*. Multilingual Matters Ltd.
East, M. (2020). What do beginning teachers make of task-based language teaching? A comparative re-production of East (2014). *Language Teaching,* 1–15. http://doi.org/10.1017/S026144481900048X
Edwards, C., & Willis, J. Eds. (2005). Teachers exploring tasks in english language teaching. Palgrave Macmillan.
Ellis, R. (2003). *Task-based language teaching and learning*. Oxford University Press.
Essig, W. (2005). Storytelling: Effects of planning, repetition and context. In C. Edwards & J. Willis (Eds.), *Teachers exploring tasks in ELT* (pp. 201–213). Palgrave MacMillan.
Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition, 18*, 299–324.

Harris, J., & Leeming, P. (2018). On Task. ABAX ELT publishers.
Hobbs, J. (2010). Practical steps towards task-based teaching. In *JALT 2010: 36th Annual International Conference on Language Teaching and Learning & Educational Materials Expo*, Aichi Industry and Labor Center, Nagoya, 19–22 November 2010. http://jalt-publications.org/proceedings/articles/1118-practical-steps-towards-task-based-teaching
Hunston, S., & Frances, G. (2000). *Pattern grammar. A corpus-driven approach to the lexical grammar of English.* Benjamins new on-line resource at https://grammar.collinsdictionary.com/grammar-pattern
Labov, W. (1972). *Language in the inner city.* University of Pennsylvania Press.
Lightbown, P., & Spada, N. (2006). *How languages are learned* (*Oxford handbooks for language teachers*).
Long, M. H. (2014). *Second language acquisition and task-based language teaching* (1st ed.). Wiley-Blackwell.
Lopez, J. (2004). Introducing task-based instruction in Brazil: Learning how to leap the hurdles. In Leaver & Willis (Eds.), *Task-based instruction in foreign language education.* Georgetown University Press.
Nunan, D. (1989). *Designing tasks for the communicative classroom.* Cambridge University Press.
Nunan, D. (2004). Task-based language teaching. Cambridge University Press.
Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press.
Prabhu, N. S. (Author), & Durairajan, G. (Ed.). (2019) *Perceptions of language pedagogy*. Orient Blackswan.
Ribe, R. (1997). *Tramas creativas y aprendizaje de lenguas: prototipos de tareas de tecera generacion.* Universitat de Barcelona.
Ribe, R., & Vidal, N. (1993). *Project work step by step.* Heinemann.
Sinclair, J. M. (1991). *Corpus, concordance, collocation.* Oxford University Press.
Sinclair, J. M., & Renouf, A. (1988). A lexical syllabus for language learning. In R. Carter & M. Mccarthy (Eds.), *Vocabulary and language teaching*. Longman.
Skehan, P. (1996). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching.* Heinemann.
Skehan, P. (1998). *A cognitive approach to language teaching.* OUP.
Skehan, P. (2002). A non-marginal role for tasks. *ELT Journal, 56*(3), 289–295.
Skehan, P. (2003) Tasks in L2 learning and teaching. *Language Teaching, 36*(1), 1–14.
Skehan, P., & Foster, P. (1997). The influence of planning and post-task activities on accuracy and complexity in task-based learning. *Language Teaching Research, 1*(3), 185–211.
Van den Branden, K. (Ed.). (2006). *Task-based language education: From theory to practice.* Cambridge University Press.
Walker, N. (2020, research in progress). Report of questionnaire results prior to UCLAN supersize your CERT conference May 2020 personal communication.
Walker, N. & Willis, J. (2020, research in progress). Task-based language learning on TEFL courses for novice teachers.
Willis, D. (1990). The lexical syllabus Collins COBUILD. https://www.birmingham.ac.uk/Documents/college-artslaw/cels/lexicalsyllabus/lexsch6.pdf
Willis, J. (1990). Collins COBUILD english course first lessons. Collins ELT.
Willis, D. (2003). *Rules, patterns and words: Grammar and lexis in English language teaching.* Cambridge University Press. Chapter 1 available at http://www.finchpark.com/courses/tkt/Unit_02/grammar-lexis-in-elt.pdf
Willis, J., & Willis, D. (1988). Collins cobuild english course level 1. Collins ELT.
Willis, D., & Willis, J. (2006). *Consciousness-raising activities in TBL.* http://www.willis-elt.co.uk/articles/
Willis, D., & Willis, J. (2007). *Doing Task-based Teaching.* Oxford University Press.
Willis, D., & Willis, J. (2010). Six propositions in search of a methodology: Applying linguistics to task-based language teaching. In S. Hunston & D. Oakey (Eds.), *Introducing applied linguistics: Concepts and skills.* Routledge.

Willis, J. (1996). *A framework for task-based learning*. ebook edition 2012. https://www.intrinsicbooks.co.uk/titles/framework.html

Willis, J. (2009). The TBL framework: The task cycle. In K. Van den Branden, M. Bygate, & J. Norris (Eds.), *Task-based language teaching—A reader*. John Benjamin's Publishing Company.

Willis, J. (2016). A flexible framework for task-based learning. https://www.academia.edu/30503745/A_flexible_framework_for_task-bas

Woodward, T. (1996). Paradigm shift and the ELT profession. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching*. Heinemann Macmillan.

Chapter 6
From Needs Analysis to Task-Based Design: Methodology, Assessment and Programme Evaluation

Roger Gilabert and Aleksandra Malicka

Abstract In this chapter, we look at the interface between needs analysis (NA) and syllabus design, with a special focus on the methodological implementation of task, language testing and programme evaluation. The goal of the chapter is twofold, since it both analyses and suggests ways in which NA may be used to inform crucial decisions in syllabus design. After a short review of how NA may contribute to task selection, sequencing and pedagogical design, the article analyses how information coming from multiple NA methods (e.g. interviews, observations or the linguistic analysis of samples) may inform decisions about the implementation of methodological principles (e.g. 'learning by doing', fostering inductive learning, providing authentic input, promoting autonomy or respecting learners' internal syllabuses) and pedagogical procedures (e.g. motivating, modelling, providing pre-emptive focus on form). The chapter then zooms in on how NA may be instrumental in making decisions about testing at the level of interactional, psycholinguistic and cognitive demands of assessment tasks. It also looks at how information gathered in NA can aid in determining the technological dimension of tasks as well as establishing performance standards on these tasks. Finally, the study also associates NA with aspects of programme evaluation, and it looks at how the performance and linguistic standards detected by NA may inform evaluation criteria for external reviewers of language programs. The chapter concludes with a reflection and future directions of the interface between NA and syllabus design.

Keywords Need analysis · Task-based methodology · Task-based assessment · Task-based programme evaluation

R. Gilabert (✉)
Universitat de Barcelona, Barcelona, Spain
e-mail: rogergilabert@ub.edu

A. Malicka
Universitat Oberta de Catalunya, Barcelona, Spain

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_6

Introduction

For over three decades, we have seen claims that language should be better taught by focusing on what learners need to do with language rather than by teaching them language as an object. Anchored in the communicative tradition of the teaching of second and foreign languages and heavily influenced by second language acquisition (SLA) findings, task-based language teaching (TBLT) has suggested that tasks should be the organizing principle of any language programme design for a number of reasons. Tasks are driven by meaning, they have a predefined outcome, they engage learners in cognitive and communicative processes to achieve specific goals and outcomes, and they are susceptible to pedagogic intervention and reasoned sequencing. Since the first theorizing steps (Long, 1985) emerged, and since the first task-based programme implementation known as the Bangalore Project (Prabhu, 1987) was launched, researchers, syllabus designers and teachers have wondered what tasks should be included in a language programme, and in what shape, and how they should be organized, taught and assessed. Since the 1990s and more centrally in the 2000s, task-based needs analysis (NA) has been proposed as an area of inquiry within TBLT in order to assist all aspects of second/foreign language programme design (Long, 2005; Serafini et al., 2015). While significant importance has been given to what needs analysis is and how it should be conducted, considerably less reflection has been devoted to how what we learn from NA should be used to make decisions about such key aspects as the selection, design, sequencing, teaching or assessment of tasks. As we suggested in Gilabert and Malicka (in press), even if NA may be challenging in some contexts[1] to the point of it being close to impossible because of the resources (individuals and time) needed to carry it out, a careful analysis of needs of any language community may have enormous benefits for all areas of syllabus design.

In this chapter, we tackle just that, the connection between NA and task and syllabus design, and we suggest ways in which NA information can be used in several areas of programme design. We do so by first defining what NA is and the areas it covers. We then summarize how information from NA has been claimed to be useful in informing task selection, pedagogic design and task sequencing.[2] The focus here is not on how NA can aid these stages of TBLT programme design in

[1] In some contexts, institutions with a shortage of teachers may not be able to liberate teachers for a few hours over several weeks in order to carry out the NA. In other institutions, they may get such different populations from year to year, requiring completely different courses, that it may render NA unworthy.

[2] One of the reviewers rightly asked whether task selection, pedagogic task design and task sequencing are going to be aligned to specific levels and/or abilities. Our answer is that in task-based course design selection, pedagogic design, and sequencing are not necessarily associated with levels, but with the complexity (both cognitive and linguistic) of each task (see Baralt et al., 2014). The levels and linguistic abilities that learners bring to the task my push designers to break down tasks into more manageable versions in a sequence at lower levels, and fewer but more demanding versions in a sequence at higher proficiency levels. (Baralt et al., 2014; Gilabert & Castellví, 2019; Gilabert & Malicka, in press). There is an ongoing debate about this in the field.

the context of a particular language or linguistic competencies; rather, the idea is to establish synergies between the affordances of NA and specific components of TBLT language programs. After that we address three other central areas involved in syllabus design: task-based methodology, task-based assessment and programme evaluation.

Defining Needs Analysis

Needs analysis is a thorough empirical investigation of learner needs in an occupational, social, academic or professional context (e.g. see Long, 2005; Serafini et al., 2015). A needs analysis may be commissioned by different types of entities, both public (e.g. government departments, public educational institutions, the military) and private (e.g. companies, institutions, organizations, private educational institutions). Given the resources required to carry it out, it is usually not done by individuals but by groups. The exact stakeholders involved in doing an NA may vary depending on the objective of the NA. For example, if the objective is to develop a task-based language programme for a particular student population, involved parties typically include experts in that particular domain, course/curriculum designers and/or applied linguists and teachers. When the objective of NA is to evaluate a language programme, insights from such parties as course designers, applied linguists or heads of department would be taken into account. Finally, if an NA sought to identify the technological needs of a specific learner community (in addition to their language needs), experts in information technology could play an important role. Although we have not accrued enough programme evaluations to show a direct link between NA and improved learning outcomes, NAs from an array of programs have reported benefits such as saving time and resources, better and more targeted adaptation to actual needs or avoidance of unnecessary aspects during course design as a result of engaging in NA (see Serafini et al., 2015 for research synthesis on NA studies).

With ‘task’, rather than language, being the central organizing principle in this endeavour, NA employs different methods and sources to identify, examine and describe the tasks and sub-tasks (Gilabert, 2005) a specific population of learners should be able to do in a foreign or second language. Some learner communities are narrow and others broad in scope, and so are the tasks they perform and the language associated with these tasks. To give a few examples, a NA may be conducted for communities such as university administrative staff in an international exchange programme, immigrants to the USA who do the naturalization interview, Japanese students spending a semester abroad in London or refugees in camps awaiting resettlement, among others. The end goal of a NA is to lay a foundation for language programs that target teaching students exactly what they need to learn, moving away from an arbitrary generic modus operandi in language learning instantiated by more traditional synthetic syllabi (Long & Robinson, 1998; Wilkins, 1976). The more specific the community, the more specific the tasks and the language associated with those tasks that learners need to master. Given the often limited time and resources

available to many communities of learners, NA is a conceptually and methodologically sound way of discovering what tasks learners need to do in an L2, as well as the language associated with them, in the development of a language programme.

The Role of NA in TBLT Programme Design

A lot of scholarly interest in NA has revolved around discovering what needs a particular language community has. To give just a few examples, some of the NA reports have informed us of what target tasks are typically done in specific language user domains (e.g. Lambert, in press; Oliver, in press; Toker & Sağdıç, in press). This is a much-needed endeavour in and of itself. However, given the conceptual and empirical advances in the TBLT domain observed in the last twenty years on one hand, and on the other hand a sustained engagement in carrying out NA to find out about particular learners communities' needs, we have seen disproportionately few conceptual and empirical investigations targeting the synergy between the two. In other words, to date there has been little speculation about how the information obtained through NA can inform key decisions at different stages of a TBLT programme design: task selection and sequencing, pedagogical design, task-based methodology, assessment and programme evaluation. Some of the relevant yet so far both theoretically and empirically unaddressed connections in this area include, but are not limited to:

- How should we use the information obtained in NA to *select tasks for the curriculum?* (E.g. Should frequently performed tasks be given priority in the design of a TBLT program? Should infrequent tasks be included at all? What about infrequent yet indispensable tasks?);
- How can insights from NA aim the process of *designing pedagogic versions of target tasks* (E.g. How can interactional, cognitive and linguistic demands be incorporated into pedagogical design?)
- How can NA assist in taking decisions about *the sequencing of pedagogic tasks*? (E.g. According to what criteria are target tasks sequenced? What factors make target tasks simple or complex? Do these factors occur separately or simultaneously in target tasks? What does a sequence of tasks mean in a particular professional domain, and how can sequences of target tasks be translated into sequences of pedagogic tasks?)
- How can insights from NA inform task-based methodology? What *methodological principles* are particularly relevant to each task? (E.g. Does NA suggest whether task methodology should promote bottom-up, implicit, inductive learning rather than top down, explicit and deductive learning?)
- How can task-based assessment benefit from information gathered from NA? (E.g. What criteria constitute an acceptable performance of a task?)
- How can the information gathered during NA be used in programme evaluation? (E.g. What can NA say about the criteria for evaluation of task procedures, cognitive, linguistic, or ICT skills?)

In Gilabert and Malicka (in press; outlined below), we discussed the potential contributions of NA to the first three of the above-mentioned areas (task selection, task design and task sequencing). The current paper aims to illuminate how NA can inform decisions related to TBLT methodology, assessment and evaluation. With these two contributions, we aim to push the NA agenda forward by establishing connections between the information obtained through NA and different stages of TBLT language programme design.

The Usefulness of NA for Task Selection, Task Design and Task Sequencing

In Gilabert and Malicka (in press), we pointed out some of the challenges and advantages of NA in relation to task selection, pedagogic task design and task sequencing. In this section, we will briefly discuss the main conclusions.

First, Gilabert and Malicka (in press) identified general issues in NA meriting examination and pointed out how such areas may be helpful in decisions regarding the selection of tasks to be included in a language curriculum. In our view, factors such as target tasks' frequency, difficulty or the need for training are crucial when it comes to selecting tasks. Before priority is given to administering the most frequent and typical tasks or those reported as most difficult, it is worthwhile starting off by consulting domain experts on the dimensions mentioned above and triangulating the results with another method, such as large-scale surveys. Also, in order to achieve an alignment between insights obtained from NA and specific characteristics of a language programme, course/syllabus designers should take account the design of tasks and how they may potentially be sequenced.

Second, Gilabert and Malicka (in press) discussed how insights from NA can be programmatically employed to the design of pedagogic tasks. In this area, they identified NA as potentially informative in the following aspects:

- interactional features (the number of participants, participant status, rules of interaction, participants' attitudinal values, concepts and norms, as well as psycholinguistic aspects);
- cognitive demands (attentional and memory demands, high- and low-order thinking skills, task difficulty);
- linguistic features (receptive/productive skills, specific vocabulary or grammatical structures, discourse features, among others).

This information is gathered in Table 6.1, and information about NA in Table 6.1 will be linked with the different areas under inspection in this paper (i.e. methodology, testing and programme evaluation).

Finally, Gilabert and Malicka (in press) shed light on how NA may serve as a basis for the sequencing of tasks in a curriculum. The information obtained about the cognitive complexity of tasks and perceived difficulty can be a conceptual starting

Table 6.1 From NA to task selection, pedagogic design and sequencing (Gilabert and Malicka, in press)

NA dimension		Dimension description	Example/information obtained through NA
1. General aspects of task	Task goal	What is the task's ultimate goal? *Associated information*: *Steps involved in task performance, task length, when performed, need for training*	Solving a problem, reaching an agreement, describing a product, process or service, etc *Associated information: In how many steps is a task performed, when it is performed and whether a task requires training or not*
	Task frequency	How frequently is a task performed?	Frequent versus infrequent tasks; tasks performed on an hourly/daily/weekly/monthly/annual basis
	Task outcome	What is the final product of a task?	A written report, a public presentation, a sale, an academic paper, a satisfactory review by a client or customer, etc.
	Topics or subject matter	What general and specific topics a task may cover?	A list of potential topics which may be tackled in tasks
	Target sub-tasks/sequence of procedures	Tasks that may be performed simultaneously or in a predetermined sequence (Gilabert, 2005), step-by-step description of the task (Long, 2015)	Whether the task is stand-alone or there are other parallel tasks accompanying it

(continued)

Table 6.1 (continued)

NA dimension		Dimension description	Example/information obtained through NA
2. Participants and interaction	Number of participants	How many participants are involved in a task?	Tasks performed individually, in pairs, in groups (small or large); across groups
	Participant status	What is the status relationship between task participants?	Participants in equal positions vs. different positions (low vs. high) within the organization, institution or company
	Rules of interaction	Set of accepted behaviours during interaction; *do's and don'ts*	The importance of listening, having control over the conversation, floor-taking rules, things to make sure happen, topics to avoid
	Participants' attitudinal values, concepts and norms	Individual characteristics participants bring to a task	Knowledge, educational training or concepts that are crucial to the successful performance of a task; being patient or sympathizing; aggressiveness, encouragement, or optimism, among others
	Psycholinguistic aspects	The way information is shared between participants and each participant's contribution to task (Pica et al., 1993)	One-way/two-way/multiple-way information flows, convergent or divergent goals, a single solution or multiple solutions, split or shared information
	Intercultural communicative competence aspects	Are there any intercultural components of the tasks that may be relevant to task design (East, 2012)	Tasks can be neutral or interculturally charged. They may require competence in negotiating differences appropriately using language, as well as relating effectively to the 'other'?

(continued)

Table 6.1 (continued)

NA dimension		Dimension description	Example/information obtained through NA
3. Physical space where task takes place	Spatial setting	The characteristics associated with the physical space in which the task takes place (Bosswood & Marriott, 1994)	Behind doors, open office space, meeting room, private online chat, an open online forum, at home, press conference room, fairgrounds, restaurant/café/bar, planes/trains/car, among others
	Psychosocial environment	Factors associated with psychosocial characteristics of the environment in which the task takes place	Noisy versus quiet; familiar versus unfamiliar physical space; culturally close versus distant setting; relaxed versus stressful environment
4. Tasks cognitive demands	Cognitive aspects of the task (Robinson, 2001; Robinson & Gilabert, 2007; Skehan, 1998, 2009)	Attentional and memory demands tasks place on those who perform them	The number of simultaneous elements/items involved in task performance Whether spatial or intentional reasoning are required Whether taking a perspective is required Whether tasks are performed under time pressure Whether familiarity with the task is important
	Higher and lower (Bloom et al., 1956) required to perform the task	Mental processes necessary for successful task completion	Sample lower-order skills: gathering info, classifying or summarizing information, etc Sample higher-order skills: establishing relationships and associations, hypothesis testing, judging, etc

(continued)

Table 6.1 (continued)

NA dimension		Dimension description	Example/information obtained through NA
	Degree of perceived difficulty of tasks Complexity factors	How easy or difficult is a task perceived to be by those who perform it	Perceived mental effort, difficulty, anxiety or stakes; Information about what factors or conditions make a task more or less complex (e.g. task features, available resources, time constraints, multi-tasking, interlocutors or external factors)
5. Tasks' languistic demands	Language associated with performing a task	Linguistic resources necessary for task completion	Skills: receptive or productive or both; Terminology: specific vocabulary items, expressions, idioms; other multiword units; grammatical features Phonology: features related to tone or intonation; Pragmatic/discursive moves, e.g. commands or requests; Other features: rhetorical devices, turn-taking; style and level of formality; language variation
6. Communication and technology	Technological/digital tools and platforms (González-Lloret, 2014; González-Lloret & Ortega, 2014)	Technology involved in performing a task and its influence on the task	How technology is built into task performance, how it potentially transforms the task in terms of difficulty, the level of digital literacy required from task participants

(continued)

Table 6.1 (continued)

NA dimension		Dimension description	Example/information obtained through NA
	Communication channels	What means are used to perform a task	– Face-to-face versus computer-mediated communication, – Verbally via videoconferencing versus by phone, via email – Conventional writing on paper or interactive online chats, among others
7. Other dimensions	Criteria for assessment	According to domain experts, what constitutes successful task performance?	Qualitative and quantitative indices/measures of task completion/performance
	Support during task performance (Mayer, 2009)	Documents, people and other resources one can resort to during task performance	Internet searches, specialized literature (text book, journal, report, technical manuals, visuals), human and online translators, colleagues
	Non-verbal aspects	Aspects of task which do not have to do with language	Dress code, body language, facial expressions, eye-gaze, gestures, distance from interlocutor, non-verbal expression of emotions

point when it comes to organizing tasks in a logical sequence. These guiding principles for sequencing can be complemented by tasks' linguistic demands which, while being secondary to the cognitive criteria, should not be ignored.

The authors concluded that by taking all three dimensions discussed here into account—task selection, pedagogic task design and task sequencing—makes it possible to design sequences of tasks ranging from decomplexified versions of target tasks all the way up to full-blown authentic approximations of them. However, TBLT programme design includes three more stages: methodology, testing and evaluation. In the current chapter, we will discuss how these three stages of TBLT programme design can benefit from needs analysis.

How NA May Inform Task Methodology

Once we have decided on what target tasks will be used as a starting point in the design of pedagogic tasks, as well as the number of tasks (how many pedagogics versions needed to prepare for each sub-task and/or target task) and their specific design parameters (what task design variables will be manipulated and modified during task design to maximize learning opportunities), as well as the sequence in which they will be presented to learners (in various combinations of cognitive complexity, perceived difficulty and linguistic demands), syllabus designers and teachers need to decide on how such tasks will be taught.

A first useful distinction is provided by Long (2015) who teases apart methodological principles from pedagogical procedures. As Ellis et al. (2019: 209) suggest: 'Methodological principles are language teaching universals, informing teachers of what should be done; pedagogical procedures are specific steps teachers follow in implementing the principles, specifying how it should be done'. Researchers have spoken in a continuum that goes from just an interest in universal methodological principles with less of an interest in pedagogical procedures that are claimed to vary across contexts (Long, 2015), to a more balanced merger of principles and procedures (Ellis et al., 2019) and those who are more concerned with mostly pedagogical choices (Prabhu, 1987; Willis, 1996).

But what can NA say about the implementation of the general methodological principles and pedagogical procedures when teaching through tasks? In the following section, we provide an answer to these questions.

How NA Can Aid the Implementation of General Methodological Principles

As we have seen, NA is a thorough inquiry into the kind of tasks learners in a community will need to be able to do in the target language of their choice. As

such, NA is coherent with the general methodological principles suggested by Long (2015) (Doughty & Long, 2003) from which we highlight the ones that can be more directly supported by NA. These include: (1) promoting integral education through 'learning by doing'; (2) using tasks, instead of text, as the unit of analysis; (3) providing rich, elaborated and authentic input; (4) encouraging inductive learning; (5) encourage focus on form and providing negative feedback; (6) respecting learners' internal 'syllabuses' and developmental processes; (7) promoting cooperative and collaborative learning; (8) individualizing instruction and promoting autonomy.

Because task-based NA provides accurate and precise information about how tasks are meant to be planned and implemented, the exact procedures to be followed and what their expected outcomes should be (see Sect. 1 in Table 6.1), NA is coherent with the methodological principle of 'learning by doing' that supports integral education (Methodological Principle 1). Through the collection of samples in a variety of formats (e.g. texts, audio and video recordings of task performances and interactions, e-mail and social media samples, documents and memos and speech samples, among others), NA (see Sects. 5 and 6 in Table 6.1) can provide 'authentic input' that then may be processed, adapted, modified and elaborated upon during task design (MP 3). Not only that, the careful collection of different textual types will provide accurate and precise information about the linguistic resources needed to complete the task. This includes, but is not limited to, specific vocabulary items, expressions, idioms, formulaic pragmatic sequences, phonological and grammatical features, pragmatic and discursive moves, stylistic devices and other rhetorical devices. The analysis of such language items can aid task design in at least two ways: firstly, in the initial design of the task, it may provide the material for pro-active focus on form (Doughty & Williams, 1998). Pro-active or pre-emptive focus on form may need to be refined and adjusted with several iterations of the tasks, but at initial stages of design, it may inform decisions about what language to include in input flooding (Arteaga et al., 2003; White, 1998), or what linguistic items to highlight through either oral or written input enhancement (Cho & Reinders, 2013; Gascoigne, 2006; Lee & Huang, 2008) or input elaboration (Oh, 2011). It may aid with the decision as to what constitutes task essential language without which the task cannot be completed. Secondly, and during programme preparation for actual practice, it may even help practitioners advance and predict the difficulties learners may have with language and that will require corrective feedback (Kartchava et al., 2020) and get ready for them. Practitioners should be aware, however, that NA cannot possibly inform the exact decision on what difficulties to address and which one to skip and leave for later stages or when to address them (reactively as students perform the task? Post-actively when they have completed their task?) or how tow address them (e.g. do recasts work better for vocabulary than for pronunciation errors? Do short explicit explanations work for grammar or is 'eliciting' the grammar from students more effective? What technique works best with each student's learning style, preferences or personality?). Those decisions will have to be made locally and contingently by the practitioners, who are more likely to be knowledgeable of their own students, as tasks performance develops.

The complete and sharp task descriptions obtained by means of NA should be able to show the holistic character of tasks which may integrate and 'respect learners' internal syllabuses and processes' thus promoting learner autonomy and allow for the personalization of instructions (MP 6 and MP8 in Doughty & Long, 2003), since the richness in terms of interactive, cognitive and linguistic aspects that NA can potentially capture about each task (see Sects. 2, 4 and 5 in Table 6.1) will be in consonance with what each learner may be able to contribute to a task. Task descriptions springing from NAs may provide hints as to how learners may cooperate and collaborate as they learn (MP7) and may aid teachers with decisions about participatory structures during task implementation that may foster such cooperation and collaboration. Since up to date we have no research into what participatory structures (e.g. pairwork, small group work, whole class work) may have the most effective impact on learning with different task types or task phases, practitioners are typically left to their own resources when it comes to deciding how to organize learners during task implementation. Such decisions need again to be made locally and contingently and by incorporating the knowledge that teachers have of their class dynamics, students' learning styles, personalities and preferences. NA can provide detailed descriptions of procedures that may specify how domain experts may cooperate and collaborate with others in order to successfully complete each task (Sect. 2 in Table 6.1). Again, this initial implementation of the methodological principle can be aided by NA and adjusted and corrected after several iterations of the same task.

Finally, the implementation of any methodological principles must take into consideration the informational and multimodal digital skills required by technology-mediated TBLT programs, and as González-Lloret (2014) has pointed out, NA may and should be able to identify such requirements (See Sect. 6 in Table 6.1).

How NA Can Aid the Implementation of Pedagogical Procedures

In the same way that the implementation of general methodological principles can be aided by the information retrieved via NA, decisions about pedagogical procedures may also benefit from information obtained through interviews and observations during NA. As stated by Ellis et al. (2019), pedagogical procedures can and have been associated with task phases and here we take up such organization and analyse how NA may support the implementation for such pedagogical procedures. These pedagogical procedures include motivating students about the task, modelling and focus-on-form in the pre-task; in the main task phase, pre-emptive and reactive focus on form, within task planning and task support; in the post-task, task repetition, focus on form and mode.

How NA May Help with Pedagogical Options in the Pre-task

NA analysis can clearly contribute to motivating students about the task, by tapping their interests and raising their expectations (Ellis et al., 2019). NA provides an open window into the intricacy of tasks, their deeply rooted connection to reality, their contingency and contextualization, and as such, they can provide arguments for teachers to justify why a specific pedagogical task is needed in order to prepare for real target tasks. Additionally, they can use the same information to present the task procedures learners will be engaging in as well as the expected outcomes of the task. In our view, such information may engage learners in 'strategic planning', that is, the conceptual and linguistic preparation before the main task (see Ellis et al., 2019: 212–15 for a detailed description of how pre-task planning may impact language use and learning).

As for modelling, teachers have the option of imagining and modelling what a task may look like on the basis of their own intuitions or they can use task observations, descriptions and samples obtained during NA in order to motivate their modelling. Modelling can happen in at least five ways: either by showing them a direct sample of how domain experts may perform the task in actuality (if permissions are granted) (Long, 2015); by showing learners a performance from previous learners doing the same task once several iterations of the tasks have been achieved (Willis, 1996); or by asking questions and getting students answer in order to motivate their independent task performance (Prabhu, 1987); by creating a pre-recorded (video, audio, text) pedagogical version of the actual target task in which the input is modified (preferably enriched and elaborated on rather than simplified) and adapted to learners; by having teachers perform a simulation of the task with their students. The two latter options give practitioners the opportunity to have better control of the input they expose their learners to and hence may avoid any potential frustration at seeing domain experts successfully perform highly complex tasks removed from the learners' current abilities.

Regarding focus on form during the pre-task, opinions range from advocates of explicit and deductive focus on form by means of which learners are explicitly exposed to and explained vocabulary, grammar, into national patterns or pragmatic/discursive aspects of the language by their teachers (Littlewood, 2007; Shehadeh, 2012) to advocates of more implicit/incidental focus on form (Doughty & Long, 2003; Long, 2015). Although our leanings are towards the implicit/incidental approach to teaching, our goal here is not to enter this fruitful and engaging ongoing debate about the best way to teach language in preparation for task performance. The issue of how to best match linguistic features, pedagogical options under specific task conditions and learners' individual cognitive, linguistic and technological abilities in order to optimize learning is an unresolved one in TBLT, teaching in general, and second language acquisition as a whole for that matter. Our concern in this chapter is not with the available and more efficient pedagogical options that exist [see recent publications by Long (2015) and Ellis et al. (2019)], but rather with how NA may assist pedagogical choices by teacher if at all. Put bluntly, NA does not say much

about what pedagogical options may be best suited for each linguistic feature or dimension. Through domain expert reports, NA provides information about what cognitive, interactive, sociocultural, linguistic and technological abilities may be required for the successful completion of each task, but it says nothing about any abilities that learners may bring to the task. The linguistic analysis (Long, 2005, 2015) following the collection of samples during NA analysis, however, may shed light on how language may be used in contextualized task performances, in contingent and often specific ways, and the consequences this may have for successful or unsuccessful communication. As such, the information retrieved from the linguistic analysis of discourse samples obtained during NA may inform whichever pedagogical options teachers choose to employ in the continuum between explicit, deductive ones (explicit explanations) to more implicit, inductive ones (input flooding, input enhancement, input elaboration, task essential language).

How NA May Help with Pedagogical Options in the Main Task

The second phase in task instruction is that of the main task (Ellis et al., 2019) or task-cycle (Willis, 1996). We saw before that focus on form may be applied pro-actively or pre-emptively and NA analysis may be instrumental in providing the language focus that tasks may require (Sect. 5 in Table 6.1). On the contrary, NA may not help with predictions as to what reactive focus on form (corrective feedback) may be needed since such focus on form will largely depend on the learners' internal syllabus and their developmental readiness for different aspects of the target language. As was suggested before, neither does NA say much about the type of corrective feedback that will be more beneficial in each case.

As for within task planning, again NA may only indirectly provide hints about the length that each task may require in terms of preparation (Ellis et al., 2019). The information contained in task descriptions, interviews and observations during NA analysis may help teachers with initial task performance time allocation. This of course will need to be adjusted with subsequent repetitions of the tasks until a reasonable within planning time span can be determined. In the same fashion, NA can provide some initial information about the kind of support that domain experts may have during task performance (Sect. 7 in Table 6.1). Such information can be partially taken up from NA analysis and recreated during the methodological implementation of the task.

In addition to more or less direct information about focus on form, within task planning and degree of support, NA can also be a source of inspiration for the pedagogical recreation of real target tasks. Conditions related to the physical space and the psychological environment of tasks (Sect. 3 in Table 6.1) may inspire and facilitate the design of task conditions and materials in order to approximate the pedagogical simulation to actual task performance (e.g. preparing a detailed scenario for a hotel reception task with multiple actions—calls, guests showing up, a small emergency—happening at the same time or the noisy conditions of a business transaction in a public market).

How NA May Help with Pedagogical Options in the Post-task

NA cannot possibly help with what may be pedagogical options in the post-task phases. By determining the cognitive demands, difficulty and linguistic demands of a task (Sects. 4 and 5 in Table 6.1), NA may assist practitioners in predicting whether task repetition will be needed in the post-task for the consolidation and improvement of task performance. But it will be the quality and the language learning affordances of a first performance of the task that will typically help teachers decide if the exact repetition of the task or procedural repetition are necessary (Kim & Tracy-Ventura, 2013).

As for focus on form in the post-task, NA may only be instrumental in providing information about most linguistic aspects (Sect. 5 in Table 6.1) that need to be focused on, and it may serve the purpose of elaborating or enriching some of the language foci introduced pro-actively during task design. NA cannot predict which errors learners will be making and how to best solve them in terms of post-task feedback and practice. It may, however, help with modelling if for example a successful performance by domain experts collected during NA (Sect. 1 in Table 6.1) is shown to learners and analysed with them in order to improve their performance in the future. NA will not help with self-reflection on task performance. NA information may also be used to determine whether mode may contribute to more careful focus on form (Gilabert et al., 2016) by either transforming a task into the written mode and hence increasing the language learning potential of the task (Manchón, 2014).

We would like to conclude this section by saying that as far as methodological decisions are concerned, NA may be both a source of information, by providing direct information for the task's methodological and pedagogical implementation (e.g. vocabulary, grammar or pragmatics needed for the completion of the task), and inspiration, by providing hints as to how some aspects of tasks may be methodologically and pedagogically applied (e.g. the analysis of discourse samples obtained through needs analysis may provide hints to teachers as to some of the difficulties their students may encounter as they perform the task, hence inspiring teachers to get ready for potential episodes of reactive feedback).

The Role of NA in Task-Based Language Assessment

Language assessment, alongside pedagogy and research, is one of the areas in which tasks have been theorized and empirically investigated. There has been a lot of interest in using communicative tasks for language assessment, which is often referred to as 'task-based language assessment' (TBLA). TBLA is 'the elicitation and evaluation of language use (across all modalities) for expressing and interpreting meaning, within a well-defined communicative context (and audience), for a clear purpose, towards a valued goal or outcome' (Norris, 2016: 232). The attractiveness of tasks for testing contexts lies in the fact that what tasks test is not *the knowledge of a language*, typically associated with traditional approaches to language pedagogy and assessment,

but rather they test a *speaker's ability to put language to use.* More specifically, following Norris et al. (2018), the interest in using tasks as assessment instruments stems from the following affordances of tasks: they focus on what L2 users can actually do with the language, they integrate language form and meaning into what gets assessed, they provide useful feedback to teachers and learners, they 'wash back' on teaching and learning by making outcomes 'real', they counter negative influence of traditional knowledge assessment, and they align testing with new, evolving language pedagogies.

Norris (2016) reviews four broad domains in which task-based assessment is applicable: (1) *tasks as standard*s (i.e. establishing learning or ability standards in a foreign language, which is done either by governmental institutions or professional bodies), (2) *tasks in proficiency assessment* (e.g. university admission testing), (3) *tasks for employment certification* (e.g. a candidate for a position demonstrating job-related language skills) and (4) *tasks for language education assessment* (tasks as pedagogic instruments used in broadly understood language education classes and programs). Associated with these different uses of tasks in assessment contexts are different purposes for which assessment takes place. There are assessment contexts in which the objective of TBLA is to support and maximize learners' performance, as is the case in the majority of language education assessments and more specifically in classroom-based formative assessment. On the other end of the spectrum is the objective to discriminate across proficiency levels, for example when verifying learners' linguistic readiness for spending a semester of studies at a university abroad.

We believe that of the four scenarios described in Norris (2016) in which task-based assessment applies, two are particularly relevant to the chapter at hand: tasks as standards and tasks for employment certification. What these two contexts share is that both are concerned with determining the language needs/abilities/competencies of particular communities of speakers. These abilities are then taken as a point of reference when it comes to establishing target tasks, and language requirements associated with them, that these speakers must demonstrate their ability in. For instance, this can take the form of demonstrating one's skills as part of the selection process for a particular position, or in one's position as an employee during employee quality assessment. Given that needs analysis is concerned precisely with that—identifying needs of learner communities which share a common goal in language learning—there seems to be a synergy between NA and TBLA insofar as insights obtained in NA can be systematically applied to the two aforementioned testing contexts. In this section, we will explore these potential synergies further at different levels: tasks' general characteristics, their internal features (interactional, psycholinguistic, cognitive), the technology used in tasks and measuring task performance.

The Role of NA in Identifying Tasks' General Characteristics

At the core of needs analysis is the idea of identifying real-life tasks relevant to a particular community of learners of a second or foreign language. NAs typically

reveal elemental information about these tasks such as task goal(s), task topics and subtopics, the sequence of procedures tasks typically follow and potentially also secondary tasks associated with main tasks (Sect. 1 in Table 6.1). In this sense, insights gathered from NA can inform TBLA when it comes to determining the scope of tasks relevant for a particular community of learners in a particular testing context. First, NA provides assessment task designers with insights about what tasks learners should be tested on to begin with and the range of general and specific topics such tasks should cover. Once the scope of tasks has been defined, NA may help determine the steps involved in the performance of an assessment task, as well as the order of these steps. Additionally, NA-derived information about task goal(s) is instrumental in determining the outcome of an assessment task (e.g. producing a written article). For example, a journalism task 'publishing an article in an online venue' may involve these steps: skimming a written resource, listening to a news story, note-taking, drawing up a first draft, writing a final draft and publishing the final draft article on a newspaper's website. Designing testing tasks with the above considerations in mind holds the potential to ensure an alignment between the competences/abilities under scrutiny in the testing context and those encountered in a real-life task. There is little use in testing learners on tasks, and abilities or competencies associated with them, that are not relevant to real-life situations in which test-takers will need to use the L2 for professional purposes.

Once the topics, subtopics and goals of tasks have been determined, the internal architecture of these tasks must be further defined. The fundamental questions associated with the design characteristics of an assessment task are: (1) What *interactional pattern(s)* does the task impose on the test-taker?, (2) What *psycholinguistic features* are associated with an assessment task? and (3) Under what *cognitive conditions* is the assessment task performed?

The Role of NA in Determining the Interactional Dimension of Assessment Tasks

In terms of *interaction*, NA can provide assessment task designers with data such as: the number of participants in a task, the relationship status between participants or the rules of interaction that should be followed in a specific context (Sect. 2 in Table 6.1). Effort should be made to translate these interactional features into the design of assessment tasks so that they simulate the interactional set-up and demands of authentic tasks. The information about the number of participants in a task can be helpful in developing assessment tasks that range in interactional patterns from (1) tasks that do not require interacting with another person, i.e. *monologic* tasks (e.g. a two-minute sales pitch on a company's new product), to (2) tasks done *in pairs* (e.g. information exchange: identifying differences between an existing and a new product), to (3) tasks done *in groups* (e.g. three-way decision-making on the optimal advertising channel for the new product), to (4) tasks done *across groups*

(e.g. multiple-way decision-making: debate between two sections of an advertising agency on the best strategy to launch the product internationally). Assessment tasks should also be designed so that they incorporate the rules of interaction associated with authentic tasks. This can take the form of establishing conversational norms for a particular task that test-takers should obey, evidenced by successful identification of times when a testee should listen to their interlocutor(s), floor-taking rules, interrupting the interlocutor, etc.

The Role of NA in Determining the Psycholinguistic Dimension of Assessment Tasks

Regarding those assessment tasks which do require interaction between test-takers (i.e. dialogic/group/across group tasks mentioned above), needs analysis can further aid in shaping the *psycholinguistic dimension of assessment tasks* (Sect. 2 in Table 6.1), which encompasses two aspects: the way information is shared between participants and each participant's unique contribution to a task. There are three ways in which this dimension of tasks can be materialized in the context of assessment tasks. Following the categorization of pedagogic tasks proposed by Pica et al. (1993), broadly speaking assessment tasks can fall into one of the following psycholinguistic scenarios:

(1) A task requires the examinees to work jointly towards the same goal, or each examinee works towards their own specific goal set by the task (convergent vs. divergent assessment task). A convergent assessment task could be joint decision-making with the final goal of reaching a consensus, whereas a task in which each examinee defends their proposal for receiving funding is an example of a divergent assessment task.

(2) Examinees have access to the same information, or each examinee holds only part of the information needed to complete the task (assessment tasks with shared vs. split information). For example, a task in which the examinees exchange opinions about a topic on the basis of visual support which is available to both of them is a shared information exchange task. By contrast, an assessment task in which one examinee describes the route from point A to point B and the other examinee has to draw the route is an example of a split information task.

(3) Examinees work towards discovering the correct solution to the task (closed task) or they work towards one of multiple acceptable solutions (open task). A closed assessment task involves, for example, discovering the right path to get somewhere on the basis of the other test-taker's indications (otherwise the completion of the task is compromised). A 'debate' is an example of an open assessment task in that it invites the test-takers to exchange opinions.

The Role of NA in Determining the Cognitive Dimension of Assessment Tasks

When it comes to the *cognitive dimension of tasks*, there are two fundamental ways in which the insights from needs analysis can be employed in the design of assessment tasks: task characteristics and task conditions (Sect. 4 in Table 6.1). *Task characteristics* refers to the design attributes or parameters of tasks that are subject to be manipulated by the task designer. It includes, for example, narrating a story in the present versus in the past, taking the first or third person perspective when narrating a story, or referring to few versus many characters. *Task conditions*, on the other hand, refers to the circumstances surrounding the performance of a task such as whether a task is performed with the provision of planning time or without it [for a full list of task characteristics and conditions, see Robinson and Gilabert (2007)]. Following from this, an assessment task can be designed so that it contains more or less of a particular task characteristic and/or task condition. A task's cognitive load can be determined on the basis of how many design characteristics it includes, how many conditions it involves (and their difficulty) and whether these characteristics and conditions appear in a simultaneously or not (see Malicka et al. (2017) and Malicka (2018), for an example of how cognitive insights obtained in NA were used to develop tasks of varying levels of cognitive challenge). By manipulating all these features, an assessment task can be designed so that it exerts a greater or lesser cognitive load on the test-taker, thus making a task simple or complex or falling somewhere on the simple–complex spectrum of cognitive possibilities.

The Role of NA in Determining the Technological Dimension of Assessment Tasks

Technology is present in multiple real-life tasks in many different domains (professional, personal and social), and technology is also increasingly present when learning a foreign language. It is because of technology's omnipresence in daily tasks within and across domains that TBLT scholars have called for the incorporation of the technological component of tasks into NA (González-Lloret, 2014, 2015; González-Lloret & Ortega, 2014) so that NAs track down not only real-life tasks performed in particular domains, but also broadly understood digital and informational needs associated with those tasks. More specifically, following González-Lloret (2014), apart from identifying tasks for particular student populations, NA efforts should be directed towards discovering *technological tool*s which mediate the communication, *digital literacies* (both learners' digital skills and those required to perform specific tasks), as well as learners' *access to technology*. Technology-mediated NAs are still an incipient avenue of research (e.g. see Decamps & Bauvois, 2001; González-Lloret, 2003; Schrooten, 2006; Yasuda, 2012). Despite scant research in this area, if information about technology used in tasks performed in professional domains is

collected in a rigorous and systematic way, it could consequently be implemented into language programme curricula and testing contexts alongside language tasks. This would result in testing having the twofold objective of measuring learner's ability to put language to practical use via tasks, as well as measuring their ability to put to practice the technological requirements indispensable for carrying out those tasks. We believe carefully designed NAs, which capture both language tasks and technological affordances, requirements and literacies associated with the tasks (see Sect. 6 in Table 6.1), have the potential to make testing more comprehensive and aligned with the complex and constantly evolving demands of the contemporary world. Simultaneously testing language and technology can provide an even more thorough picture of an examinee's skills, which we believe is particularly relevant for the domains of assessment tasks identified earlier in this section: tasks as standards and tasks for employment certification. Needless to say, the synergy of language and technology would require a rethinking of ways of measuring performance which would need to tap into both skills sets.

The Role of NA in Measuring Performance on Assessment Tasks

Many NAs have reported discovering performance standards associated with tasks as an NA goal. This is usually done by consulting with domain experts what constitutes 'good/successful task performance' in a particular context and task. This information is relevant for task assessment in that it can be the basis for developing assessment criteria for tasks used in testing contexts (Sect. 7 in Table 6.1). This can involve such endeavours as developing discrete items or families of items for a particular task that assessment should tap into (e.g. task completion, linguistic accuracy, fluency, etc.), devising representative and fair ways in which performance should be measured according to these items (e.g. qualitative and/or quantitative performance indices), guidelines on how criteria should be interpreted qualitatively (e.g. establishing bands), as well as the consequences of particular interpretations of task performance for the test-taker (e.g. eligibility for a particular position).

Assessment tasks that build on insights obtained from NA have the potential to mirror authentic situations and are therefore valid indices of candidate preparedness to deal with requirements of tasks encountered in real-life situations.

The Role of NA in Programme Evaluation

There is a considerably large literature on programme evaluation (Birbeck, 2010; Hashimoto et al., 2010; Hedberg et al., 2002; Huber & Harvey, 2013) and a much more limited literature on evaluation of task-based programs (González-Lloret, 2014;

Lopes, 2015). Programme evaluation seeks to assess the progress and quality of a project, as well as to evaluate its impact on teaching and learning processes. It does so by getting critical feedback on the implementation of the process, its procedures and management and its products or outputs. Whether more top-down or bottom-up, it typically involves both end users of the programme (student and teachers), academic experts in FL and educational experts, as well as professional evaluators. It provides an overall picture of the design, and it helps parties involved and external reviewers the opportunity to detect strong points as well as weaknesses. In doing so, it may also provide clues for improvement. As far as possible, programme evaluation may be used to validate outcomes and to measure the transferability of the language learned in the programme to other contexts outside the programme.

While NA is not concerned with how a programme may be evaluated, NA is definitely a window into a linguistic community's overall needs, and as such it can provide a backdrop against which to evaluate a language programme. Second, language task-based NA can help programme designers and teachers obtain a full picture of all important tasks that a community may need to perform in the target language in second language context or in a combination of first and target language in foreign language contexts, as well as their sub-tasks and interconnection with other tasks. It provides an overall map of the goals, procedures, topics, expected outcomes and relative frequency and difficulty of each task. External reviewers may use such descriptions to check the degree to which those expectations were met, and so NA information may be useful when building evaluation criteria and rubrics. Programme evaluation studies like González-Lloret (2014) or Lopes (2015) show that criteria for assessment and evaluation may be based on the rich and detailed information provided by NA (all sections in Table, but with a particular useful contribution of Sects. 1, 4, 5, 6 and 7) since, as Lopes (2015: 13) puts it, information includes: 'the detailed description of the task, including demonstration of the abilities to be acquired, type of product created, product requirements or prerequisites, situation or theme, process (the different steps of the task in chronological order), division of roles, consolidating activities suggested and success factors or assessment criteria'. Information obtained through NA about linguistic and language proficiency expectations and achievements (Sects. 5 and 7 in Table 6.1) and acquisition of ICT skills (Sect. 6 in Table 6.1) may be also used for evaluation purposes. Such information can be utilized to provide practical methodological and pedagogical suggestions for practitioners, indicating best practices, which may have a meaningful impact on teachers' attitudes and practices.

Conclusions

In this chapter, we have seen how NA may potentially contribute to all dimensions involved in syllabus design: task selection, pedagogical design, task sequencing, methodology, assessment and programme evaluation, and we have provided explanations of how NA may inform all those dimensions in meaningful and productive

ways. In our view, these connections justify even more the need for NA to be at the base of language programs, since every minute invested in NA will reduce the design workload and assist decisions in every aspect of syllabus design. We have also identified some areas that may not be directly informed by NA but that might be 'inspired' by information coming from NA (e.g. the conditions related to the physical space and the psychological environment of tasks that may be used during task recreation and simulation).

Regarding task-based methodology, we have seen that NA and its multiple methods such as interviews, observations and the analysis of discourse samples may support syllabus designers and teachers in carefully choosing methodological principles such as 'learning by doing', the provision of rich and authentic input and pre-emptive focus on form. The same applies to pedagogical procedures associated with the different phases of task implementation. NA may help motivate students about the task, contribute to modelling decisions and focus on form and task planning. In the main task, we have seen that NA may indirectly help predict the kinds of issues that may generate corrective feedback, and it may back initial decisions about within task planning. As far as the post-task is concerned, we have also seen how it may contribute to determining the linguistic aspects that may be targeted in the post-task in conjunction with linguistic issues springing from actual task performance by students. On the other hand, NA may not be able to provide any useful information regarding whether a task needs to be repeated or how exactly mode may be used to consolidate language aspects (e.g. by turning a spoken performance into a written text).

In the area of language testing, we have seen how the information gathered from NA can aid in determining the internal architecture of tasks. The contribution of NA to assessment task design has been tackled from three complementary perspectives: interactional, psycholinguistic and cognitive demands assessment tasks pose on the test-takers. We have also suggested that the insights obtained in NA can be helpful in determining the technological dimension of assessment tasks and for establishing criteria for measuring performance on tasks used for testing purposes.

Regarding programme evaluation, we have seen that NA may help provide a full map of all tasks and sub-tasks needed by a linguistic community, and it may help with the setting up of evaluation criteria by informing about performance and linguistic standards of tasks.

While this chapter has laid out some of the foundations of the interface and transfer from NA to syllabus design, certainly a lot more reporting from TBLT courses is needed in order for us to accumulate enough knowledge that may eventually help syllabus designers and teachers with the challenging but exciting endeavour of setting up a task-based language programme. No research to date has been conducted into the decision-making process of converting NA information into actual TBLT syllabuses and programs, and so a whole new area is open for exploration and research.

As a word of encouragement for practitioners and course designers, even if the transition from NA to actual task design may be seen as highly demanding, such a complex and intense process can be broken down into smaller components, so that teachers do not feel overwhelmed by the amount of training needed. Training in NA

and task design of course can happen prior to the implementation of NA to task design, but it may also be distributed over time and progressively increased with each step in programme design. From our experience in course and task design, we also believe that the higher the involvement practitioners in NA, the easier the incorporation of NA findings into task design and the more effective and well-informed actual teaching will be.

In our previous work on this area (Gilabert & Malicka, in press) and here, we have admitted to the fact that NA may sometimes not be easy to conduct, even to the point of impossible, since it requires a considerable amount of time, effort and institutional support for it to work, but we also believe that the enormous pay-off in boosting design and saving time is worth every minute of NA.

References

Arteaga, D., Gess, R., & Herschensohn, J. (2003). Focusing on phonology to teach morphological form 585 in French. *Modern Language Journal, 87*(1), 58–70.

Baralt, M., & Gilabert, R., & Robinson, P. (2014). An introduction to theory and research in task sequencing and instructed second language learning. In M. Baralt, R. Gilabert, & P. Robinson (2014) (Eds.), *Task Sequencing and Instructed Second Language Learning* (pp. 1–34). London: Bloomsbury Academic

Birbeck, D. (2010). Benchmarking and peer review of teaching practice for the purpose of evidencing excellence in teaching and learning. *Ergo, 1*(3), 5–10.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals. In *Handbook I: Cognitive domain.* David McKay Company.

Bosswood, T., & Marriott, A. (1994). Ethnography for specific purposes: Teaching and training in parallel. *English for Specific Purposes, 21*, 59–79.

Cho, M., & Reinders, R. (2013). The effects of aural input enhancement on L2 acquisition. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 133–148). University of Hawaii, National Foreign Language Resource Center.

Decamps, S., & Bauvois, C. (2001). A method of computer-assisted language learning: The elaboration of a tool designed for an 'un-schoolable' public. *Computer Assisted Language Learning, 14*(1), 69–96.

Doughty, C., & Long, M. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning and Technology, 7*(3), 50–80.

Doughty, C., & Williams, J. (1998). *Focus on form in classroom second language acquisition.* Cambridge University Press.

East, M. (2012). Addressing the intercultural via task-based language teaching: Possibility or problem? *Language and Intercultural Communication, 12*(1), 56–73.

Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2019). *Task-based language teaching: Theory and practice.* Cambridge University Press.

Gascoigne, G. (2006). Toward an understanding of incidental input enhancement in computerized L2 environments. *CALICO Journal, 24*, 147–162.

Gilabert, R. (2005). Evaluating the use of multiple sources and methods in needs analysis: A case study of journalists in the Autonomous Community of Catalonia (Spain). In M. Long (Ed.), *Second language needs analysis* (pp. 182–199). Cambridge University Press.

Gilabert, R., & Castellví, J. (2019). Task and syllabus design for morphologically complex languages. In J. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* [*Cambridge handbooks in language and linguistics* (pp. 527–549)]. Cambridge University Press. http://doi.org/10.1017/9781108333603.023

Gilabert, R., & Malicka, A. (in press). From needs analysis to task selection, design, and sequencing. In M. Amadian & M. Long (Eds.), *The Cambridge handbook of task-based language teaching*. Cambridge University Press.

Gilabert, R., Manchón, R., & Vasylets, L. (2016). Mode in theoretical and empirical TBLT research: Advancing research agendas. In a special issue edited by A. Mackey. *Annual review of applied linguistics* (*ARAL*).

González-Lloret, M. (2003). Designing task-based call to promote interaction: En Busca de Esmeraldas. *Language Learning & Technology, 7*(1), 86–104.

González-Lloret, M. (2014). The need for needs analysis in technology-mediated TBLT. In M. González-Lloret & L. Ortega (Eds.), *Task-based language teaching* (Vol. 6, pp. 23–50). John Benjamins Publishing Company.

González-Lloret, M. (2015). What is technology-mediated TBLT? In M. González-Lloret (Ed.), *A practical guide to integrating technology into task-based language teaching* (pp. 1–16). Georgetown University Press.

González-Lloret, M., & Ortega, L. (2014). Towards technology-mediated TBLT: An introduction. In M. González- & L. Ortega (Eds.), *Technology-mediated TBLT: Researching technology and tasks* (pp. 1–22). John Benjamins Publishing Company.

Hashimoto, K., Pillay, H., & Hudson, P. (2010). An evaluation framework for sustaining the impact of educational development. *Studies in Educational Evaluation, 36*(3), 101–110. Retrieved from: http://doi.org/10.1016/j.stueduc.2010.12.002

Hedberg, J., Wills, S., Oliver, R., Harper, B., & Agostinho, S. (2002). Developing evaluation frameworks for assessing quality ICT based learning in higher education. In P. Barker & S. Rebelsky (Eds.), *Proceedings of ED-MEDIA 2002, World Conference on Educational Multimedia, Hypermedia & Telecommunications.* Association for the Advancement of Computing in Education. Retrieved from: http://ro.uow.edu.au/cgi/viewcontent.cgi?article=1229&context=edupapers

Huber, E., & Harvey, M. (2013). Time to participate: Lessons from the literature for learning and teaching project evaluation in higher education. *Studies in Educational Evaluation, 39*(4), 240–249. Retrieved from: http://doi.org/10.1016/j.stueduc.2013.10.004

Kartchava, E., Gatbonton, E., Ammar, A., & Trofimovich, P. (2020). Oral corrective feedback: Pre-service English as a second language teachers' beliefs and practices. *Language Teaching Research, 24*(2), 220–249. https://doi.org/10.1177/1362168818787546

Kim, Y., & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? *System, 41*, 829–840. https://doi.org/10.1016/j.system.2013.08.005

Lambert, C. (in press). Fostering the integration of humanitarian entrants in Australia: A task-based needs analysis. In M. Ahmadian & M. Long (Eds.), *Cambridge handbook of task-based language teaching*. Cambridge University Press.

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*(3), 307–331.

Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching, 40*(3), 243–249.

Long, M. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hylstenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77–99). Multilingual Matters.

Long, M. H. (2005). *Second language needs analysis.* Cambridge University Press.

Long, M. H. (2015). *Second language acquisition and task-based language teaching.* Wiley Blackwell.

Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom SLA* (pp. 15–41). Cambridge University Press.

Lopes, A. M. (2015). Critical issues in the evaluation of an international project dedicated to technology-mediated TBLT (PETALL). *Teaching English with Technology, 15*, 4–18.
Malicka, A. (2018). The role of task sequencing in fluency, accuracy, and complexity: Investigating the SSARC model of pedagogic task sequencing. *Language Teaching Research*. https://doi.org/10.1177/1362168818813668
Malicka, A., Gilabert, R., & Norris, J. (2017). From needs analysis to research tasks: Insights from an English for specific purposes context. *Language Teaching Research, 21*, 1–29.
Manchón, R. M. (2014). The internal dimension of tasks: The interaction between task factors and learner factors in bringing about learning through writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning—Insights from and for L2 writing* (pp. 27–53). John Benjamins.
Mayer, R. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.
Norris, J. (2016). Current uses for task-based language assessment. *Annual Review of Applied Linguistics, 36*, 230–244.
Norris, J., Sasayama, S., & Malicka, A. (2018). Meta-analysis of task complexity effects on L2 performance: Implications for assessment. Paper presented at the Conference European Association for Language Testing and Assessment, Ruhr-Universität Buchum, Germany.
Oh, S. Y. (2011). Two types of input modification and EFL reading comprehension: Simplification versus elaboration. *TESOL Quarterly, 35*(1), 69–96.
Oliver, R. (in press). Developing a task-based approach: A case study of Australian Aboriginal VET students. In M. Ahmadian & M. Long (Eds.), *Cambridge handbook of task-based language teaching*. Cambridge University Press.
Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction. In G. Crookes & S. Gass (Eds.), *Tasks and language learning. Integrating theory and practice*. Multilingual Matters.
Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press.
Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27–57.
Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *IRAL-International Review of Applied Linguistics, 45*(3), 161–176.
Schrooten, W. (2006). TBLT and ICT: Developing and assessing interactive multimedia for task-based language teaching. In K. Van den Branden (Ed.), *Task-based language teaching in practice*. CUP.
Serafini, E. J., Lake, J. B., & Long, M. H. (2015). Needs analysis for specialized learner populations: Essential methodological improvements. *English for Specific Purposes, 40*, 11–26.
Shehadeh, A. (2012). Broadening the perspective of task-based language teaching scholarship: The contribution of research in foreign language contexts. In A. Shehadeh & C. A. Coombe (Eds.), *Task-based language teaching in foreign language contexts: Research and Implementation* (pp. 1–20). John Benjamins.
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510–532.
Toker, Ş., & Sağdıç, A. (in press). A needs analysis for Syrian refugees in Turkey. In M. Ahmadian & M. Long (Eds.), *Cambridge handbook of task-based language teaching*. Cambridge University Press.
White, J. (1998). Getting learners' attention: A typographical input enhancement study. In C. Doughty & J. Williams (Eds.), *Focus-on-form in classroom second language acquisition* (pp. 85–113). Cambridge University Press.
Wilkins, D. (1976). *Notional syllabuses*. Oxford University Press.
Willis, J. (1996). *A framework for task-based learning*. Longman.
Yasuda, S. (2012). *The implementation of genre-based tasks in foreign language writing instruction: A longitudinal study of writers' rhetorical awareness, writing quality, and lexicogrammatical choices* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses.

Chapter 7
Differentiating Task Repetition from Task Rehearsal

Gavin Bui and Rhett Yu

Abstract Asking second language (L2) learners to repeat the same or slightly altered tasks is a common practice in task-based language teaching (TBLT). Prior research suggests that, when properly designed, task repetition can induce natural re-occurrence of a task interesting to learners, hence less fatigue and boredom than dry rehearsal. Repeating a task has also been associated with heightened L2 performance in previous studies. What remains inadequate in the field, however, is an ill-defined construct of task repetition as it appears to be often construed synonymously with task rehearsal. This chapter is a response to Bui's (Processing perspectives on task performance. Benjamins, Amsterdam, the Netherlands, pp. 63–94, 2014) call to differentiate rehearsal from task repetition as two different constructs, with the former involving learners' awareness of future performance and the latter shunning such forewarning. Then a mini-meta-analysis of prior task repetition and rehearsal studies is presented to tease out the differentiating effects of the two constructs on L2 complexity, accuracy, lexis and fluency performance. Based on the patterns identified from the available information in those studies, theoretical and pedagogical implications are discussed.

Keywords Task-based language teaching · Task repetition · Task rehearsal · Task readiness · Task performance · CAF

Introduction

Task-based language teaching (TBLT) has come under the spotlight as a language pedagogy as it prioritizes meaning over forms, emphasizes concrete outcomes that require language use and stresses real-world relevance for transferrable skills. An important feature of TBLT lies in its nature as a researched pedagogy in which psycholinguistics plays a central role but other relevant theories (such as sociocultural theories) also contribute to its maturation and fruition. Among the different strands of

G. Bui (✉) · R. Yu
Department of English, The Hang Seng University of Hong Kong, Siu Lek Yuen, Hong Kong
e-mail: gavinbui@hsu.edu.hk

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_7

research in TBLT, various types of task planning have attracted great attention in the field. Ellis (2005, 2009) broadly defined three types of planning: rehearsal, (pretask) strategy planning and online planning; he cited a wide range of prior research to show that these three planning conditions prepare second language (L2) learners for the task in different but often complementary ways. For example, strategic planning is usually found to benefit linguistic complexity and speech fluency while online planning could raise grammatical accuracy. The effects of different types of planning on L2 performance undoubtedly provide second (L2) or foreign language (FL) teachers with more insights into their classroom practice.

While Ellis' (2005, 2009) taxonomy has become a standard in planning studies, Bui (2014) pointed out two limitations. On the macrolevel, the scope of planning as preparation for task performance is somewhat limited. He argued that content familiarity and procedural familiarity, for instance, serve as implicit 'preparedness' for enacting a task. Bui termed the various forms of familiarity with the content, the task and the procedures as 'task-internal readiness' because they are inherent within the learners and require no additional planning time. In contrast, the three types of planning in Ellis (2005, 2009) are 'task-external readiness' as they are extra preparation opportunities. On the micro-level, Bui contended that Ellis and other researchers have not differentiated task rehearsal from task planning, which have been typically construed synonymously in the literature. In a series of papers (Bui, 2014; Bui & Huang, 2018; Bui & Teng, 2019), Bui proposed to differentiate task rehearsal from task repetition on a conceptual level. This paper will further examine empirical evidence in the TBLT literature and attempt to arrive at some preliminary conclusions as to how task rehearsal and task repetition exert differentiating effects on L2 speech complexity, accuracy and fluency (CAF or CALF to include lexis, See Bui & Skehan, 2018; Housen & Kuiken, 2009; Michel, 2017 for reviews) and why they should merit distinct statuses as task preparedness (Ellis, 2019).

Conventional Views of Task Repetition/Rehearsal

Rehearsal, although defined as 'a special type of pre-task planning…[in which the] performance of a task at one time can be seen as providing planning for performance of the same task at a second time' (Ellis, 2005, p. 476), has been simply treated as a form of task repetition where the first performance transfers certain skills to the next. As mentioned above, task rehearsal and task repetition have been used as exchangeable terms in the TBLT research. This equivalent view, however, oversimplifies what rehearsal implies and how it might impact the learners. The following subsections attempt to differentiate the two constructs.

Theoretical Underpinning of Repeating a Task

The effects of repeating a task on L2 performance or acquisition have been informed by numerous theoretical perspectives. One of the most influential theoretical bases cited in the discussion of task repetition is Levelt's (1989) psycholinguistic model of speech production. This model includes three stages: conceptualization, formulation and articulation. Conceptualization sets the goal of the speech and prepares the speaker with non-linguistic content. During formulation, the speaker chooses appropriate lexical items and a syntactic frame to map onto the preverbal message and creates a 'covert speech' with a phonological plan. At the final stage, articulation, the speaker coordinates motor mechanisms to produce an overt speech. Levelt stipulates that the conceptualization stage requires controlled processing for both native and L2 speakers as it is cognitive demanding to generate ideas on various occasions. While the remaining two stages prove to be automatic processes for native speakers, L2 speaking requires attentional resources which are often quite limited (Baddeley, 2003; Baddeley & Hitch, 1974; Skehan, 1998, 2014). Therefore, L2 speakers have to channel their attentional capacity to some but not all of the performance areas, such as complexity, accuracy *or* fluency. This is documented as the limited attention capacity (LAC) hypothesis (Skehan, 2014, p. 7; Skehan et al., 2012), or the trade-off hypothesis in Skehan's (1998) early term, which argues that joint focuses on complexity, accuracy and fluency in a task lead to cognitive overload for L2 learners; selected attention is the norm among L2 speaking. Skehan's theory makes an interesting contrast to Robinson's (2001) cognition hypothesis which states that higher task complexity would help guide learner attention to form, leading to jointly raised accuracy and complexity.

During the initial task performance, learners are more likely to focus on the conceptualization stage, e.g. generating ideas (Bui & Teng, 2018), especially when the learner focuses on meaning expression in a task. The repeated performance, in contrast, allows learners to shift their focus to formulation, articulation and even monitoring, as their cognitive capacity has been freed up with the prior planning of the content. It is important to note that while the literature has shown some agreement in the improvements to CAF as a result of task repetition, studies vary greatly in the findings of the effects on these three aspects, respectively. Indeed, few studies show simultaneous improvements in all three performance areas for learners under the task repetition condition (e.g. Ahmadian & Tavakoli, 2011; Bui et al., 2019; Wang, 2014), with most reporting enhanced complexity and fluency and mixed results on accuracy. As Ellis (2005) points out, 'if learners are viewed as having a limited processing capacity, they will find it difficult to attend to both complexity and accuracy and so will prioritise to one of these' (p. 502).

Another relevant theory to the benefit of repeating a task would be the skill acquisition theory (SAT), which has drawn from cognitive psychology (DeKeyser, 2007; DeKeyser & Criado, 2012). SAT stipulates that all skills can be acquired through similar stages by proceduralizing declarative knowledge to ultimately achieve automatization of the skill. The repetitive nature of task engagements allows learners to

proceduralize certain elements, for example, reoccurring phrases and grammatical patterns, and therefore releasing attentional capacity, performing increasingly effortlessly in subsequent repetitions and enhancing fluency and accuracy. Studies have compared the effects of exact task repetition and procedural repetition on CAF (e.g. De Jong & Perfetti, 2011; Fukuta, 2016; Kim & Tracy-Ventura, 2013; Lynch & Maclean, 2000; Patanasorn, 2010), and while the results are mixed, procedural repetition has been reported to enhance mostly accuracy and fluency. A caveat has to be made, though, that the number of repetitions in task literature is far from sufficient for genuine proceduralization of skills. Therefore, the performance of task repetition or rehearsal has to be thought of as an interim stage, somewhere along the continuum from a controlled to an automatic process, depending on the frequency of repetition, and thus, the progress, of practice.

Types of Repetition

Although the concept of repetition of a task is quite straightforward, there have been slightly different operationalizations of this construct. These operationalizations can be categorized according to the types, intervals and frequency of repetition.

Types of Repetition Condition

Patanasorn (2010) proposed three types of repetition, with different combinations of whether the content, the procedure or both are repeated in a task. Content repetition repeats the content of the task but not the procedure. Procedural repetition repeats the procedure of the task with different contents. Task repetition repeats both the content and the procedure of the task (i.e. an exact repetition). Earlier studies on the effects of task repetition (e.g. Bygate, 2001; Lynch & Maclean, 2000) on task performance were in the form of what Patanasorn (2010) called 'task repetition', by which she meant exact repeated tasks with both identical content and procedures. However, she later demonstrated that content repetition (same content, different task) and procedural repetition (same task, different content) improve learners' global proficiency and accuracy, respectively, but that task repetition did not show statistically significant enhancements in any aspects of language production. These results led her to believe that the repetition of a single aspect of the task may be more useful than offering a complete replica of the task. However, subsequent studies following Patanasorn's distinction have shown mixed results on the effects of different operationalization of repetition. For example, Kim and Tracy-Ventura (2013) found that both task and procedural repetition yield improvements in accuracy and syntactic complexity. Also, Fukuta (2016) studied the attention orientation of learners from task repetition and procedural repetition and showed that the former led to better performance in accuracy and lexical variety than the latter. These studies reflect that

the condition of repetition does not necessarily encourage learners to prioritize their attentional resources on specific aspects of the task.

Types of Repetition Intervals

Task conditions in the literature also differ in terms of time intervals between repetitions. Bui et al. (2018) categorize three broad operationalizations of repetition: immediate task repetition, intervaled task repetition and multiple task repetitions across long intervals.

An immediate task repetition typically requires a learner to perform the same task consecutively. For example, Lynch and Maclean (2000) asked L2 learners to give poster presentations six times to different interlocutors. The results indicated that intermediate learners performed better in terms of fluency and accuracy. In another study, Wang (2014) requested the learners to tell a story again to an imaginary listener immediately after narrating along with the source material (i.e. story shown in a video). The finding showed improvements in complexity, accuracy and fluency. Lambert et al. (2017) engaged the participants in aural-oral tasks for six times as repetitions and found that speech fluency was enhanced.

Previous studies also investigated intervalled task repetitions spaced one day to ten weeks apart. For example, Bygate (1996) administered a three-day intervalled task repetition and found positive effects on speech accuracy and fluency. Ahmadian and Tavakoli (2011) and Fukuta (2016) both studied the effects on CAF after repetition with a one-week interval but reported mixed results. While Ahmadian and Tavakoli (2011) found enhancements in all three areas of speech production, Fukuta (2016) only observed improvements in accuracy and lexical variety, adding that exact task repetition is more effective than procedural repetition. More recently, Bui et al. (2019) explored task repetition under five interval conditions between the initial and repeated performance. The EFL learners in their study performed a picture description task and repeated the same, unanticipated task with either no interval (immediate repetition), a one-day, a three-day, a one-week or a two-week interval. The results were that task repetition per se exerted a positive effect on L2 performance regardless of the length of intervals. The interval conditions appeared to 'mediate the effects of task repetition in terms of fluency and structural complexity with speed fluency benefitting most from immediate or small intervals between initial and repeated performances' (p.1). They also found that a one-week interval was the task condition most conducive to the improvement in structural complexity and repair fluency.

The final type of task repetition involves repeating the same task multiple times across an extended period, usually over a week or a whole semester (as regular class training). Bygate (2001) tested the effect of task repetition (content repetition) and task-type repetition (procedural repetition) under an experimental condition of multiple repetitions across a ten-week interval and found that speech complexity and fluency were improved in the repeated task performance after ten weeks. Gass et al. (1999) compared the effect of task repetition and procedural repetition on general

proficiency, accuracy and fluency. They discovered that, at an interval of two to three days, the third and final repetitions displayed improvements in general proficiency, accuracy of the Spanish variants of 'to be', morphosyntax, lexical density and lexical sophistication.

Frequency of Repetition

Studies of task repetition have also varied in the frequency of repetitions, i.e. the number of times one repeats a task. Most of the studies in the task literature have a frequency ranging from one (e.g. Wang, 2014) to six times (e.g. Lambert et al., 2017). Theoretically, there is no upper limit to the number of repetitions, but in reality, there is a concern for boredom and fatigue if one has to perform the same learning task multiple times.

Repeating a task one time often leads to improvement in performance, possibly due to an instant familiarization of the task content or the task type, which constitutes task-internal readiness (Bui, 2014). For example, Bui et al. (2019), Bygate (1996, 2001) and Wang (2014) all required their participants to only repeat the same task (or task type) once, and the findings all showed improvements in fluency and complexity (Wang even found an effect on increased accuracy). That might suggest that even one repetition is sufficient to invoke task or content familiarity, hence task-internal readiness, leading to an overall enhancement in performance.

Studies with multiple repetitions are more likely to discover gradual effects on CAF or even the trend of the change in CAF so as to determine the optimal number of repetitions for effective language learning. For example, Lambert et al. (2017) found that the speech rate of the participants improved markedly over the first three performances, but the improvement only lasted until the fifth performance. Moreover, they found that the frequency of overt self-repairs decreased in the fifth and the sixth performance. They argued that the participants' accuracy and efficiency in linguistic encoding had improved by then and therefore fewer self-corrections or reformulations were necessary. Their study also questioned the participants' perception of the numbers of repetitions, and of those who felt that five repetitions were not necessary; they reported that a repetition of three to four times is sufficient.

Problems with the Conventional Views of Task Repetition

As can be seen in the previous section, the operationalization of task repetition varies in terms of task type, repetition type, length of interval and repeating frequencies, and has therefore made it difficult to compare the results of different studies. In particular, there is a paucity of report on the awareness of the learners of future performances, as task repetition (involving unwitting learners) has been used synonymously with rehearsal (with informed participants). Ellis (2005), for example, reviewed articles

which he considered to be about rehearsal, with most of the cited studies (e.g. Bygate, 1996, 2001; Gass et al., 1999) having studied task repetition rather than rehearsal, as defined by Bui (2014) and Ellis (2019) himself. Bui (2014) highlighted the distinction between task repetition and task rehearsal, with the main difference lying in 'whether one knows if s/he is going to do the task again' (p. 67). In Bui's theoretical framework of task readiness, task repetition represents a form of implicit planning (or, in Bui's term, task-internal readiness), where learners can potentially benefit from topic and procedural familiarity; whereas task rehearsal offers explicit planning opportunities (or task-external readiness) for learners to practice for the next round of performance.

Unfortunately, as Ellis (2019) commented, 'Bui's (2014) suggestion that we should distinguish between 'rehearsal' and 'repetition' has not been acted on to date' (p. 17). Most studies in the literature (e.g. Bygate, 1996, 2001; Gass et al., 1999) do not make a distinction between task repetition and rehearsal, as the description of the procedures only includes information regarding task type, repetition type, length of interval and repeating frequencies, but never participants' knowledge of potential future repetitions. In other words, task repetition has been used synonymously with rehearsal in the literature.

This distinction, however, is a 'potentially important distinction' (Ellis, 2019, p. 18), as the pre-task awareness may direct learners' attention resources to the formulation stage (in Leveltian terms) in the planning of their subsequent speech performances, leading to more interesting observations on the effects on CAF (See Sect. 3). The following sections will review relevant research to identify possible differences in the effects resulting from a task repetition and a task rehearsal condition, respectively.

Differencing Task Repetition from Rehearsal on Performance

Given the qualitative difference between rehearsal and repetition, one would wonder 'which has a stronger influence on the improvement of task performance' (Bui, 2014, p. 67). This section discusses the possible effects of the two constructs on speech production.

The key difference between rehearsal and repetition lies in whether the learner is conscious of the task preparation. Therefore, to discuss their effects on CAF, it should be useful to look at how the presence of attention contributes to learning. Tomlin and Villa (1994) offered a fine-grained analysis of attention, where they divide attention into three parts: alertness, orientation and detection. Alertness means the readiness to receive incoming stimuli. Orientation is the process of directing attentional resources to a particular type of input and ignoring other input. Finally, detection is the selection and registration of such sensory stimuli in memory. They argued that detection does not require awareness. In other words, learners can benefit from certain language input without realizing their effects. This can be seen from studies such as Lynch and Maclean (2000) in which half of the participants engaging

in repeated task performance reported they did not consciously make improvements in their subsequent language output, whereas the transcripts of their performance showed otherwise. Concerning attention to the level of detection, both constructs should provide students with benefits in future oral production as they have registered certain elements in the input to feed into the next repetitions.

However, what rehearsal adds to the learner's preparedness, which repetition does not, is the awareness of future performance(s). This constitutes what Bui (2014) terms a form of 'task-external readiness', where learners are given extra preparation and thus external manipulations for a task. The construct is analogous to N.C. Ellis's (2015) view of explicit learning, which he considered just as important as implicit learning, especially in the context of L2 learning. If learners are aware of next enactments of the same task, they would (possibly) start to rehearse (or practice). They would reflect on what they could have done better from their performance last time and work on improving different aspects of their oral production. This performance thus acts as their chance to experiment with the language and to consciously learn from their mistakes, in the hope of improving them in future repetitions. That makes it clear why Bui (2014) categorized rehearsal under what he calls task-external readiness (alongside strategic planning and online planning), where it is essentially a form of hands-on planning.

Levelt's (1989) 'blueprint of speaking' is another theoretical model of speech that might shed light on the effects on oral task performance under the two contrasting conditions. Anticipating beneficial outcomes, learners will make a deliberate effort in carrying certain 'rehearsed' elements to the next task performance, triggering the monitoring mechanism in Levelt's model, where learners strive to ensure an accurate speech production. However, as studies about careful online planning (e.g. Ahmadian & Tavakoli, 2011; Ellis, 2003; Ellis & Yuan, 2005) have demonstrated, planning within a task severely degrades learners' fluency in oral production (but raises accuracy). This is likely to be caused by conscious use of explicit rules to monitor an otherwise natural speech, leading to pauses or fillers (if monitoring behaviours occur in formulation) and reformulation or false starts (if the monitoring occurs after articulation). However, under the rehearsal condition, this form of online planning might elicit a specific strategy from learners, namely that they, being aware of their limited attentional resources, might pre-emptively focus on a single aspect of oral production and allocate attentional resources to such an aspect during the performance, thus alleviating cognitive load. This has implication for varying effects under the two conditions on CAF: task repetition would result in improvements in multiple aspects of oral production (usually two to three aspects), while task rehearsal would enable learners to strategically enhance their speech performance with a limited scope (possibly only one aspect).

However, it should be noted that the difference between rehearsal and repetition might become negligible should the repetition interval be too long for memory to facilitate monitoring (e.g. a two-week interval in Bui et al., 2019). This is to say, the repeated task needs to follow shortly after the initial task for the benefits of rehearsal to take effect. Assuming learners do not practice on purpose during the time interval between the previous enactment and the next, attention paid to specific areas in the

task performance would fade. In other words, the practice effect or rehearsal in the previous performance should not be sufficient for learners to proceduralize any sort of oral production skills that can be carried over to the next performance. In this case, the effects of rehearsal and repetition on CAF should be similar. It is even possible that the task repetition condition might be more beneficial in causing acquisition in the long term than rehearsal, given repeated training scattered across a long time (e.g. weekly for one semester). This is because rehearsal promotes task-external readiness, which would benefit mainly performance; while task repetition contributes to task-internal readiness, where learners might, through unconscious proceduralisation, acquire language skills in the long term. Longitudinal studies about the role of task repetition and rehearsal in language acquisition are needed in the future.

Effects of Task Repetition on CAF

Though most prior studies on task repetition did not report whether learners were informed of the future performance, some exceptions did exist, as summarized in Table 7.1. Bygate (1996), for example, conducted a small-scale experiment, where participants were asked to watch a cartoon video for about 90 s and then to retell it immediately. The same task was repeated after three days without warning. He found that the learners improved in both accuracy and fluency, with a marked widening of lexical repertoire and a 75% increase in the use of subordinate clauses in their second

Table 7.1 Studies with task repetition

Study	Rehearsal /repetition	Task types	Repetition types	Intervals (repeat)	Dependent variables	Results
Bygate (1996)	Repetition	Narration	Exact	3 days (×2)	CALF	TR > NR in AF
Bygate (2001)	Repetition	Narration, interview	Exact + procedural	10 weeks (× 2)	CALF	TR > NR in CF
Ahmadian and Tavakoli (2011)	Repetition	Narration (immediate)	Exact	1 week	CAF	CAF improved with online planning
Wang (2014)	Repetition	Narration (simultaneous)	Exact	Immediate (×2)	CALF	TR > NR in CAF ns in L
Bui et al. (2019)	Repetition	Narration	Exact	Immediate, 1 day, 3 days, 1 week, and 2 weeks	CALF	TR > NR CF with weak A, ns in L

TR Task repetition, *NR* Non-repetition, *C* Complexity, *A* Accuracy, *F* Fluency, *L* Lexical diversity

performance. Later, Bygate (2001) used a more complex experimental design to test the long-term effects of task repetitions on oral production. Forty-eight participants were assigned into two treatment groups (narratives or interviews) and a control group. Each group engaged in both exact task repetition and procedural repetition after a ten-week interval. The results indicated that participants who repeated the same task performed better in terms of fluency and complexity, but there were no statistically significant improvements for the other two groups.

Ahmadian and Tavakoli (2011) used four different experimental groups to research the effects on oral production with combinatorial conditions in terms of time pressure, online planning and task repetition. Sixty participants were divided into four groups of different task conditions, two of which involved the task repetition component (together with pressured online planning or careful online planning). Participants in these two groups were asked to watch a 15-min silent video and narrate it immediately. They were asked to repeat the exact task in a week without being warned of such repetition. The authors found that the group with task repletion and carefully online planning as conditions simultaneously showed improvements in accuracy, complexity and fluency.

Wang (2014) studied intermediate to advanced students using immediate task repetition as one of the five experimental conditions. The learners were asked to narrate a video they had seen immediately, and they were not told of the second performance until they had finished the first. The results for the repetition group were statistically significant in complexity, accuracy and fluency, with large effect sizes.

Bui et al. (2019) were probably the first study to involve different spacing conditions, from immediate repetition to a two-week interval, between the initial and the reiterated task. The second, repeated task came as a surprise to participants in all these interval conditions as they were intentionally kept unaware of it. With all task conditions taken together, task repetition significantly raised breakdown fluency but not repair fluency, structural complexity but not lexical complexity (as in *D*, or lexical diversity). An interesting finding in this study has been that accuracy was slightly improved in the repeated task when it was measured in the number of errors per 100 words; task repetition did not impact on the 'ratio of error-free clauses' measure.

Effects of Rehearsal on CAF

As mentioned, most of the studies in the literature have not explicitly stated if their participants were notified the possibility of future performances. Even with the ones that do, their experimental conditions are all in the form of task repetition. However, some studies have subtle indications that can lead readers to deduce that the participants in their studies were in some way alert of such possibility (a rehearsal as the condition). As Ellis (2019) comments, 'in some studies …where the same tasks were repeated multiple times, it will become evident to learners that they may have to perform the task again' (p. 18). With the exception of clear indication of a test

Table 7.2 Studies with rehearsal

Study	Rehearsal/repetition	Task types	Repetition types	Intervals (repeat)	Dependent variables	Results
Lynch and Maclean (2000)	Rehearsal (recycle/retrial)	Communicative (poster carousel)	Procedural (poster carousel)	Immediate (×6)	Accuracy, fluency	RH > NR in AF
Patanasorn (2010)	Test practice → rehearsal Rehearsal (p. 67)	Narration	Exact vs. procedural vs. content	3 weeks (×3)	Global accuracy and fluency, accuracy of simple past tense	PR or CR > TR in A or F
De Jong and Perfetti (2011)	Rehearsal	Narration	Exact vs. Procedural (planning available before next TR)	Immediate (after short interviews) (×3)	Fluency	TR > PR in F
Thai and Boers (2016)	Rehearsal	Narrative talk	Exact	Immediate (×3)	Fluency, structural complexity, accuracy	TR > NR in F
Lambert et al. (2017)	Rehearsal with procedural (task set)	Instruction, narration, opinion, communicative	Exact	Immediate (×6)	Fluency (4 measures)	TR > NR in F

RH Rehearsal, *TR* Task repetition, *PR* Procedural repetition, *CR* Content repetition, *NR* Non-rehearsal, *C* Complexity, *A* Accuracy, *F* Fluency

Table 7.3 Unspecified studies with repeated performance

Study	Rehearsal /repetition	Task types	Repetition types	Intervals (repeat)	Dependent variables	Results
Gass et al. (1999)	unknown	Narration (simultaneous)	Exact + Procedural	2–3 days (×3)	AL (morpho-syntax accuracy + lexical sophistication)	R > NR Proficiency Partly accuracy, TTR, lexical density and sophistication
Kim and Tracy-Ventura (2013)	Unknown	Communicative	Exact vs. procedural	Three days (×3)	CAF	PR > TR in syntactic complexity ns in AF, lexical complexity
Fukuta (2016)	Unknown	Narration	Exact vs. procedural	1 week (×2)	CALF	TR > PR in A & L ns in C & F

R Repeating, *NR* Non-repeating, *TTR* Type-token ratio, *C* Complexity, *A* Accuracy, *L* Lexical diversity, *F* Fluency

practice before the task (e.g. Patanasorn, 2010), the experiments in the studies share the following features to be deemed a rehearsal condition: (a) the task is repeated for more than 3 times; (b) the repetitions take place within the same day (usually immediate) with some hints.

Lynch and Maclean (2000) studied how task rehearsal influences the global accuracy and fluency among L2 learners in an English for specific purposes course. The task used in the study is called the 'poster carousel' (Lynch & Maclean, 1994). It first required participants to pair up, and each pair was given a different research article. Then the pairs would make a poster based on the article. The task involved the six pairs standing at different parts of the room, and one of the participants in the pair began by going around the room and asking questions about other posters. The partner that stayed (the host) was responsible for answering the questions from students of other pairs. This process was repeated six times until all six posters were consulted once and the partner returned to his/her own poster, at which point is the host's turn to repeat the same task. This aural-oral task, as Lynch and Maclean caveated, is not the same as what Bygate (1996) termed 'task repetition', as it is not a strict duplication of a task. The authors instead explained that in their case, 'the basic communication goal remains the same, but with variations of content and emphasis depending on the visitor's questions' (Lynch & Maclean, 2000, p. 277). This description matches what Patanasorn (2010) called 'procedural repetition', where the content of each repetition is different, yet the procedure (question and answer) remains the same. Given the design of the task, participants who walked around to ask questions about posters would have anticipated asking similar questions to the following posters station, for example, 'What is this article about?' or 'This is interesting. Can you tell me more about the implications for this study?'. This also goes for 'the host', who had probably prepared a set of answers for certain common questions from visitors. Therefore, it can be inferred that the learners were engaged in some sort of rehearsal while they were performing the task. The two participants that were studied, despite reporting contrasting language self-monitoring, show a general improvement in accuracy, and yet fluency (measured by speech rate) became gradually stable across the carousel task. Transcripts of their task performances reflect that they showed gradual enhancement in accuracy in terms of lexical use, syntactic structure and pronunciation, with the rate of speech improved from the first enactment and slowly remained stable from the second enactment onwards. The authors ascribed this to the unfamiliarity with the task content on the first trial. Moreover, it is observed that, across the six repetitions, the high proficiency learner was able to first enrich the explanation and then make it more concise in later repetitions. In other words, she/he was able to condense the language and achieve the same communicative goal, implying a possible improvement in syntactic complexity in the first few repetitions. In short, they concluded that such task rehearsal (in procedure but not necessarily in content) benefits learners in terms of accuracy and syntactic complexity, but not fluency.

De Jong and Perfetti (2011) employed a 4/3/2 task design to increase time pressure for learners when they repeat a task. The repetition groups (repetition and repetition II) repeated the same task for 4 min, 3 min and 2 min in the same training session on the same day, and they performed the 4/3/2 tasks three times in total for two weeks. This

means that, over a course of two weeks, the learners would have performed similar tasks for nine times. Therefore, at some point during the second 4/3/2 task, participants should have expected a possible similar procedure in the future. This allowed them to expect future performance and started rehearsing in early task performances. The results in the post-test, both immediate (for repetition group) and delayed (for repetition II group), showed that fluency was improved, as opposed to the control group, which did not. The findings also suggested that the benefits of fluency could be maintained over four weeks and transferred to new topics. Though De Jong and Perfetti (2011) only focused on L2 fluency, Thai and Boers' (2016) similar study did explore how the 4/3/2 task rehearsal could benefit L2 CAF. Their results were in line with De Jong and Perfetti in that only greater fluency was achieved but not accuracy or complexity. Thai and Boers discovered that these learners resorted to 'a high amount of verbatim duplication from one delivery of their narratives to the next, which explains why relatively few changes were attested in performance aspects other than fluency' (p. 369).

Another study that can be assumed to have used rehearsal instead of task repetition as the experimental condition is Lambert et al. (2017). The study aimed to find out if task rehearsal could affect L2 fluency in the short term. A total of 32 English learners (Japanese native speakers) were assigned into four groups of eight. The participants were then paired up to perform four tasks in a task set (instruction, narration, opinion and a dialogue task) twice, once as the speaker and once as the listener. After they have finished the task set, they changed partners and repeated the same process, until they have repeated the task set for five times with different partners. Learners were assigned into groups before the experiment, and the pairing was done before the task rehearsal began. Moreover, as the task set needs to be repeated six times (twelve if speaker and listener role are considered separately), the learners should have figured out early in the experiment that they have to repeat the same task in the future, and hence, a task performance would constitute a rehearsal for the next. The results showed that fluency was able to improve gradually starting from the second repetition up until the fifth, at which point the authors deemed it as the optimization of the effects of task rehearsal on fluency. It should be noted that this study, like De Jong and Perfetti (2011), also had only measures related to fluency and therefore the task rehearsal effects on complexity and accuracy remained opaque.

Studies Without Demarcation Between Repetition and Rehearsal

Research information that has been reviewed in Sects. 4 and 5 is largely based on inference from the descriptions given in the methodology section of the mentioned studies. That being said, many other papers do not include sufficient detail for a replication of the experiments concerning the difference between rehearsal and task repetition. This section reviews a sample of these studies.

Gass et al. (1999) used similar task materials used in Skehan and Foster (1997, 1999) to test if the benefits of task repetition can be emulated in a new context. A group of English learners (with L1 Spanish) was separated into two experimental groups (exact task repetition and procedural repetition) and a control group (without task repetition), where the first experimental group watched the same Mr. Bean video for a total of three times at a two to three days interval, with the second experimental group watching different episodes following the same procedure. The results showed improvements for exact task repetition in areas such as overall proficiency, morphosyntax and lexical sophistication. However, those benefits were not transferred to a new task in the post-test.

Kim and Tracy-Ventura (2013) studied 32 female Korean junior high students with high EFL proficiency. The participants were divided into two experimental groups (exact task repetition and procedural repetition) and were required to work on three information exchange tasks, with the exact task repetition group repeating the same task and the other repeating with different content, at a one-day interval for three days. The results generally provided no statistically significant results to support that one type of repetition is better than the other, nor that task repetition treatment improves overall accuracy, complexity or fluency in any way.

Fukuta (2016) studied 28 EFL learners in a junior high school in Japan. The students formed an experimental group and a comparison group. They engaged in a narrative task of six-frame cartoons. The experimental group repeated the exact task with the same cartoon one week later. The comparison group repeated the same task type with different pictures. The results showed that there were remarkable improvements in accuracy and lexical variety for the experimental group, whereas fluency and complexity measures showed no statistically significant differences.

Conclusion and Implications

This chapter reviews previous definitions of task repetition and their limitations in the ambiguity of learner awareness of the repetition. It is argued that task rehearsal should be differentiated from task repetition, as the former involves task-external readiness with potential explicit learning while the latter constitutes task-internal readiness with potential implicit learning. Unfortunately, except very few studies (e.g. Bui et al., 2019), the majority of task repetition research to date has not specified whether their participants were forewarned about future performance; hence the potential differentiating effects being overlooked. Based on a small number of studies in which the judging criteria were clear or could be inferred with some confidence, the following preliminary conclusion could be made:

1. Task repetition with participants uninformed about future performance seems to boost a more balanced range of performance areas in complexity, accuracy and fluency, such as Wang (2014) with strong CAF effects and Bui et al. (2019)

with strong complexity and fluency effects and weak but statistically significant accuracy effect.

2. Task rehearsal with explicit instruction on the next iteration(s) appears to direct learner attentional focus to certain performance areas, such as fluency in De Jong and Perfetti (2011) and Thai and Boers (2016), complexity in Kim and Tracy-Ventura (2013), and accuracy in Lynch and Maclean (2000).
3. Immediate repetition has shown consistent effects on CAF, but different spacing conditions or lengths of the intervals may lead to different performance outcomes. Bui et al. (2019) found that fluency benefited most from the shorter intervals while the one-week interval was the most conducive condition for other aspects of speech. Replication of this study with different task types is needed to further examine the best interval between repeated tasks.
4. From a skill acquisition theory perspective, multiple repetitions are encouraged in the classroom (Lambert et al., 2017). The optimal frequency for task repetition at different proficiency levels warrants further investigations.
5. There appears to be more production or output-based task repetition and task rehearsal research than comprehension or input-based research. TBLT should benefit from more research insights from the latter.
6. It appears that the task repetition and task rehearsal literature typically focuses on clause-based measures for (syntactic) complexity; L2 lexical complexity has only been occasionally employed (Skehan, 2009). It is suggested that future research should consider lexical complexity more systematically to include lexical diversity, lexical sophistication and lexical density (see Bulté & Housen, 2012 for the definitions and Bui, 2019 online for operationalization of these constructs in empirical research).
7. It becomes obvious that future research in this area should clearly report whether the learners are engaged with task repetition, task rehearsal or both at different stages.

Pedagogically, both task repetition and task rehearsal have been shown to benefit L2 language development in terms of their ability to mitigate processing demands, to direct attention to form and to enhance proceduralization and automaticity in an L2. To achieve these ends, rehearsal could first of all be utilized for learners with relatively lower proficiency as preparedness (Ellis, 2019; Skehan, 2014) for an upcoming task to boost their confidence as well as actual L2 performance. The explicit instruction on the next round(s) of the same task may help reduce anxiety about an impromptu task. However, dry rehearsals like this may cause boredom and fatigue, especially for more proficient learners who are capable of improvisation for certain tasks. Then, task repetition can chip in to help more able learners as unwitting participants consolidate or extend their performance. At a more advanced level, procedural repetition with altered content can be adopted to increase the linguistic and cognitive challenges. Secondly, the frequency and the interval of repetition are essential considerations for the implementation of these task conditions. The research discussed earlier seems to suggest a 3–4 times repetition schedule after which the benefits attenuate (Lambert et al., 2017). Then, how to optimally space out repeated

tasks is the next decision to make. Prior studies recommend shorter intervals (such as immediate repetition) for enhanced speech fluency, but they also recommend a one-week interval for language restructuring and improvement (Bui et al., 2019). Longer intervals than a week may still be useful, but they do not seem to be the most conducive for overall L2 performance. To sum up and conclude, teachers should make balanced arrangements of task repetition and task rehearsal in the lesson design with due consideration to learner proficiency and emotional responses along the scale from low to high linguistic and cognitive demands that are imposed on L2 learners. That way we can scaffold language performance and acquisition in a gradual and cyclical manner that characterizes L2 learning.

References

Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research, 15*(1), 35–59.

Baddeley, A., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation, 8*, 47–89.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839.

Bui, G. (2014). Task readiness: Theoretical framework and empirical evidence from topic familiarity, strategic planning, and proficiency levels. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 63–94). Benjamins.

Bui, G., & Huang, Z. (2018). L2 fluency as influenced by content familiarity and planning: Performance, methodology and pedagogy. *Language Teaching Research, 22*(1), 94–114.

Bui, G., & Teng, F. (2018). Exploring learners' behavioral patterns in two task-readiness conditions: A qualitative study. *Chinese Journal of Applied Linguistics, 42*(2), 129–149.

Bui, G., & Teng, F. (2019). Task planning and task-readiness. In J. I. Liontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1–7). Wiley. https://doi.org/10.1002/9781118784235.eelt0972.

Bui, G., & Skehan, P. (2018). Complexity, fluency and accuracy. In J. Liontas (Ed.), *TESOL encyclopedia of English language teaching*. Wiley.

Bui, G., Skehan, P., & Wang, Z. (2018). Task condition effects on advanced level foreign language performance. In P. A. Malovrh & A. Benati (Eds.), *The handbook of advanced proficiency in second language acquisition* (pp. 219–237). Wiley.

Bui, G., Ahmadian, M., & Hunter, A. (2019). Spacing effects on repeated L2 task performance. *System, 81*, 1–13. https://doi.org/10.1016/j.system.2018.12.006

Bui, G. (2019, online). Influence of learners' prior knowledge, L2 proficiency and pre-task planning time on L2 lexical complexity. *IRAL—International Review of Applied Linguistics in Language Teaching*. https://doi.org/10.1515/iral-2018-0244.

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Investigating complexity, accuracy and fluency in SLA* (pp. 21–46). John Benjamins.

Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136–146). Macmillan.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23–48). Longman.

De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning, 61*(2), 533–568.
DeKeyser, R. M. (2007). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 97–112). Erlbaum.
Dekeyser, R., & Criado, R. (2012). Automatization, skill acquisition, and practice in second language acquisition. *The encyclopedia of applied linguistics.*
Ellis, N. C. (2015). Implicit and explicit language learning: Their dynamic interface and complexity. In *Implicit and explicit learning of languages* (pp. 1–24). John Benjamins.
Ellis, R. (2005). *Planning and task performance in a second language.* John Benjamins.
Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics, 30*(4), 474–509.
Ellis, R. (2019). Task preparedness. In Z. Wen & M. Ahmadian (Eds.), *Researching L2 task performance and pedagogy: In honour of Peter Skehan* (pp. 15–38).
Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In *Planning and task performance in a second language* (pp. 167–192). John Benjamins.
Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research, 20*(3), 321–340.
Gass, S., Mackey, A., Alvarez-Torres, M. J., & Fernández-García, M. (1999). The effects of task repetition on linguistic output. *Language Learning, 49*, 549–581.
Housen, A., & Kuiken, F. (2009). Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics, 30*(4), 461–473.
Kim, Y., & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? *System, 41*(3), 829–840.
Lambert, C., Kormos, J., & Minn, D. (2017). Task repetition and second language speech processing. *Studies in Second Language Acquisition, 39*(1), 167–196.
Levelt, W. J. M. (1989). *ACL-MIT Press series in natural-language processing: From intention to articulation.* The MIT Press.
Lynch, T., & Maclean, J. (1994). Poster carousel. In K. Bailey & L. Savage (Eds.), *New ways of teaching speaking* (pp. 108–109). TESOL.
Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research, 4*(3), 221–250.
Michel, M. (2017). Complexity, accuracy, and fluency in L2 production. *The Routledge Handbook of Instructed Second Language Acquisition, 50*, 68.
Patanasorn, C. (2010). *Effects of procedural content and task repetition on accuracy and fluency in an EFL context*. Northern Arizona University.
Robinson, P. (Ed.). (2001). *Cognition and second language instruction.* Cambridge University Press.
Skehan, P. (1998). *A cognitive approach to language learning.* Oxford University Press.
Skehan, P. (Ed.). (2014). *Processing perspectives on task performance.* John Benjamins.
Skehan P. (2009). Lexical performance by native and non-native speakers on language-learning tasks. In B. Richards, M. H. Daller, D. D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition.* Palgrave Macmillan.
Skehan, P., Bei, X., Li, Q., & Wang, Z. (2012). The task is not enough: Processing approaches to task-based performance. *Language Teaching Research, 16*(2), 170–187.
Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research, 1*(3), 185–211.
Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49*(1), 93–120.
Thai, C., & Boers, F. (2016). Repeating a monologue under increasing time pressure: Effects on fluency, complexity, and accuracy. *TESOL Quarterly, 50*(2), 369–393.
Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*(2), 183–203.

Wang, Z. (2014). On-line time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27–62). John Benjamins.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity, and accuracy in L2 monologic oral presentation. *Applied Linguistics, 24*, 1–27.

Chapter 8
Task Complexity and Language Proficiency: Its Effect on L2 Writing Production

Veena Nair and Shruti Sircar

Abstract The paper reports a study that examined the impact of the interaction of task complexity manipulations and language proficiency on second-language writing performance. Increased cognitive complexity in tasks has been shown to benefit writing in terms of syntactic and lexical complexity; the quantum of benefit defined by language proficiency notwithstanding. The study manipulated task complexity and studied the impact on written production of learners at two different levels of proficiency in English and wished to contribute to the debate that Robinson's Cognition Hypothesis (2001a, b) and Skehan & Foster's Limited Attentional Capacity Model (1997) sparked off about the facilitative or adverse impact of task complexity on production. It also wished to see how language proficiency interacts with task complexity to affect L2 written production. On the basis of an English proficiency test that included reading, writing, speaking, vocabulary and grammar, 30 learners of age range 23–38 years were categorized as lower ($N = 15$) and higher proficiency ($N = 15$) groups. The lower proficiency group were students of Bachelor's degree in English, Engineering, Science or Computers and had an average of 10 years of English instruction. The high proficiency participants were pursuing their doctoral studies in Humanities with at least 12–15 years of English medium instruction. Five tasks were manipulated for cognitive complexity: [−complex] tasks required learners to describe a product that their company is planning to launch in the market (a descriptive task), and the [+complex] version required learners to take a standpoint on which product a customer should go with: an attractive, not so high quality one or an unattractive one but high quality. The standpoint needed to support with arguments (an argumentative task). The written production was assessed on five linguistic measures—syntactic complexity, syntactic variety, lexical density, lexical variety and accuracy. The main findings of the study were: (i) proficiency affected written performance in all dimensions except frequency of reference markers, high proficiency learners with better scores than lower proficiency learners, (ii) complexity

V. Nair (✉)
Mar Baselios College of Engineering and Technology, Thiruvananthapuram, India

S. Sircar
Department of Linguistics and Contemporary English, EFL University, Hyderabad, India

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_8

increase affected the two groups differently—it increased syntactic complexity and lexical variety in higher group, and not in lower group; it decreased lexical density and variety in lower group, and (iii) accuracy remained unaffected, calling into question Skehan's main claim of the Limited Capacity Model. A task difficulty questionnaire administered post the tasks showed an increase in thinking and perception of stress and decreased interest in both groups, suggesting that the task complexity manipulations are also psychologically real. The findings of the study validate Robinson's *Cognition Hypothesis* that a focused attention on complexity is not at the expense of accuracy especially for high proficiency learners, though the effect of task complexity on written performance is modulated by language proficiency.

Keywords Task complexity · Linguistic complexity · Proficiency · Writing · Cognition · Accuracy

Introduction

The study[1] reported in the paper examines the effects of task complexity and language proficiency on L2 written production. Task complexity was manipulated on dimensions of resource directedness (Robinson, 2001a, b), and proficiency was determined on the basis of a language test. The effects of manipulation were examined on specific dimensions of the written performance, i.e. complexity (syntactic and lexical), variety (syntactic and lexical) and grammatical accuracy.

Two models of task complexity were put to test: Skehan and Foster's Limited Attentional Capacity Model (Skehan, 1998, 2001, 2003; Skehan & Foster, 1999, 2001) and Robinson's Triadic Componential Framework, also known as the Cognition Hypothesis (Robinson, 2001c, d, 2005, 2007a, b). The essential point of debate in the two models is which task features affect the allocation of attentional resources during task completion and which aspects of production get affected by the manipulation of task features.

An aspect that we examined in the study, to tease out the different claims of the two models was *language proficiency*: Do cognitive demand manipulations in the task affect second-language learners at different proficiency levels in a similar fashion? Or do the learners with lower proficiency get adversely affected by task demand manipulations?

The paper is organized as follows: Section "Introduction" discusses the information-processing theory which underlies the project of task-based language teaching, Section "The Allocation of Attentional Resources and Task Demands" looks at the dimensions and variables of task complexity and compares them within Skehan and Foster's Limited Attentional Capacity Model and Robinson's Triadic Componential Framework. Section "The Limited Attentional Capacity Model and the Triadic Componential Framework" looks at proficiency as a variable in

[1] This study is a part of a Ph.D. project submitted by Veena Nair at EFLU University, Hyderabad.

the task complexity–task performance interaction. Section "Proficiency as a Variable" presents the aims of the study and explains the specific measures which have been proposed to assess written production of L2 learners. Sections "The Study" and "Methodology" present the design, the methodology and the results of the study, followed by a discussion of the implications of the study for further research in Section "Results and Discussion".

The Allocation of Attentional Resources and Task Demands

An individual cannot easily perform two simultaneous tasks; this is why talking on the cell phone and driving is banned across the world. Attentional resources cannot be allocated to both tasks simultaneously, particularly when two different outputs in terms of behaviour are expected from the tasks. Trying to do both is not possible, unless automatized. Usually, while attending to one of the tasks, the other gets affected by an attentional bottleneck. According to various 'Bottleneck' theories of attention (e.g. Broadbent, 1958, 1971; Deutsch & Deutsch, 1963, 1967; Treisman, 1969), information processing capacity is limited, and processing of unattended stimuli stops, either at the early or later stage of the information-processing sequence. An alternative to structural theories of attention is the 'capacity theories' (e.g. Kahneman, 1973; Posner & Snyder, 1975; Wickens & Kessel, 1980). The human processing system is believed to have limited resources or capacity. 'Capacity Theories' allow more flexibility in allocation of attention. As per the task demands, attention can be allocated to different activities.

Kahneman (1973) described attention as a reservoir of mental energy from which resources are drawn to meet situational attentional demands for task processing. Mental effort or attention is controlled by task demands, i.e. mental effort increases proportionately with task difficulty/complexity, and learning results in reduction of mental effort required for performing a task and producing a certain outcome. With the increase in task demands, the excess resources are initially allocated in order to achieve the required level of task performance. However, with the continuous increase in the task complexity, demands continue to increase, and soon a point is reached where the amount of required resources becomes greater than the limit of available resources. Additional increases in task demands are not supported by increases in allocated resources. Skehan and Foster's Limited Attentional Capacity Model follows Kahneman's capacity model.

Nevertheless, at times, two tasks can be simultaneously performed with no difference in quality of output, provided suitable conditions are provided to perform each task separately. If the tasks are extremely challenging and require maximum use of the available resources, performance on both tasks may decline (frequently referred to as dual-task deficit or divided attention deficit). Capacity models of attention were developed mainly to understand and explain divided attention performance. Primarily there are two categories of capacity models. First, which support a single,

undifferentiated attentional resource and the second, which propose multiple, independent attentional resources. Kahneman's model (1973) involved an undifferentiated resource. This model could not explain why some task combinations were easier than others. It could also not explain why some complex tasks could be performed together without any effect on the output at all. Further theoretical developments for capacity models led to models such as Wickens' multiple resource model (1989).

According to Wickens' model, different dimensions can draw on different resource pools, and competition for attention may not be the norm. There would be competition only if two tasks feed on the same resources. Later models (Navon, 1989; Neumann, 1987) have used the idea of Wickens' different resource pools but abandoned the idea of capacity limitations. Processing difficulties and limitations are now treated as consequences of interference between stimuli drawing from the same pool but requiring different responses/outputs, or because of a kind of 'crosstalk' among resources, and not due to any capacity constraints (Sanders, 1998). Wickens argued that when completing different tasks, people draw their attention from different resource pools depending on processing mechanisms (i.e. encoding or responding), codes (i.e. spatial or verbal), modalities (i.e. visual or auditory), or responses (i.e. manual or vocal) that each task requires. It is claimed that a competition for attention occurs not between but within the resource pools. Robinson's Triadic Componential Framework of task complexity follows Wickens' model.

The Limited Attentional Capacity Model and the Triadic Componential Framework

Although it is generally agreed upon that tasks make different demands on our attentional resources, Limited Attentional Capacity Model developed by Skehan and Foster (Skehan, 1998, 2001, 2003; Skehan & Foster, 1999, 2001) and Robinson's Triadic Componential Framework (Robinson, 2001c, d, 2003a, b, 2005, 2007a, b) make contrasting predictions about linguistic performance in relation to the attentional demands of tasks.

Limited Attentional Capacity Model

Skehan's Limited Capacity Model (1996) is founded on theories of working memory (Carter, 1998; Gathercole & Baddeley, 1993) and Kahneman's concept of attention or mental effort as a finite but multidimensional resource. Skehan extends what was initially used to explain divided attention and dual tasks to attentional resource allocation to different aspects of task performance, i.e. fluency, which is content-focused, and accuracy and complexity, which are form-focused. One is believed to be achieved always at the expense of the other, i.e. each vie for attention. When the

task demand increases, learners first allocate attentional resources to the content of the task, and what remains is assigned for meeting the complexity and accuracy demands of the task. When content demands are very high, the limits of attentional resources are reached, and in the absence of (or a very limited) attentional resources, complexity lapses to the automatically processed simpler language and lowers the accuracy of non-automatic forms.

To further elaborate how the automatic and controlled processing is put into action, Skehan (1996) bases his work on a model of language learning where L2 knowledge is stored in a dual system, a rule-based system and an exemplar-based system. The former, consisting of abstract representations of language patterns, requires more processing and is ideally suited for more controlled, less fluent language performance. The latter, being lexical in nature, consists of words and fixed units. This system contains the linguistic knowledge which can easily and quickly be accessed or automatically processed, and hence, it is adequate for fluent language performance. Accuracy and complexity therefore require learners to draw on their rule-based system and thus require syntactic rather than semantic processing, while fluency draws on their exemplar-based system (Skehan, 1998).

Studies supporting Skehan and Foster's model: According to studies in Skehan's paradigm (Crookes, 1989; Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997; Wigglesworth, 1997; Yuan & Ellis, 2003), for cognitively simple tasks (with planning time), there is a marked increase in fluency, relatively strong effect on complexity, and very little effect on accuracy. Studies have investigated whether familiarity with the task type interacts with task complexity (Foster & Skehan, 1996) and the effect of pre-task and online planning time (Crookes, 1989; Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997; Wigglesworth, 1997; Yuan & Ellis, 2003). Cognitively simple tasks (with planning time) showed a marked increase in fluency, relatively strong effect on complexity and very little effect on accuracy, thereby supporting Skehan's limited capacity model.

However, Skehan's Limited Capacity model was unable to explain the phenomenon of multitasking. This led to the propagation of a dual or multiple processing model, which is the basis of Robinson's Triadic Componential Framework.

The Triadic Componential Framework

Integrating information-processing theories (Schmidt, 2001), interactionist explanations of L2 task effects (Long, 1996) and psychological models such as Wickens' model of dual-task performance (Wickens, 1989, 1992), Robinson claims that learners can simultaneously access multiple and non-competitional resource pools. As a result, manipulating task complexity by increasing the cognitive demands of a task can lead to simultaneous improvements of complexity and accuracy (1995a, b, 2001c, d, 2003a, b, 2007c) since the input is processed more deeply and elaborately (Gilabert, 2007).

The Cognition Hypothesis (2001) proposes that learners can concurrently fulfil multiple task demands as long as they draw from different pools of attentional resources. This theory includes taxonomy of factors that may influence attentional allocation during task-based L2 performance, and the Triadic Componential Framework distinguishes cognitive factors of *task complexity* from interactive factors of *task condition* and from learner factors of *task difficulty*.

Task complexity: The triadic framework proposes two distinct dimensions of *task complexity*: 'resource-directing' dimensions and 'resource dispersion' dimensions.

Resource-directing dimensions are those in which the demands on language use made by increases in task complexity can be met by manipulating the manner (directing their resources) in which the information is presented. When a task becomes cognitively more challenging, the demands may be met by directing more attentional resources towards the linguistic form. For instance, a task which requires us to justify our beliefs, argue for a stance, predict actions using causal logic, give reasons, infer from given set of facts and substantiate interpretations, is cognitively complex than a simple descriptive or narrative task. To fulfil the conceptual and performative demands of the task, more complex features of language need to be noticed and employed. For instance, an argumentative writing task would require a learner to use lexically marked vocabulary units (e.g. verbs '*claim*', '*propose*', '*argue*') and syntactically complex adverbial clauses (e.g. '*if... then*' clauses). In contrast, a simple narrative task would need simple or coordinate clauses. Since formal aspects of language receive attention, consciously or unconsciously, complexity and accuracy would also be pushed up, though fluency might be affected negatively.

Task complexity with respect to resource-directing dimension can be manipulated on the following factors: ± here and now, ± few elements and ± no reasoning demands. In our study, we looked at ± no reasoning demands.

Resource dispersion dimensions are those where increasing complexity replicates the processing conditions under which real-time language is often used. Practice along these dimensions could be argued to facilitate real-time access to an already established and developing repertoire of language. Robinson (2005: 167) argues that an increase on resource-dispersing variables (e.g. ± planning time, ± prior knowledge) diverts attentional allocation over various linguistic and non-linguistic task aspects. As a result, the linguistic output of L2 performers suffers (resulting in lowering of linguistic complexity, accuracy and fluency) due to the extra cognitive load of a complex task that focuses the attention on other task features than language.

Though many studies (Ellis & Yuan, 2004; Mehnert, 1998; Robinson, 2005; Skehan, 1996; Skehan & Foster, 2005) have been conducted in the area of planning, none of them looked at how planning in the presence or absence of pre-tasks can affect language production. In a previous study (Nair, 2008), it was found that complex tasks get benefitted (with increased structural complexity and lexical variety) when planning time is provided. This paper will not look at these aspects.

Task conditions are concerned with the participation dimension, for instance, the information flow in classroom interactions (one-way or two-easy tasks), open/closed,

convergent /divergent, etc., and participant variables such as gender of participants in a group/pair, familiarity/unfamiliarity with each other, etc.

Task difficulty refers to learner perceptions of the level of difficulty of task resulting from the abilities and skills they bring to the task, i.e. intelligence, working memory, aptitude and many affective factors like motivation, anxiety, confidence, etc.

According to Robinson, both task difficulty and task condition cannot form a basis for sequencing as they are largely learner dependent and can be determined only during the course rather than before the course (Robinson, 2001a, b).

Studies supporting Robinson's Cognition Hypothesis: Here, we report a few findings of studies on L2 writing which support the Cognition Hypothesis. In Kuiken and Vedder's studies (2007, 2008), the findings, in general, authenticate the improvement of accuracy of L2 development. The general findings of the study done by Ishikawa (2006) on manipulating task complexity on the ± here and now dimension found that increased task complexity resulted in increase in the complexity, accuracy and fluency of written language production. Yuan and Ellis (2003) studied the effects of pre-task planning, online planning and no-planning on complexity, accuracy and fluency of Chinese narrative writings. They found that pre-task planning led to increased fluency and syntactic variety, and online planning led to increased accuracy. Similarly, pre-task planning produced greater fluency and complexity of learners' written performance in Kang (2005) (Table 8.1).

Only a few studies (Gilabert, 2007; Levkina, 2008; Michel et al., 2007) have examined the interaction of two cognitive variables, and no written studies have specifically looked at the interaction of resource-directing (narrative vs. argumentative) and resource dispersion (± pre-task) manipulations. Simultaneous manipulation was found to lead to mixed results in the studies. As only a few studies have looked at the interaction between different dimensions, this research is therefore an attempt at analysing a new dimension within the field of task complexity studies. In this study, we looked at the interaction between task complexity [± Resource directing (± reasoning demands) and task difficulty (proficiency)].

Proficiency as a Variable

The Robinson model claims that task complexity does not adversely affect proficiency; however, if language proficiency is not very high, it would require added allocation of attention. In that case, would accuracy still remain unaffected?

One of the issues that these task complexity studies have not looked at in a very concerted way is the proficiency of the learners. This factor has been theorized in the cognition model under task conditions as learner variable, i.e. what the learner brings to the task rather than a feature of the task or the conditions within which the task has to be performed. Language proficiency has been addressed in a few recent studies. Some studies that have looked at proficiency variable are Kawauchi (2005), Ishikawa (2006), Kuiken and Vedder (2007), Kuiken and Vedder (2008),

Table 8.1 Summary of research on task complexity (resource-directing) and writing production

Studies	Context of study	Task complexity	Result
Robinson et al. (1995)	Learners of Mandarin Chinese	Narrative ± planning	Fluency: no change Accuracy: no change Complexity: increase
Kuiken et al. (2005)	Dutch learners of Italian	Number of elements (letter writing)	Complexity: ns Accuracy: decrease
Ishikawa (2006)	Japanese high school learners of English	Here and now/there and then (narrative)	Accuracy: increase Syntactic complexity: increase Fluency: increase
Kuiken and Vedder (2007)	Dutch university students of Italian and of French origin	Number of elements (letter writing)	Accuracy: increase Complexity: increase
Kuiken and Vedder (2008)	Dutch university students of Italian and of French origin	Number of elements (letter writing)	Accuracy: increase Structural complexity and lexical variety: no change
Abdollahzadeh and Kashani (2011)	High proficiency and low proficiency Iranian EFL participants	Here and now/there and then (narrative)	Complexity and accuracy: increase for High proficiency learners Fluency: no change

Studies on oral tasks (Deng, 2005; Gilabert, 2005, 2007; Ishikawa, 2006; Iwashita et al., 2001; Niwa, 2000; Rahimpour, 1997; Robinson, 1995a, b, 2001c, d, 2005, 2007a, b; Robinson et al. 1995; Shiau & Adams, 2011) showed trade-off between accuracy and fluency on increase in complexity; however, complexity in writing was found to increase with task complexity

Abdollahzadeh and Kashani (2011), Malicka and Levkina (2012). In most of these studies, task complexity seems to have benefitted the high proficiency learners more.

In Table 8.2, we see beneficial effects of task complexity especially for high proficiency learners (though a trade-off is seen between accuracy and fluency).

The Study

The findings reported above provide support for Cognition Hypothesis, which claims that increasing the complexity of a task along the resource-directing dimension can direct learners' attention to essential features in the task input. The learners will be forced to 'notice' these features in order to meet the task demands, and this 'noticing' (Schmidt, 1990) would enhance their task performance in terms of complexity of form, as well as push interlanguage development. Since the task input creates a situation of need for complex form, the learners will be compelled to use forms of language that might not have been used when they had to do a less complex task.

Table 8.2 Proficiency studies in task complexity

Research	Subjects	Task	Results
Wigglesworth (1997)	400 ESL learners at different levels of proficiency in Australia	Planning time (speaking)	Complexity and accuracy: increase in high proficiency learners Fluency: increase in low proficiency learners
Kuiken et al. (2005)	62 Dutch learners of Italian	Letter writing. no. of elements	Complexity: no effect Accuracy: increase when task complexity and proficiency both high Low proficiency group: no effect
Kawauchi (2005)	12 high intermediate and 11 advanced Japanese EFL learners	Three picture-based narrative tasks	Fluency: (−planning): was low < high < advanced but in (+planning): low < high = advanced Complexity: no effect Accuracy: increase only in low proficiency
Ishikawa (2006)	52 Japanese high school learners of English	Narrative, here and now/there and then	Accuracy: increase Structural complexity: increase Lexical complexity: decreases with increase in proficiency Fluency: increase
Kuiken and Vedder (2007)	84 Dutch university students of Italian and 75 students of French	Letter writing. no. of elements	Accuracy and complexity: increase
Kuiken and Vedder (2008)	91 Dutch university students of Italian and 76 students of French	Letter writing. no. of elements	Accuracy: increases Structural complexity and lexical variety: no effect
Abdollahzadeh and Kashani (2011)	32 high proficiency and 36 low proficiency participants	Writing. Here–now and there–then	Complexity and accuracy: increase in complex task for high proficiency learners Fluency: no effect
Malicka and Levkina (2012)	37 (20 advanced and 17 pre-intermediate undergraduates)	Speaking, ± few elements, ± spatial reasoning	High proficiency: accuracy and complexity increases. Fluency decreases Low proficiency: fluency increases. Accuracy and Complexity: no effect

Research Questions

The study sought to find answers to two broad research questions:

1. What is the effect of task complexity on learners' written performance?
2. Does language proficiency interact with task complexity to affect L2 written production?

Apart from this, we also looked at whether task complexity and language proficiency affected learner perception of task difficulty.

Task Conditions

Task complexity and language proficiency were the independent variables, and written performance was the dependent variable.

Task complexity was manipulated for reasoning demands. Learners were administered a set of 5 [+complex] tasks which required learners to justify their choice of statement, i.e. it required argumentation, and a set of 5 [−complex] tasks which were description of objects or products.

Two groups of learners, who differed in their proficiency in English, participated in the study. Their proficiency levels were established on the basis of an English proficiency test which assessed learners' knowledge of grammar, vocabulary, reading and writing.

Measures Used to Examine L2 Written Production

The written performance of the learners was assessed for five linguistic features: syntactic complexity, syntactic variety, lexical density, lexical variety and accuracy.

Syntactic (i.e. grammatical) complexity refers to the range and the degree of *sophistication* of the forms that appear in production (Ortega, 2015: 82), and this concept has long been regarded as an important construct in L2 proficiency. Syntactic complexity is multidimensional in nature, and a variety of measures tap into different dimensions of the construct. The syntactic complexity measure varies across studies depending on what these are used to measure. Brown (1973) looked at length of utterance in terms of morphemes to understand the emergence of syntax in child speech. Bardovi-Harlig (1992) used it as a measure of coordination, Robinson (2007a, b) as a frequency of various verb forms and Elder and Iwashita (2005) as a measure of subordination as mean of number of clauses per T-unit. The most commonly used measure, however, is the mean length of T-unit (Gilabert, 2005; Ishikawa, 2006; Rahimpour & Hosseini, 2010; Robinson, 1995a, b), which we use in our study. Our measure for syntactic complexity was mean T-unit length. MTUL is the total number of words divided by the total number of T-units in a text (Hunt, 1965).

Syntactic variety measures the command of a range of syntactic structure and is measured as a type-token ratio. Guiraud's index (Guiraud, 1954) was used to control for text length. For a long text, this procedure will result in a higher lexical richness than what would have been obtained with a simple TTR.

$$\text{TTR} = \frac{\text{Type of verb patterns}}{\sqrt{2} \times \text{Verb pattern token}}$$

Lexical density measures information packaging which is indicated as a ratio of content words (noun, verb, adjective and adverb) and function words (prepositions, interjections, pronouns, conjunctions and count words) in a text. The measure was similar to one used by Robinson (1995a, b) where the percentage of open-class words to closed-class words per utterance was counted.

Lexical variety measures the lexical richness of the text. For a text to be highly lexically varied, the speaker or writer has to use many different words, with little repetition of the words already used, and thus a richer vocabulary. It is expressed as a ratio of type of words and total number of words.

Accuracy is the number of error-free T-units per T-units (Arent, 2003; Rahimpour, 2008).

Methodology

Participants

Two groups of learners, differing in their proficiency in English, participated in the study. The proficiency of the groups was established on the basis of an English proficiency test which assessed learners' knowledge of grammar, vocabulary, reading and writing, and the difference was statistically significant ($p < 0.0001$). One group which will be referred to as 'Higher Proficiency' (HP) group comprised 15 (ten females and five males) learners pursuing their doctoral programmes in English in Hyderabad. The score of this group was 80.42 ($M = 67.4$, SD = 4.92). The second group, henceforth referred to as the 'Lower proficiency' (LP) group, comprised 15 learners (eight females and seven males) pursuing their Bachelor's degrees in varied fields like English, Engineering, Science or Computers from different colleges in Hyderabad and had a score of 40.3, ($M = 34.4$, SD = 4.27).

All learners had had a minimum of 12 years of instruction in English and had studied in schools where the medium of instruction was English. The learners came from varied backgrounds with Tamil, Telugu, Kannada, Hindi, Marathi and Kashmiri as their L1; however, this study does not study the influence of the L1 on L2 writing, and therefore, they are treated as one group.

Task Stimuli

The learners were expected to read and understand the prompt and write a paragraph or letter as instructed by the task.

[−complex] task involved descriptive writing, where the language needed was simple sentence structures, an extensive use of adjectives, that are verifiable and scalar (smooth, clear, narrow, etc.), and a limited use of phrasal or clausal embeddings. On the contrary, a argumentative writing [+complex] required learners to justify an opinion or substantiate a claim, which involved complex sentences, phrasal embeddings, use of evaluative adjectives, and nuanced use of verbs which indicate the intention of the author, thus making the text syntactically more complex and lexically dense. While descriptive writing requires a dependence on the perceptual aspects of the object to be described and finding the right attributes (lexically and semantically) to describe the product, argumentative writing requires a two-step process, where the content needs to be thought out along with the way the content needs to be laid out in order to fulfil the discourse requirement, i.e. persuasion.

We present below a sample task 'social networking sites' and its [+complex] and [−complex] versions.

> [−complex].
> One of your friends does not know much about social networking sites. Send a friend request inviting him/her to join the site. Also send a mail describing briefly what social networking is all about (definition, advantages, disadvantages of social networking, etc.).

> [+complex].
> Pakistan recently banned three major social networking sites: Orkut, Facebook and Twitter on religious grounds. Other countries like China and Indonesia have also banned certain social networking sites. According to them, these sites propagate immorality and violence. "Moral policing should be a characteristic of Social network-ing sites". Do you agree? Give reasons.

Procedure

Five sets of tasks were used, each with a [+complex] and a [−complex] version. No learner got both versions of the task, and therefore, there was no practice effect. For each writing task, a time frame of 80 minutes was provided. The study was conducted over 10 weeks, and learners wrote two essays per week. Initially, 50 subjects were

Table 8.3 Descriptive statistics for learners' written production

	Measures		Descriptive task −complex	Argumentative task +complex
1	Syntactic complexity[a]	HP	13.75 (1.98)	16.54 (1.54)
		LP	10.27 (1.42)	10.05 (1.78)
2	Syntactic variety[b]	HP	1.19 (0.15)	1.34 (0.09)
		LP	1.11 (0.15)	1.20 (0.21)
3	Lexical variety[c]	HP	0.52 (0.07)	0.49 (0.05)
		LP	0.41 (0.05)	0.46 (0.02)
4	Lexical density[d]	HP	2.44 (0.35)	2.34 (0.31)
		LP	1.88 (0.34)	2.19 (0.23)
5	Accuracy[e]	HP	0.93 (0.06)	0.87 (0.13)
		LP	0.55 (0.16)	0.67 (0.17)

[a]Mean of words per T-unit; [b]Type-token ratio (Guiraud's index) for verb patterns; [c]Type-token ratio of content words to functional words; [d]Type-token ratio of different types of words to total number of words; and [e]Number of error-free T-units

taken for the study. Eight subjects did not submit all sets of essays and were therefore eliminated from the actual study. Twelve learners were outliers and were therefore not considered for this study.

Results and Discussion

All five task topics used were similar in familiarity, and therefore, task topics would not affect writing output in significant ways. No significant topic effect was found for both the HP group ($F(4, 70) = 8.51$, $p < 0.07$) and the LP group ($F(4, 70) = 9.34$, $p < 0.09$).[2]

Table 8.3 provides the descriptive statistics for complexity, accuracy and discourse measures of L2 learners' written narratives in descriptive (−complex) task versus argumentative (+complex task). The mean scores and SD in parentheses are reflected.

Higher proficiency group: As seen in Table 8.3, in high proficiency learners, syntactic complexity was significantly higher in [+complex] tasks (Mean = 16.54, SD 1.65) than in [−complex] tasks (Mean = 13.75, SD = 1.98), [$F(1, 140) = 16.4$, $p < 0.05$)]. In the case of syntactic variety also, there was a significant increase [$F(1, 140) = 9.43$, $p < 0.05$] in [+complex] tasks (Mean = 1.34, SD = 0.098) as compared to [−complex] tasks (Mean = 1.19, SD 0.154). Lexical variety and lexical density decreased when tasks become complex, but the drop was not significant.

Lower proficiency group: When task complexity was increased, the syntactic complexity of the task performance of low proficiency learners did not record a

[2] One-way ANOVA was done individually for LP and HP groups.

significant increase: compare mean of 10.27 (SD = 1.42) in [+complex] condition with mean of 10.05 (SD = 1.78) in [−complex] condition. Interestingly, in lexical variety and density, in [+complex] condition, there was a significant decrease. Lexical variety decreased in [+complex] tasks from a mean of 0.46–0.41, though statistically not significant. Lexical density significantly decreased from a mean of 2.18 to 1.88 in the [+complex] condition [$F(1, 140) = 8.99, p < 0.001$].

In comparing high and low proficiency learners, we found that syntactic complexity and variety increased only in the writings of the high proficiency group. In case of the low proficiency group, we found a significant decrease in lexical density and lexical variety in [+complex] tasks when compared to the [−complex] tasks. *Accuracy showed no significant decrease or increase in the groups.*

Interaction effects: By means of ANOVA, significant effects of the proficiency level (HP and LP), task complexity ([+complex] and [−complex]) and interaction of proficiency and task complexity were examined, as provided in Table 8.4.

The results indicate a significant effect of **proficiency level** with respect to syntactic complexity [$F(1, 140) = 545, p < 0.05$], syntactic variety [$F(1, 140) = 11.8, p < 0.05$], lexical variety ([$F(1, 140) = 50.8, p < 0.05$], lexical density [$F(1, 140) = 3318, p < 0.001$] and accuracy [$F(1, 140 = 731, p < 0.05$]. The HP group wrote essays with more complex structures and with more varied structural patterns when compared with the writing of the LP group. The HP group outperformed the LP group as the former made fewer errors, used more variety in words and used more function words when compared to content words.

With regard to **task complexity**, a significant effect was found for syntactic complexity [$F(1, 140) = 53.0, p < 0.000$], accuracy [$F(1, 140) = 14.2, p < 0.000$], with students performing better in the [−complex] than in the [+complex] condition. [+complex] and [−complex] versions did not significantly affect syntactic variety, lexical variety and lexical density.

Significant interaction effect between **proficiency level** and **task complexity** was seen for syntactic complexity [$F(4, 140) = 20.1, p < 0.000$], syntactic variety [$F(4, 140) = 7.31, p < 0.01$] and lexical variety [$F(4, 140) = 11.1, p < 0.000$]. [+complex] and [−complex] versions significantly varied across proficiency level on syntactic complexity, syntactic variety and lexical variety.

The results indicate a significant effect of proficiency level with respect to syntactic complexity, syntactic variety, lexical variety, lexical density and accuracy.

During the review of planning time and here and now studies, trade-off effects had been hypothesized between different dimensions of production. Skehan and Foster (1997) had shown that gains in complexity are always at the expenses of accuracy. In oral productions, fluency was always at the expense of accuracy (Yuan & Ellis, 2003), and the reason given was a competition for attention. However, evidence from the present study and earlier studies (Table 8.5) shows that *a focused attention on complexity is not at the expense of accuracy especially for high proficiency learners.*

In earlier studies on proficiency and task complexity (cf Table 8.6), the interaction between proficiency and task complexity showed no trade-off effect between

Table 8.4 Interaction effects between proficiency and task complexity

Measures	Proficiency level			Task complexity			Level * task complexity		
	F	df	*p*	*F*	df	*p*	*F*	df	*p*
Syntactic complexity	545	1, 140	0.000	53.0	1, 140	0.00	20.1	4, 140	0.000
Syntactic variety	11.8	1, 140	0.000	0.810	1, 140	ns	7.31	4, 140	0.01
Lexical variety	50.8	1, 140	0.000	3.13	1, 140	ns	11.1	4, 140	0.000
Lexical density	3318	1, 140	0.000	1.51	1, 140	ns	2.28	4, 140	ns
Accuracy	731	1, 140	0.000	14.2	1, 140	0.00	0.813	4, 140	ns

df is degree of freedom, ns—not significant at $p < 0.05$; *$p < 0.05$; **$p < 0.01$

Table 8.5 Performance on [+complex] tasks with respect to [−complex] tasks in study

Measure	High proficiency	Low proficiency
Syntactic complexity	↑	*ns*
Syntactic variety	↑	*ns*
Lexical variety	*ns*	↓
Lexical density	*ns*	↓
Accuracy	*ns*	↓

↑ significant increase; ↓ significant decrease; *ns* not significant

Table 8.6 Performance on [+complex] tasks with respect to [−complex] tasks for high proficiency learners with respect to low proficiency learners in earlier studies

		Kawauchi (2005)	Kuiken et al. (2005)	Ishikawa (2006)	Kuiken and Vedder (2007)	Kuiken and Vedder (2008)	Abdollahzadeh and Kashani (2011)	Malicka et al. (2012)
Complexity	Syntactic	↑	*ns*	↑	↑	*ns*	↑	↑
	Lexical	↑	*ns*	↓	–	*ns*	–	↑
Accuracy		↑	↑	↑	↑	↑	↑	↑
Fluency		↓	–	↑	–	–	*ns*	↓

↑ significant increase; ↓ significant decrease; *ns* not significant

complexity and accuracy on complex tasks for high proficiency learners. The trade-off was more prominent in fluency, where it was seen that with increases in task complexity, though complexity and accuracy increased, fluency decreased.

In this study, high proficiency learners gained in complexity by the increase in task complexity. This study has shown that in high proficiency learners, though complexity increases in complex tasks, there was no significant increase or decrease in accuracy. However, in low proficiency learners, lexical variety, lexical density and accuracy showed significant decrease in [+ complex] tasks, a finding similar to Skehan's model.

On all the linguistic measures, high proficiency learners had better scores than the low proficiency learners. In general, high proficiency learners used longer and more error-free T-units (see syntactic complexity scores of the learners). They also used more variety in words (lexical variety) and used more function words (lexical density). In both HP and LP, lexical variety and lexical density dropped in [+complex] tasks.

The results clearly suggest that both groups react to task complexity in different ways, and it affects their linguistic performance differently. Many researchers who believe in a multiple resources theory also seem to suggest that complexity and accuracy compete for attention. For example, Wickens' (1984) claims that when two tasks are being carried out simultaneously and draw from the same resource pool, confusion between tasks may lead to poor performance. However, the fact that complexity and accuracy do not seem to compete for attention would necessitate

an argument that probably attention to micro-level of form (editing, monitoring and error correction) draws from a different resource pool than attention to complexity and variety.

From the results, it can be concluded that proficiency is an important variable and needs to be controlled in such experimentation. Syntactic complexity and variety increase only for high proficiency groups. In case of the low proficiency groups, a significant decrease in lexical density and lexical variety in [+complex] tasks was found when compared to the [−complex] tasks. Accuracy shows a significant difference in the groups with a trade-off between complexity and accuracy for low proficiency learners in resource dispersion dimension. While the results on syntactic complexity, syntactic variety and accuracy in high proficiency learners support Robinson's model of attention and task complexity, the findings on accuracy, lexical density and lexical variety in low proficiency learners lend partial support to Skehan's model of limited attentional memory.

Task Difficulty Questionnaire

Robinson (2001a, b), Gilabert (2005) assessed perceptions of task difficulty, anxiety, interest level, motivation and perception of ability in their study. Results of these studies showed that learners rated the complex version of a task to be more difficult overall, and more stressful than the simple version. It was found in these studies that learner's perception of ability correlated with their fluency in performing the task.

The subjects were administered a questionnaire after each writing task to assess perception of task difficulty (ease, relaxed, confidence, interest and thinking). There were questions on how difficult, how stressful the task was, how confident they (the learners) felt about their performance on the task, how interested they were in doing the task, and whether the task demanded deep thinking or not. The learners had to rate each writing task on the questions on a 5-point Likert scale (5—Strongly agree and 1—Strongly disagree).

High proficiency learners' perception of task difficulty showed a difference in the task conditions though with respect to perception of ease, confidence and interest variables there was no significant difference. The perception of relaxed feeling significantly ($p < 0.023$) decreased in [+complex] conditions. However, thinking significantly ($p < 0.001$) increased in [+complex] tasks.

High proficiency learners required significantly more thinking in [+complex] tasks than [−complex] tasks and found the [+complex] tasks significantly less interesting than [−complex] tasks.

There was positive correlation for ease and relaxed in [−complex] tasks for high proficiency learners (0.57*). However, in the case of low proficiency learners, there was a positive correlation for ease and relaxed in [+complex] tasks (0.64*). In high proficiency learners for interest and thinking, a positive correlation (0.58*) was seen for [+complex] tasks stating that when learners were interested in the task, they involved in more thinking. What is interesting is that in low proficiency learners,

both tasks have shown similar levels of thinking and interest, recording no difference between [+complex] (0.99**) and [−complex] (0.93**) tasks.

Conclusion

In a task-based curriculum, a task need neither be too simple nor be too complex (Nunan, 1989). With the developing interlanguage, the tasks need to become gradually complex so that the task pushes the learners to move from one interlanguage stage to the next. The TBLT subsumes Swain's idea of comprehensible output (1985) which claimed that by producing language in communicative contexts, a learner moves from one stage to the next. Also, in performing a task, the meaning (semantic processing) needs to get a form (syntactic processing), thereby pushing the boundaries of learner's grammatical/discoursal knowledge.

The kind of tasks that learner perform lead them to push their output and test their hypothesis. Complex tasks (tasks with high cognitive load) provide these opportunities. The cognitive load of tasks will be matched by cognitive effort made by the learners in performing the tasks. When the cognitive demands imposed by tasks can be met by learners, the effort to communicate will push the learners to the limits of their current resources and beyond stimulating language development (Robinson, 1995a, b). This has been argued by Krashen in his book, The *Input Hypothesis* (1985) that learners acquire language best by understanding input that is slightly beyond their current level of competence ($i + 1$). Thus, being pushed in output is desirable, and it involves some effort on the learner to analyse the unused or new linguistic forms, thereby leading to interlanguage development.

However, when cognitive load of a task is 'way above' the resources available to a learner, the increased cognitive complexity has detrimental effects—by lowering overall task performance in terms of complexity, accuracy and fluency. Therefore, gradually increasing task complexity (while grading tasks) is necessary in a task-based syllabus so that there is a uniform distribution of cognitive load. When learners are repeatedly exposed to tasks which are complex (in a small quantity), the structures that are to be used in the task (complexity and accuracy) become automatized and release attentional capacity to meet other aspects of the task. Automatization is possible only when learners see patterns in language and language use, and this enables them to organize information and store them as chunks in the long-term memory. When information is processed, sorted and stored in the long-term memory, the stored information is better able to scaffold the learning of new information. Thus, old information uses up less working memory space and requires less cognitive effort. Less load on the working memory results in a higher activation of less accessed linguistic forms and their consequent use (Robinson, 1995a, b).

Not only input and output, but other variables like motivation, confidence, anxiety, etc., also play a key role in acquisition (Dornyei & Otto, 1998). Low motivation and high anxiety can combine to raise the affective filter and form a 'mental block' preventing comprehensible input from being used for acquisition (Krashen, 1985).

With developing interlanguage, tasks need to become gradually complex so that they push the learners to move from one interlanguage stage to the next. Also, being pushed in output is desirable, and it involves some effort on the part of the learner to analyse the unused or new linguistic forms, thereby leading to interlanguage development. Therefore, task sequencing holds major significance in interlanguage development and language learning.

Implications for Teaching of Writing

The *Cognition Hypothesis* has clear implications for classroom instruction. This is because the conceptual and linguistic requirements of a task encourage task selection in the classroom for teaching. It will also help learners understand what aspects are being focused when a task which is based on a particular dimension is taught. This will help teachers in understanding what to focus on and how to focus on different aspects. The results from the study show that learners at different proficiency levels require different kinds of scaffolding and instruction.

Acknowledgements Thanks to the Principal of Kendriya Vidyalaya, Hyderabad School, Class IX students and teachers of Kendriya Vidyalaya, proficiency course students of EFL University and the research scholars of EFL University for participating in this study. Thank you Dr. Geetha Durairajan and Dr. Lina Mukhopadhyay for all the help.

References

Abdollahzadeh, S., & Kashani, A. F. (2011). The effect of task complexity on EFL learners' narrative writing task performance. *Journal of English Language Teaching and Learning, 8*. Retrieved from http://languagejournal.tabrizu.ac.ir/Files/Journal/2012-07-22_11.42.30_1-abdollahzadeh.pdf on August 28, 2013.

Arent, R. (2003). Promoting revision and development in L2 writing through a combination-based curriculum. *The Korea TESOL Journal, 6*(1), 1–26.

Bardovi-Harlig, K. (1992). A second look at t-unit analysis: Reconsidering the sentence. *TESOL Quarterly, 26*(2), 390–395.

Broadbent, D. E. (1958). *Perception and communication.* Pergamon.

Broadbent, D. E. (1971). *Decision and stress.* Academic Press.

Brown, R. (1973). *A first language: The early stages*. Harvard University Press.

Carter, R. (1998). *Mapping the mind*. University of California Press.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition, 11*, 367–383.

Deng, X. (2005). A case study of task complexity & individual learner's oral production. *US-China Foreign Language, 3*(9), 49–54.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70*, 80–90.

Deutsch, J. A., & Deutsch, D. (1967). Comments on " Selective attention: Perception or response?" *Quarterly Journal of Experimental Psychology, 19*, 362–367.

Dornyei, Z. & Otto, I. (1998). Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics, 4*, 43–69. Retrieved from http://eprints.nottingham.ac.uk/39/ on July 18, 2008.

Elder, C., & Iwashita, N. (2005). Planning for test performance: Does it make a difference? In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 219–238). John Benjamins.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition, 26*(1), 59–84. https://doi.org/10.1017/S0272263104026130

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299–323. https://doi.org/10.1017/S0272263100015047

Gathercole, S., & Baddeley, A. (1993). *Working memory and language*. Lawrence Erlbaum.

Gilabert, R. (2005). *Task complexity and L2 narrative production* (Unpublished Ph.D. dissertation). Retrieved from http://www.tesisenxarxa.net/TDX-1220105-085713/index_an.html#documents on December 15, 2007.

Gilabert, R. (2007). Effects of manipulating task complexity in self-repairs during oral production. *International Review of Applied Linguistics, 45*, 215–240.

Guiraud, H. (1954). *Les CaractèresStatistiques du Vocabulaire*. Presses Universitaires de France.

Hunt, K. (1965). *Grammatical structures written at three grade levels*. National Council of Teachers of English.

Ishikawa, T. (2006). The effect of task complexity and language proficiency on task based language performance. *The Journal of Asia TEFL, 3*(4), 193–225.

Iwashita, N., McNamara, T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information processing approach to task design. *Language Learning, 51*(3), 401–436.

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kang, J. Y. (2005). Written narratives as an index of L2 competence in Korean EFL learners. *Journal of Second Language Writing, 14*, 259–279.

Kawauchi, C. (2005). The effects of strategic planning on the oral narratives of learners with low and high intermediate L2 proficiency. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 143–164). John Benjamins Publishing Company.

Krashen, S. (1981). *Second language acquisition and second language learning*. Pergamon Press.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. Longman.

Kuiken, F., Mos, M., & Vedder, I. (2005).Cognitive task complexity and second language writing performance. In S. Foster-Cohen, M. del Pilar Garcia Mayo, & J. Cenoz (Eds.), *Eurosla yearbook* (pp. 195–222). John Benjamins Publishing Company.

Kuiken, F., & Vedder, I. (2007). Cognitive task complexity and linguistic performance in French L2 writing. In M. P. Garcia-Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 117–135). Multilingual Matters.

Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing, 17*(1), 48–60.

Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance and proficiency*. John Benjamins Publishing Company.

Levkina, M. (2008). *The effects of increasing cognitive task complexity along [± planning time] and [± few elements] on L2 oral production* (Unpublished M. A. dissertation). University of Barcelona.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Malicka, A. & Levkina, M. (2012). *Measuring task complexity: Does L2 proficiency matter?* In A. Shehadeh & C. Coombe (Eds.), *Task-based language teaching in foreign language contexts: Research and implementation* (pp. 43–66). John Benjamins.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition, 20*, 83–108.
Michel, M. C., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics, 45*, 241–260.
Nair, V. (2008). *Cognitive challenge in tasks and effects on second language writing* (Unpublished M. Phil. Thesis). CIEFL, Hyderabad, India.
Navon, D. (1989). The importance of being visible: On the role of attention in a mind viewed as an anarchic intelligence system. *European Journal of Cognitive Psychology, 1*, 191–238.
Neumann, O. (1987). Beyond capacity: A functional view of attention. In A. F. Sanders & H. Hever (Eds.), *Perspectives on perception: Action* (pp. 46–70). Erlbaum.
Niwa, Y. (2000). *Reasoning demands of L2 task and L2 narrative production: Effects of individual differences in working memory, intelligence, and aptitude* (Unpublished M.A. dissertation). Aoyama Gakuin University, Tokyo.
Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA. *Applied Linguistics, 30*(4), 555–578.
Nunan, D. (1989). *Designing tasks for the communicative classroom.* Cambridge University Press.
Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition, 21*, 109–148.
Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*(4), 492–518.
Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing, 29*, 82–94.
Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition* (pp. 55–85). Erlbaum.
Quirk, R., & Greenbaum, S. (1973). *A university grammar of English.* Longman.
Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (2010). *A comprehensive grammar of the English language.* Longman.
Rahimpour, M. (1997). *Task condition, task complexity and variation in oral L2 discourse* (Unpublished Ph. D. dissertation). University of Queensland, Australia.
Rahimpour, M. (2008). Implementation of task-based approaches to language teaching. *Pazhuhesh-e-Zabanha-Ye Khareji Journal, University of Tehran, 41*, 45–61.
Rahimpour, M., & Hosseini, P. (2010). The impact of task complexity on L2 learners' written narratives. *English Language Teaching, 3*(3), 198–205.
Robinson, P. (1995a). Task complexity and second language narrative discourse. *Language Learning, 45*, 99–140.
Robinson, P. (1995b). Attention, memory and the 'noticing' hypothesis. *Language Learning, 45*, 99–140.
Robinson, P. (2001a). Task complexity, cognitive resources and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287–318). Cambridge University Press.
Robinson, P. (2001b). Task complexity, cognitive load, and syllabus design. In P. Robinson (Ed.), *Cognition and second language instruction.* Cambridge University Press.
Robinson, P. (Ed.). (2001c). *Cognition and second language instruction.* Cambridge University Press.
Robinson, P. (2001d). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*, 27–57.
Robinson, P. (2003a). Attention and memory in SLA. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 631–678). Blackwell.
Robinson, P. (2003b). The cognition hypothesis of adult, task-based language learning. Second Language Studies, 21, 45–107. Retrieved from http://www.hawaii.edu/sls/uhwpesl/21(2)/Robinson.pdf on 8 May, 2008.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics, 43*, 1–32.

Robinson, P (2007a). *The cognition hypothesis, task design and adult task-based language learning*. Retrieved from http://www.hawaii.edu/sls/uhwpesl/21(2)/Robinson.pdf on13 January, 2007.

Robinson, P. (2007b). Task complexity, theory of mind and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics, 45*, 193–214.

Robinson, P. (2007c). Criteria for grading and sequencing pedagogic tasks. In M. del Pilar & G. Mayo (Eds.), *Investigating tasks in formal language learning* (pp. 7–27). Multilingual Matters.

Robinson, P. (2013). *The Routledge encyclopedia of second language acquisition*. Routledge.

Robinson, P., Cadierno, T., & Shirai, Y. (2009). Time and motion: Measuring the effects of the conceptual demands of tasks on second language speech production. *Applied Linguistics, 30*(4), 533–554.

Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *International Review of Applied Linguistics, 45*, 177–192.

Robinson, P., Ting, S., & Urwin, J. (1995). Investigating second language task complexity. *RELC Journal, 25*, 62–79.

Sadeghi, K., & Mosalli, Z. (2012). The effect of task complexity on fluency and lexical complexity of EFL learners' argumentative writing. *International Journal of Applied Linguistics & English Literature, 1*(4), 53–65.

Sanders, A. (1998). *Elements of human performance*. Lawrence Erlbaum Associates.

Sasayama, S. (2011). Cognition hypothesis and second language performance: Comparison of written and oral task performance. *Second Language Studies, 29*(2), 107–129.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 17–46.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 1–32). Cambridge University Press.

Shiau, Y. S. & Adams, R. (2011). The effects of increasing reasoning demands on accuracy and complexity in L2 oral production. *University of Sydney Papers in TESOL*, *6*, 121–146.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 16*, 43–68.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp.167–185). Pearson Education.

Skehan, P. (2003). Task-based instruction. *Language Teaching, 36*, 1–14.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510–532.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research, 1*, 185–211.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49*(1), 93–120.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language learning* (pp. 183–205). Cambridge University Press.

Skehan, P., & Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (Vol. 11, pp. 193–218). John Benjamins Publishing Company.

Swain, M. (1985). Communicative competence: Some role of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–256). Newbury House.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*, 370–391.

Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 99–119). Longman.

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (Vol. 11, pp. 239–273). John Benjamins Publishing Company.

Treisman, A. M. (1964). Verbal cues, language, and meaning in selective attention. *American Journal of Psychology, 77*, 206–219.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review, 76*(3), 282–299.

Treisman, A. M., & Riley, J. G. (1969). Is selective attention selective perception or selective response? A further test. *Journal of Experimental Psychology, 79*(1, Pt.1), 27–34. http://doi.org/10.1037/h0026890

Wickens, C. D. (1984). Processing resources and attention. In D. Damos (Ed.), *Multiple-task performance* (pp. 3–34). Taylor & Francis.

Wickens, C. D. (1989). Attention and skilled performance. In D. H. Holding (Ed.), *Humans skills* (pp. 71–105). Wiley.

Wickens, C. D. (1991). Processing resources and attention. In D. L. Damos (Ed.), *Multiple task performances* (pp. 3–34). Taylor and Francis.

Wickens, C. D. (1992). *Engineering psychology and human performance*. Harper Collins.

Wickens, C. D. (2007). Attention to the second language. *International Review of Applied Linguistics, 45*(3), 177–192.

Wickens, C. D., & Kessel, C. (1980). Processing resource demands of failure detection in dynamic systems. *Journal of Experimental Psychology: Human Perception and Performance, 6*(3), 564–577. https://doi.org/10.1037/0096-1523.6.3.564

Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing, 14*(1), 85–106.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*, 1–27.

Chapter 9
From Cognitive Grammar to Pedagogic Grammar: Macrostrategies for Designing Form-Focused Tasks

N. P. Sudharshana

Abstract Several studies have emphasized on the need for explicit instruction in ESL/EFL contexts for achieving higher success levels. Accordingly, different kinds of output-based contextualized form-focused tasks (e.g. structure-based communicative tasks, interpretation tasks and consciousness-raising tasks) have been proposed in the task-based language teaching framework. In the field of linguistics, on the other hand, cognitive grammar (CG) is emerging as a better alternative to traditional approaches since CG is able to offer meaning explanations for most grammatical phenomena, traditionally thought of as unexplainable. Though there have been some attempts at applying CG to language pedagogy, they have mostly been individual and discrete. This paper aims to offer a set of macrostrategies to draw more effectively on CG to design contextualized form-focused tasks. It is argued that such strategies help teachers first arrive at a comprehensive set of pedagogical grammar principles which in turn can be exploited in various tasks with a focus on specific forms. The paper illustrates an application of CG to pedagogical grammar in teaching participial adjectives as target items in English.

Keywords TBLT · Form focus · Cognitive grammar · Pedagogical grammar

Introduction

Historically, teaching grammar in ESL/EFL contexts has undergone several transformations as it has been a subject of several debates and theories. In the structural approach, there was an explicit focus on grammar, and largely a deductive method—presenting rules/definitions first and then examples and exceptions, if any—was followed (see e.g. Moulton, 1970 for structural approaches to language and language learning). Once the teacher 'presents' rules, learners are then asked to 'practice' the target form (e.g. sentence structures). Practice activities (e.g. gap filling, matching or MCQs) as discrete exercises were used with the assumption that 'imitation'

N. P. Sudharshana (✉)
Department of Humanities and Social Sciences, IIT Kanpur, Kanpur, India
e-mail: sudh@iitk.ac.in

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_9

and 'practice' would help in language output. With the advent of communicative language teaching (CLT) approaches, the views towards grammar teaching changed. In a strong version of CLT, particularly those inspired by L1 acquisition research (e.g. Krashen & Terrel, 1983; Prabhu, 1987)—explicit teaching of grammar was completely eschewed. It was argued that learners need not be taught grammar explicitly; they just need to be 'exposed' to the target language forms, and in due course, they would acquire the forms just as children do in their L1. However, studies have shown that mere focus on communication may not lead to accuracy in L2 contexts (e.g. Ellis, 1997; Pienemann, 1984; Schmidt, 1990). So, a weaker version of CLT approach has proposed that teaching grammar explicitly may lead learners to 'notice' specific linguistic forms and also notice 'gap' in their linguistic knowledge. This, in turn, may lead to acquisition of implicit knowledge in the long run when they are developmentally 'ready'. Following this, there have been several attempts to integrate meaning focus and form focus in the ESL/EFL contexts.

Form Focus in Communicative Contexts

Different approaches to form focus in communicative contexts have been proposed. Nassaji and Fotos (2011) note that such approaches can be broadly classified into two groups, viz. input-based and output-based. The input-based approaches focus on providing appropriate input and making target forms available to learners for comprehension. It is argued that acquisition happens when learners process the input and output is seen as not necessary. The output-based approaches, on the other hand, focus on creating opportunities for learners to use target forms while performing on communicative tasks. The assumption here is that mere comprehension is not enough and it is necessary for learners to produce target forms to acquire them. The former category includes input processing (e.g. VanPatten, 2004), input enhancement (e.g. Smith, 1991) and noticing (e.g. Schmidt, 1990) approaches among others. The latter includes cognitive-interactionist approaches (e.g. Long, 2015), structure-based communication tasks (e.g. Loschky & Bley-Vroman, 1990) and interpretation and consciousness-raising (CR) tasks (e.g. Ellis, 2015) among others.

Reference Grammar to Pedagogical Grammar

When teachers decide to include grammar in the ESL/EFL pedagogy, an important question is which grammar model they should choose to derive the principles of pedagogical grammar in class. In ESL/EFL contexts, a distinction is made between 'pedagogical grammar', a set of rules and explanations derived exclusively for teaching/learning purposes and 'reference grammar', an extensive descriptive account of rules and regulations (Verity, 2018). Traditionally, grammar taught in most ESL/EFL contexts is still based on the structural approaches to language. In this

approach, categories (e.g. parts of speech) and structures (e.g. tenses) are defined, and their 'correct uses' are prescribed. The views preseted in 'standard grammar books' and 'standard dictionaries' have to be strictly followed since any deviations from these norms are 'errors'. However, traditional grammar is inadequate: it does not offer comprehensive explanations for several grammatical phenomena (e.g. use of passive voice), many aspects of grammar (e.g. prepositions) are considered 'idiosyncratic', and no meaning-based explanation is given (see e.g. Lindstromberg, 2001 for an analysis of prepositions). There have been some attempts to draw on generative grammar; however, they are largely restricted to syntax and lay over-emphasis on L1-L2 differences.

Cognitive grammar (CG), an umbrella term for various theories in cognitive linguistics, of late has emerged as a better alternative. CG has been found to be most relevant for language pedagogy among other linguistic theories since some of its salient principles are in congruence with communicative approaches to language teaching. Building on this aspect, this paper proposes a framework to draw on CG to arrive at a comprehensive pedagogical grammar which in turn can be exploited to design and use form-focused tasks.

The paper is organized thus: Section "Introduction" introduces basic principles of CG; Section "CG and Language Pedagogy" reviews previous studies in CG and language pedagogy; Section "Macrostrategies for Drawing on Principles of CG" proposes a set of macrostrategies for effectively drawing on principles of CG; Section "From CG to PG and Actual Classroom Tasks: An Illustration" presents illustrative tasks on participial adjectives; finally, Section "From CG to PG: Challenges and the Way Forward" indicates some possible challenges in implementing the CG-based framework and provides suggestions for future research.

Principles of Cognitive Grammar

In the late 1980s, several cognitive approaches to language were proposed. Prominent among them is cognitive grammar (CG) (Langacker, 1987, 1991, 2008a). The cognitive approaches have the common premise that language is part of general cognition. In this regard, this theory contrasts with generative grammar which argues that language faculty is separate and distinct. Within this theory, three principles have been argued to be most relevant for pedagogy (see e.g. Langacker, 2008b). These are discussed briefly below.

Grammar is Meaningful

In CG, each linguistic unit (vocabulary and grammatical structures included) has two dimensions—phonological and meaning. From this follows the most important claim, 'grammar is meaningful' (see e.g. Langacker, 2008a). In CG, there is no

distinction between 'vocabulary' and 'grammar' since both are 'meaningful'. This contrasts sharply with the traditional approach where a distinction is made between 'content' and 'function' words. The only distinction made in CG is that the meaning of some elements is less abstract while that of some others may be more abstract. This factor cuts across traditional categories of words known as 'parts of speech'. For instance, the noun 'honesty' may be more difficult to define than the preposition *on* (when used in a sentence like 'The books are placed on shelves in the library'). The principle that grammar is meaningful has a significant implication for pedagogy: this can reduce the list of exceptions and minimize rote learning associated with 'random' and 'unexplainable' grammatical phenomena.

Meaning is Conceptualization

If all linguistic elements are meaningful, what is the nature of this meaning? In CG, meaning making is a conceptualizing activity. In other words, the meanings are not out in the world, but they are formed in our mind based on sensory experiences. We 'perceive' the outside world through our senses, primarily our eyes. When we view a particular scene, several cognitive processes such as 'selection', 'abstraction', 'schematization' and 'categorization' come into play. Let us imagine there is a coffee cup on the dining table. Speakers of English will probably describe this scene as 'a coffee cup on the dining table'. But would they ever say something like: 'the dining table under a coffee cup'? The speakers of English 'construe' this spatial scene as a small object (here 'a coffee cup') being *placed on* and *supported by* 'a larger horizontal surface' (here 'the dining table'). This particular schema (a highly abstract representation) is encoded by the preposition *on* in English. When native speakers come across similar situations later, their viewing is governed by the preposition *on.*

The argument that meaning is conceptualization has two implications. First, there is no one-to-one correspondence between what is out in the world and how it is encoded in language. This means a particular scene may be 'construed' and 'encoded' differently within the same language and across languages. For instance, if John is at the front and Mary at the end of a queue, both these sentences are possible in English depending on the perspective imposed by the speaker—'John is in front of Mary' and 'Mary is behind John'. Regarding cross-linguistic variations, in English, the prepositions *on* and *in* encode 'support' and 'containment' scenes, whereas comparable situations are encoded with the single preposition *en* in Spanish (Choi amp; Bowerman 1991; Bowerman & Choi, 2001: 485).

The second implication is that most lexical items have multiple conceptualizations associated with them and hence are polysemous. Usually, there is a 'core' and 'prototypical' sense, and then, there are meaning extensions based on knowledge structure, metaphor, metonymy and image schemas (Lakoff, 1987: 113–114). For instance, in case of spatial uses of the preposition *on,* the scene described above (a

smaller object supported by a larger horizontal surface) is the core prototypical sense from which the following extensions[1] are made:

(1)

a. a clock **on the wall** (vertical support against gravity);
b. a bandage **on the leg**, a balloon **on a stick** (support with adhesive/string);
c. a ring **on the finger** (a piece of jewellery supported by human body, suggesting around);
d. a cork **on the bottle** (part–whole relationship);
e. an apple **on the skewe**r (impalement).

Language is Usage-Based

In CG, though meanings are formed as part of conceptualization process, it is not a mental exercise; it is actually rooted in the actual instances of language use. It is argued that in reality each and every use of a word or a structure is based on the specific context and without context it is not possible to interpret any word or structure. The term 'context' here refers to spoken and written discourse. Naturally, a discourse includes information other than pure linguistic usage such as sociocultural conventions, prior knowledge, worldly affairs and power relations between interlocutors among others. All of these (termed 'encyclopaedic knowledge') is part of the meaning of a particular word or structure. Therefore, in CG, no distinction is made between 'original (denotative)' and 'extended (connotative)' meaning of a word or structure. For instance, the word 'dove' is said to literally mean a bird and having the connotations of 'purity' or 'peace'. In CG, all are part of the meaning of the word 'dove'.

This has a far-reaching implication for pedagogy: in most cases, dictionaries are not helpful in fully understanding a word or a structure since they list only 'denotative meaning' and leave 'connotative meaning' to be interpreted from the contexts where the word may occur.

CG and Language Pedagogy

The basic three principles discussed above have been found to be relevant for language pedagogy. As a result, many researchers have attempted to use CG to offer 'meaningful' explanations for grammatical phenomena hitherto thought of as

[1] 'Prototype(s)' or 'prototypical member(s)' refers to the best example of a particular category. 'Radial categories' emerge as extensions from this prototype based on some common features. The extended categories may not share any common feature(s)' (Sudharshana, 2018: 145). See Lakoff (1987) for details.

'arbitrary' and 'full of exceptions'. A few examples of such studies are summarized below.

Applying CG to Language Pedagogy

The studies in this area can be categorized into two broad groups: First, there have been several descriptive accounts of English grammar based on the CG principles (e.g. Lindstromberg, 2010; Radden & Dirven, 2007), which have attempted to explain traditional grammatical categories (such as noun, verb and prepositions), structures (such as passives, relative clauses) and also have looked at discourse as a whole. These accounts offer meaning explanations for linguistic elements using cognitive linguistic tools and analyses.

The second set of studies have offered explanations for grammatical elements and have also attempted to teach them in second/foreign language classroom contexts. An area that has been extensively studied is prepositions. Lindstromberg (1996) proposes a framework to teach the preposition *on* combining explanations from CS and Total Physical Response method. The methodology consists of using pictures which are schematic representations of spatial relations between objects; starting with core meaning and then explaining extensions from it and comparing and contrasting uses of closely related prepositions to highlight differences in conceptualization. Other areas include tense and aspect (Bielak & Pawlak, 2013) partitive particles (Achard, 2008), modal verbs (Tyler, 2008), metaphor awareness (Boers, 2004; Boers & Demecheleer, 1998), polysemy of verbs/phrasal verbs (Csabi, 2004) and polysemy of prepositions (Tyler & Evans, 2004) among others. The full review is beyond the scope of this paper. See Boers and Lindstromberg (2006) and Putz (2007) among others for a comprehensive review of studies.

Macrostrategies for Drawing on Principles of CG

The studies discussed above largely focus on individual areas and are experimental in nature. There is a need for a broad or a macroframework to derive principles of pedagogical grammar from the principles of CG. Such a framework will enable practicing teachers understand how to use CG more effectively. Also, it will reduce rote learning regarding grammar rules, common practice in many ESL/EFL contexts. These macrostrategies (MS) are also likely to help teachers integrate CG with task-based approaches to teaching and learning (see Nunan 1998, 2004 among others for details on teaching grammar in TBLT frameworks).

MS1: Present Grammar as a Meaningful Phenomenon and not as a Set of Idiosyncratic Rules

In CG, all grammatical categories (e.g. nouns, verbs) and structures (e.g. passives) are meaningful. The grammar can be explained in terms of image schemas (highly abstract representations based on sensory perceptions), metaphor and metonymy, prototypes and extension categories, form-form and form-meaning connections and force dynamics.[2] The differences among the prepositions *on, in* and *at* in the following sentences, for instance, can be explained using image schemas:

(2)

a. The coffee cup is on the table.
b. The book is in the bag.
c. We met the guests at the reception.

The preposition *on* conceptualizes the reference object (here 'table') as a two-dimensional horizontal surface; the preposition in (here 'bag') as a three-dimensional container; and the preposition *at* (here 'reception') as a one-dimensional point-like object. These uses are prototypical 'spatial' uses from which 'temporal' uses emerge as extensions. For instance, the preposition *in* is used with relatively longer durations (e.g. in a year/summer/2019), *on* with shorter durations (e.g. on Sunday/24 of this month) and *at* with the shortest (e.g. at 3 pm).

While all of the above can be clubbed under synchronic motivations, in CG, diachronic motivations[3] are also invoked particularly to explain idioms and proverbs. For instance, the idiom 'make no bones about something' means 'have no difficulty/hesitation to talk about something'. Historically, it seems to have originated from a culinary context: if there are no bones in a soup bowl, it is easy to consume it without any difficulty/hesitation. By metaphoric extension, this can refer to non-culinary contexts to mean 'if there are no bones, then there are no problems talking

[2] 'Metaphor' refers to conceptualising an abstract domain in terms of a relatively more concrete domain. For instance, emotions are often described in terms of 'liquids' and human body as 'a container'; cf: children are brimming with joy. 'Metonymy' is using a part to represent the whole or vice versa. For instance, in the sentence 'The Supreme court has rejected bail application of the terror-accused', the whole ('The Supreme Court') refers to a part ('One or more Supreme Court judges'). 'Force dynamics' refers to 'the linguistic representation of force interactions and causal relations occurring between certain entities' within a specific situation (Talmy, 2000: 12). For instance, in the sentence 'John's under a lot of pressure to perform', an outside authority is exerting force on John who is in a relatively subordinate position. Form-form connections refer to resemblance of sound between words in phenomena such as alliteration (e.g. Donald Duck), assonance (e.g. Jack and Jill) and rhyming (e.g. see you later, alligator). Form-meaning connections refer to similarity in spelling/pronunciation among words that express similar/related set of meanings. For instance, most words describing a sudden and forceful action in English start with the consonant cluster 'sp' as in *spit, spike, splutter,* etc.

[3] 'Synchronic motivation' refers to explaining the word meanings based on metaphor, metonymy or image schemas among others. 'Diachronic motivation' refers to explaining the word meanings based on historical usage, etymology or language change.

about an issue'. See Upadhaya & Sudharshan (ch 10) in this volume for details on metaphor.

MS2: Introduce Grammatical Elements in Context

It has been a general argument in the communicative approach that grammar be introduced in context and teachers move away from discrete point exercises used in the structural approach. However, in the CG, this has a greater significance. One of the main arguments, as we noted above, is that grammar does not exist without a context, an actual instance of spoken or written language use. What does 'introduce grammar in context' actually mean? The teacher needs to select oral/written texts carefully so that they contain adequate representation of the target grammatical element(s). If a single text is not sufficient, a 'pedagogic corpus' (Willis & Willis, 2007) may be developed. Teacher can collect or ask texts from various sources and of different genres that have occurrences of the target element in many kinds of contexts. In fact, teachers can ask learners to bring texts to the classroom.

The teacher can also use techniques of input flooding and/or highlighting. These strategies are argued to induce learners to notice target elements which in turn may lead to successful acquisition (see Nassaji & Fotos, 2011, ch 3 for details). The texts should not serve just as a pretext to introduce a grammatical item; in fact, the target element has to be explained based on its occurrence(s) in the texts, and further tasks also need to 'focus' on the usage in similar contexts. Thus, highlighting as a noticing cognitive strategy will help. For instance, passive voice (e.g. The book was written by John), for instance, is generally introduced in the popular grammar and course books simply as one of the structural options available to users. However, in real life, passive voice is used specifically to hide the information about the doer ('subject') because that information is irrelevant (e.g. After 6 months, essentials were dispatched to the valley), obvious (e.g. The bank was robbed at the gunpoint yesterday), or the speaker wants to hide that information purposefully (e.g. 'The check has been sent', a White House official about money sent to a fallen soldier's father). Passive voice may be grammatically correct but inappropriate in contexts where the focus has to be on the doer (e.g. The organizations are actively seeking volunteers vs Volunteers are being sought actively by organizations). Instead of merely presenting rules about active to passive conversion, this kind of contextual analysis will help learners understand why passive is used in real life and what are the different occasions in which it can be used.

MS3: Explain Each Grammatical Construction as a Matter of Choice and Show that Other Plausible Alternatives Exist in Most Cases

CG observes that there is no one-to-one correspondence between outside world events and how they are encoded in a language. This may be because of our perceptual processes. For instance, the following two sentences are different in terms of 'zooming in' and 'zooming out' though they have the same set of words:

(3)

a. An empty box on the shelf in the kitchen
b. In the kitchen on the shelf an empty box

The sentence (3a) starts with a narrow focus ('an empty box') and then expands the backdrop area gradually. In contrast, (3b) starts with a broad focus and then zeroes in on a specific object. The second sentence foregrounds 'an empty box' and is used when the speaker wants to add more about the object (say 'In the kitchen on the shelf lay an empty box. What it contained nobody knew').

This may also be because of available linguistic means in that particular language. Sometimes there is more than one way of linguistically representing the same situation (e.g. John visits his aunt regularly vs John is visiting his aunt regularly), and in some cases, a language may have limited means of encoding a specific set of events. Therefore, teachers need to explain why a particular construction has been used in the given context. One can also explore other possibilities and examine if the interpretation differs. For instance, one can say, 'the disease has spread through a community' or 'spread across a community'. The former focuses on how intense the disease has been while the latter focuses on how widespread it has been. This again emphasizes the need to introduce grammar in context.

MS4: Move from 'Correct–Incorrect' Paradigm to 'More Appropriate–Less Appropriate'

In an extension of MS3 above, teachers need to move from prescriptive approaches to descriptive approaches. Traditionally, it is stated that one 'has' to use simple present—and not present continuous—to talk about one's occupation (e.g. John works as a teacher). However, we need to note that using present continuous is not ungrammatical (e.g. John is working as a teacher), but the interpretation is different: while the use of simple present refers to something more permanent, the use of present continuous can refer to a temporary arrangement. Therefore, there is a strong need for teachers to stop categorizing usage examples as 'correct' and 'incorrect'; instead, teachers need to emphasize that a particular construction may only be 'more appropriate' or 'less appropriate' in the given context. In some cases, a structure

may be grammatically correct but may not be appropriate in the given context or vice versa. For instance, sentences like 'These are books', 'Those are bags' do not make much sense without a proper context though they are grammatically correct. Similarly, in casual speech, one may find many instances that supposedly violate grammar rules (e.g. progressive use of a stative verb like 'love' in: 'I'm loving it').

MS5: Present Related/Competing Grammatical Elements Together

This is in continuation of MS2 and MS3 above. We have noted above that often there are multiple options for representing the given situation in a language. It may be useful to discuss competing/interchangeable grammatical elements together so that their usage patterns become clear. The same strategy can be adopted in the case of polysemy. In CG, it is argued that there is a 'core meaning' and then there are several extended uses. For purposes of illustration, let us look at present participle (-ing participle) and past participle (-ed participle). Both have many uses and occur in comparable syntactic contexts: (i) they help encode 'aspect' in English with auxiliary verbs (-ing participle is used in progressive tenses whereas -ed participle in perfect tenses); (ii) both are used as attributive (The movie is boring,) and predicative adjectives (John is interested in history); and (iii) both can occur in non-finite clauses in complex sentences (e.g. Walking along the pavement, John saw a sick puppy; Silenced by his own party members, the mayor walked out of the meeting). Because of these similarities, particularly as adjectives, participles are known to create difficulties for second/foreign language learners. While traditional approaches treat these uses as different, in CG, they are discussed together (e.g. Radden & Dirven, 2007). See Section "From CG to PG and Actual Classroom Tasks: An Illustration" below for more details on using participles in form-focused tasks.

MS6: Raise Language Awareness

In the ESL/EFL pedagogy, several researchers have highlighted the advantages of fostering language awareness (e.g. Ellis, 2003; Kumaravadivelu, 2001). In CG, this acquires a new dimension. As noted above, in CG, it is assumed that each and every grammatical element and structure has a specific meaning. While focusing on this aspect, teachers can foster language awareness of learners. For example, wherever necessary, cross-linguistic comparisons can be made; cf: 'being cool headed' in English vis-à-vis *thande dhimaag-se soochna* 'think with a cool head' in Hindi. Suitable tasks may be designed to focus on cross-linguistic variations and thus raise metalinguistic awareness. CG can also help increase 'critical language awareness'. For instance, many people find the phrase 'a disabled person' offensive and instead

prefer to use 'a person with disabilities'. This is because a past participle (here 'disabled') when used before a noun (e.g. the disabled soldier)denotes a more permanent state. In contrast, the phrase 'a person disabled in a war' where the past participle is used after a noun refers to a specific event and does not attribute a permanent quality to the subject. A branch of this field is 'critical metaphor analysis' where metaphor use in every-day discourse is analysed to investigate inherent biases, if any, against a particular race, religion or a community (see e.g. Charteris-Black, 2006).

From CG to PG and Actual Classroom Tasks: An Illustration

In this section, I attempt to illustrate how one can move from CG to pedagogical grammar and further design form-focused tasks in communicative contexts. The target elements chosen are present and past participial adjectives.

Present and Past Participial Adjectives

English has two participles, viz. '-ing' and '-ed'/'-en'. The former is known as 'present participle' and the latter as 'past participle'. The 'ing' participle occurs in a variety of contexts as shown in (4) below:

(4)

a. The children are practising for a recital. (progressive tense)
b. Do you know three *interesting* facts about coffee? (attributive adjective—prenominal)
c. The three things *interesting* about coffee are mentioned in this book. (attributive adjective—post-nominal)
d. The critics have found the recently released movie quite *interesting*. (predicative adjective)
e. *Walking* along the pavement, I bumped into an old friend of mine. (subordinate clause conjunction)

The '-ed/-en' participle occurs in various contexts as shown in (5) below:

(5)

a. The new tax law has *benefitted* millions of people in the country. (perfect tense)
b. The college administration will replace all the *broken* chairs. (attributive adjective—prenominal)
c. The chairs *broken* during the recent student agitation are in the old classroom. (attributive adjective—post-nominal)
d. The critics are not *interested* in reviewing the recently released movie. (predicative adjective)

e. The bank on our university campus was *robbed* last night. (passive)

The adjectival uses of these participles are particularly found to be challenging for ESL learners. For instance, Williams and Evans (1998: 140) observe that learners interchange present and past participial forms: cf. *I am so boring in school* (meant to say 'I am so bored in school').

How are these treated in non-CG approaches? Based on some popular grammar books in ESL/EFL contexts (e.g. Azar & Hagen, 2017; Celce-Murcia & Larsen-Freeman, 1999; Larcen-Freeman & Celce-Murcia, 2016), the following points can be noted about these:

(6)

a. Participles are derived from verbs by adding '-ing' and '-ed'/'-en' to the base forms.
b. Adjectival participles are different from their verbal counterparts. That is, the '-ing' participle in progressive tenses is different from the adjectival uses (e.g. The magician is amusing us vs The magician is amusing). Similarly, the '-ed/-en' perfective participle in passives is different from the adjectival uses (e.g. The vase is broken by John vs The vase is broken).
c. The '-ing' participle has an active meaning; the subject does the actual action or has the quality talked about. For instance, in the sentence 'The movie was boring', the subject 'movie' possesses the quality of being 'boring to watch'. The participle does not mention anything about the experiencer/object (here in the above sentence who finds the movie boring is not specified).
d. The '-ed/-en' participle has a passive meaning; the subject is actually an experiencer/object. For instance, in the sentence 'John was bored', the subject 'John' is the experiencer. The participle does not mention anything about the doer/subject (here in the above sentence who makes John bored is not specified).

There are a few major issues here. The traditional accounts present adjectival uses of participles as different from other uses. However, such a distinction is not always clear as these accounts themselves admit. They explain 'the meaning' of adjectival uses (6c-d above) but do not explain how they get this meaning.

On the other hand, CG analyses participles holistically, and there is an attempt to explain various uses as interconnected. The CG account of participles is presented in the next section.

Participles in CG

A major issue is arriving at a rule explaining the major difference between the two participles as adjectives. Both of them appear in almost similar syntactic contexts as is evident from examples in (4b-d) and (5b-d) above. This may probably be the reason for learner errors observed by Williams and Evans (1998). How does CG explain the uses of participles? The participles have received considerable attention

in the CG (see, for instance, Langacker, 1987, 2001, 2008a, 2009; Raden & Dirven, 2007). Langacker (2008a) observes that while a verb 'profiles a complex relationship scanned sequentially through conceived time' (p. 122), a participle abandons such a sequential scanning and takes up a 'holistic construal of the verbal process' (p. 120). In other words, while verbs encode an action/state in a time lapse or completion manner, the participles lose the time lapse dimension and instead present a summary view of an event or the event becomes like a static property. Thus, through a common principle, both the participles are derived from verbs.

How are the two participles different from each other? The participles differ in the kind of specific perspective they impose on the events: the '-ing' participle has 'limited immediate scope' or 'an onstage region' of a longer process (Langacker 2008a: p. 120); the '-ed' participle 'looks at an action from the past' (Radden & Dirven, 2007: 155) and 'refers to situations that are finished or completed' (Radden & Dirven, 2007: 155).

As mentioned earlier, CG looks at participles holistically: the adjectival uses are not treated as discrete but are connected with other uses in a radial network of categories. As we noted earlier in (4) above, the '-ing' participle is used in progressive tenses. This is because the participle focuses on here and now part of a long event, and the focus of progressive tenses is also the same. This core sense of 'ongoing process' leads to use of '-ing' in clauses that serve as backdrop (as a durative) against another event: cf. *Walking in the park, John and Mary spent quality time together.* When we need to note that the 'backdrop' incident is complete, we use '-ing' participial form of the perfective auxiliary 'have': cf. *Having completed their walking in the park, John and Mary went back home.* This focus on 'ongoing' process helps explain its adjectival uses. The present participle focuses on the 'ongoing part of a longer process' and by metonymic extension on 'Trajector' of the action. 'Trajector' refers to the person/object in primary focus in an event. In contrast, the person/object which has secondary focus or serves as a reference point is 'Landmark' (see Langacker, 1987 among others for more details on 'trajector' and 'landmark'). Therefore, when the speakers want to emphasize on the agent of an action, they use present participles or 'active participles' (Langacker, 2009: 237): cf. *The movie is quite interesting* (here 'the movie' is 'trajector', and the landmark 'people' is implicit).

The past participle focuses on a completed event or a resultant stage; therefore, it is used in perfective tenses: cf. *John and Mary have completed their tasks.* By metonymic extension, the past participle focuses on the person/object affected by the action or 'Landmark'. Therefore, it is used in the passive construction: cf. *John has been sent home by his teacher.* The past participle highlights 'the end of the verbal process' or by metonymic extension 'the final participant or the final state' (Langacker, 2008a: 121). Therefore, when the focus is on the 'landmark' or those who are affected by the action, the past participle or 'passive participle' (Langacker 2008a: 118–122) is used: cf. *The people are not interested in this movie* (here, 'the people' are 'landmark' who are affected by the implicit trajector 'the movie).

We can see that CG not only gives explanations for adjectival uses of participles but also sees various uses of participles as related. CG also helps explain a few issues that are usually neglected in traditional grammar. While both the participles can be

used in prenominal and post-nominal attributive position, there is a subtle difference in interpretation. Let us look at the examples below:

(7)

a. a killing machine versus a machine killing people;
b. a damaged building versus a building damaged during the earthquake.

Syntactically, in the prenominal position, there is no complement, while a complement is compulsory in the post-nominal position: cf. *a machine killing. The prenominal participle, particularly the past participle, focuses on 'the stable result' (Radden & Dirven, 2007: 155–6), and the post-nominal participle focuses on 'the temporary event causing the result' (Radden & Dirven, 2007: 155–6). This explains the difference in interpretation between these two. The prenominal creates a more stable and permanent attribute for the modified. That is why, some phrases such as 'disabled person' are considered socially inappropriate. The post-nominal refers only to a single process and therefore does not create a strong and permanent attribute for the modified: cf. *a person disabled in an accident*.[4]

Task Design and Sequencing

Now, drawing on the CG account of participles explained in the previous section and using macrostrategies explained in Section "Macrostrategies for Drawing on Principles of CG above, some sample tasks are discussed in this subsection. This paper chooses interpretation tasks and CR tasks (Ellis, 1995, 1997, 2003, 2015) to illustrate form-focused tasks on participles. In an interpretation task, the focus is on comprehension of target items. Learners are presented with oral or written texts, and tasks require learners to respond non-verbally or with minimum language. For instance, there may be a poem on a child feeling happy going to school the first time, and a task may require the learners to circle 'a happy face' or 'a sad face' against a set of questions such as—How does the child feel when playing with friends? Here,

[4] Now, an interesting question—is there any link between the present participle and the gerund? After all, both have the same morphological form (e.g. *a working mother—Working hard ensures success*), and both are derived from verbs (here 'work'). Langacker (2008a) observes that things (that are traditionally referred to as 'nouns') and relationships (that are traditionally referred to as 'verbs') that have become atemporal (e.g. participles) 'represent a natural grouping since both construe a situation in summary fashion.' (p. 119). Participles still focus on the process. If the focus shifts from the process to a thing, they become nominalized. This shifting can happen in one of the two ways: (i) focus on the final participant in case of past participles (e.g. the betrothed, the damned); or (ii) conceptualizing the whole process as a single entity (Langacker, 2008a: 120). The latter can be clearly observed in deverbal nouns. For instance, the verb 'perform' encodes several things. Usually, there is a performer and an audience, and then, the actual act has a beginning and an end. The noun 'performance' blurs all the stages during the act and conceptualizes the entire episode as a whole: cf. *The performance of the singer in the last hour was below expectation.* It must be noted here that some stative verbs are not used as gerunds (e.g. interesting, boring). This issue needs more detailed analysis and is beyond the scope of this paper.

the focus is on adjectives 'happy' and 'sad'. In a CR task, on the other hand, the focus is on production and metalinguistic analysis. Learners are presented with an oral or written text which has ample occurrences of the target elements, and while reading their attention is diverted to target elements through various techniques/tasks such as underlining. They are asked to guess why a particular structure/word is used in a specific way. They may be given additional texts and asked to verify whether or not their guesses are correct. For instance, students may be presented with a text that has many occurrences of the antonyms with 'in-' and 'un-' prefixes and asked to find out patterns of distribution (see Ellis, 2003: 160–167 for more details including examples of these tasks).

The tasks on participles can be designed keeping in mind the following principles. The relevant macrostrategy for each principle is also listed.

(8)

a. Present both past and present participles together. Since participles occur in a variety of domains, teachers may choose any one domain (e.g. adjectival uses) but present both together. (MS5)
b. Choose interesting theme(s) and text(s) according to the level of the target group of learners. Do not present single decontextualized examples. For instance, to present adjectival uses of participles teachers can choose the theme of 'hunting for a job' with a group of young adults. This leads to a discussion on 'interesting/boring jobs'. (MS2)
c. Link different uses of participles by designing a series of tasks. For instance, the first one could be on verbal participles in perfective tenses, the second one on passives and the third on adjectival uses. (MS 1)
d. Include tasks that encourage learners to guess the rationale behind uses of participles in the given contexts. Teachers can interchange elements in examples (e.g. a disabled veteran vs a veteran disabled in a war) and ask students to guess if there is any change in the interpretation. (MS 6)

Some sample tasks and their suggested sequence are listed in Appendix 1 below. These tasks are designed for learners at the intermediate level of proficiency (A2 to B1 of CEFR), and the theme is 'hunting for an ideal job' (see Sudharshana, 2017 for details on tasks).

The first task in the sequence could be an interpretation task that requires learners to simply put a tick or cross mark. Learners may be given a questionnaire with options such as 'I am inspired by the work of space scientists' and could be asked to tick one of the options—always, sometimes or never. This task introduces '-ed' forms of adjectives, and learners need to comprehend them before they answer.

The second task is a minimal production task and builds on Task 1. This introduces '-ing' forms. Learners are asked to list some jobs that they find 'interesting/boring/amazing', etc. This again requires comprehension of target elements, but learners need not use them in their speech.

The third task is another interpretation task. Here, a dialogue or a text that includes both kinds of participles can be given to learners. The comprehension questions

that demand minimal verbal responses need to be set. However, the comprehension questions should necessarily encourage learners to understand target elements before answering the questions.

The fourth and fifth tasks are CR tasks. Here, learners are presented with reading texts that include ample instances of both participles and are asked to underline the target elements. Learners can later be asked to guess the meaning of the underlined expressions.

From CG to PG: Challenges and the Way Forward

Designing form-and-meaning focus tasks using CG principles has several advantages. This approach deemphasizes rote learning of grammar and makes way for meaningful teaching of several grammatical phenomena which are not dealt with in traditional grammar. However, there are some issues which still need further exploration. The biggest challenge is to arrive at a comprehensible explanation of grammatical elements that is easily accessible to teachers and learners. This issue is more relevant when teaching young learners. Second, there is a need for experimental studies to try out CG-inspired pedagogic methods in actual classroom and examine their efficacy vis-à-vis traditional approaches. Also, though a lot of research has happened in task design, it is difficult for practicing teachers to ensure form focus in communicative contexts. Teacher training programmes need to address these issues. Nevertheless, it is important to recognize that CG is a good and meaningful alternative available to make grammar teaching more meaningful and less stressful.

Appendix 1: Sample Form-Focused Tasks Based on CG

CR Task 1

Read the following text and list present and past participials separately. On the basis of the list you have made, can you identify contexts in which two kinds of participials occur? Can you form a rule about their uses?

Do you like your current job? According to a recent study, about 70% of people are not happy with their jobs. Also, majority of people spend a large part of their day at work, yet they are completely dissatisfied. What could be the reason for this dissatisfaction? Some experts say that the answer could be in how people choose their career options and jobs.

People are usually not aware of what the job is actually about and whether they can really manage it while applying for it. Job aspirants look at an ad and decide whether to apply or not mostly on the basis of salary and promotions. While the remuneration and perks are important, they cannot sustain a person's interest in a

job for a long period. What is more important is that a person should find the work environment rewarding and interesting.

CR Task 2

Read carefully each pair of sentences below and identify contexts for the use of '-ing' and '-ed' participial adjectives. Does the rule you have formed in Task 2 above apply to these sentences? Why/why not?

1. (a) The court expressed concerns over the frightening COVID situation in the country.
   (b) A family with a frightened kitty waited for a rescue team for 12 hours amidst raging wildfires.

2. (a) The fans are excited after the star shared the poster of his upcoming movie on Twitter.
   (b) Women navy officers who will be the first combat warriors said that the job profile is exciting and thrilling.

3. (a) It is time to switch to our burgundy hair colour from your boring black.
   (b) Are you bored of lock-down routine? It is time to binge-watch all episodes of our latest Web series.

References

Achard, M. (2008). Teaching construal: Cognitive pedagogic grammar. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive linguistics and second language acquisition* (pp. 432–455). Routledge.

Azar, B. S., & Hagen, S. A. (2017). *Understanding and using English grammar* (5th ed.). Pearson Education.

Bielak, J., & Pawlak, M. (2013). *Applying cognitive grammar in the foreign language classroom: Teaching english tense and aspect*. Springer-Verlag.

Boers, F. (2004). Expanding learners' vocabulary through metaphor awareness: What expansion, what learners, what vocabulary? In M. Achard & S. Niemeier (Eds.), *Cognitive linguistics, second language acquisition, and foreign language teaching* (pp. 211–232). Mouton de Gruyter.

Boers, F., & Demecheleer, M. (1998). A cognitive semantic approach to teaching prepositions. *ELT Journal, 52*(3), 197–204.

Boers, F., & Lindstromberg, S. (2006). Cognitive linguistic applications in second or foreign language instruction: Rationale, proposals and evaluation. In G. Kristiansen, M. Achard, R. Dirven, & F. J. R. Ibanez (Eds.), *Cognitive linguistics: Current applications and future perspectives* (pp. 305–355). Mouton de Gruyter.

Bowerman, M., & Choi, S. (2001). Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 475–511). Cambridge University Press.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course*. Heinle & Heinle Publishers.

Charteris-Black, J. (2006). Britain as a container: Immigration metaphors in the 2005 election campaign. *Discourse & Society, 17*(5), 563–581.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 41*(1–3), 83–121.
Csabi, S. (2004). A Cognitive linguistic view of polysemy in English and its implications for teaching. In M. Achard & S. Niemeier (Eds.), *Cognitive Linguistics, second language acquisition, and foreign language teaching* (pp. 233–256). Mouton de Gruyter.
Ellis, R. (1995). Interpretation tasks for grammar teaching. *TESOL Quarterly, 29*(1), 87–105.
Ellis, R. (1997). *SLA research and language teaching*. Oxford University Press.
Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
Ellis, R. (2015). *Understanding second language acquisition* (2nd ed.). Oxford University Press.
Krashen, S. D., & Terrel, T. D. (1983). *The natural approach: Language acquisition in the classroom*. Prentice Hall Europe.
Kumaravadivelu, B. (2001). Toward a post method pedagogy. *TESOL Quarterly, 35* (4), 537–60. Accessed April 29, 2020 from http://www.bkumaravadivelu.com/articles%20in%20pdfs/2001%20Kumaravadivelu%20Postmethod%20Pedagogy.pdf
Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press.
Langacker, R. W. (1987). *Foundations of cognitive grammar, vol. 1: Theoretical perspectives*. Stanford University Press.
Langacker, R. W. (1991). *Foundations of cognitive grammar, vol. 2: Descriptive application*. Stanford University Press.
Langacker, R. W. (2008a). *Cognitive grammar: A basic introduction*. Oxford University Press.
Langacker, R. W. (2008b). Cognitive grammar as a basis for language instruction. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 66–88). Routledge.
Langacker, R. W. (2009). *Investigations in Cognitive grammar*. Mouton de Gruyter.
Langacker, R. (2001). Discourse in Cognitive Grammar. *12*(2), 143–188. https://doi.org/10.1515/cogl.12.2.143
Larcen-Freeman, D., & Celce-Murcia, M. (2016). *The grammar book: Form, meaning, and use for English language teachers*. NG Learning.
Lindstromberg, S. (1996). Prepositions: Meaning and method. *ELT Journal, 50*(3), 225–236.
Lindstromberg, S. (2001). Preposition entries in UK monolingual learners' dictionaries: Problems and possible solutions. *Applied Linguistics, 22*(1), 79–103.
Lindstromberg, S. (2010). *English prepositions explained*. John Benjamins Publishing Company.
Long, M. (2015). *Second language acquisition and task-based language teaching*. Wiley Blackwell
Loschky, L. & Bley-Vroman, R. (1990). Creating structure-based communication tasks for second language development. *University of Hawai'i Working Papers in ESL, 9* (1), 161–212.
Moulton, W. G. (1970). *A linguistic guide to language learning* (2nd ed.). The Modern Language Association of America.
Nassaji, H. & Fotos, S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. Routledge.
Nunan, D. (1998). Teaching grammar in context. *ELT Journal, 52*(2), 101–109.
Nunan, D. (2004). *Task-based language teaching*. Cambridge University Press.
Pienemann, M. (1984). Psychological constraints on language teachability. *Studies in Second Language Acquisition, 6*(2), 182–214.
Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press.
Putz, M. (2007). Cognitive linguistics and applied linguistics. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of Cognitive linguistics* (pp. 1139–1159). Oxford University Press.
Radden, G., & Dirven, R. (2007). *Cognitive English grammar*. John Benjamins Publishing Company.
Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129–158.
Smith, M. S. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research, 72*, 118–132.

Sudharshana, N. P. (2017). Teaching grammar in a task-based framework: What Cognitive semantics can offer. *Humanising Language Teaching, 19* (6), Accessed on May 31, 2019 from http://www.hltmag.co.uk/dec17/mart06.htm

Sudharshana, N. P. (2018). Reading comprehension in ESL contexts: An applied cognitive semantics perspective. In R. J. Ponniah & S. Venkatesan (eds.), *The idea and practice of reading* (pp. 141–162). Springer Nature.

Talmy, L. (2000). *Toward a Cognitive semantics. Vol. 1, Concept structuring systems.* MIT Press.

Tyler, A. (2008). Cognitive linguistics and second language instruction. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive linguistics and second language acquisition* (pp. 456–489). Routledge.

Tyler, A., & Evans, V. (2004). Applying Cognitive linguistics to pedagogical grammar: The case of over. In M. Achard & S. Niemeier (Eds.), *Cognitive linguistics, second language acquisition, and foreign language teaching* (pp. 257–280). Mouton de Gruyter.

VanPatten, B. (2004). Input processing in second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–31). Lawrence Erlbaum.

Verity, D. P. (2018). 'Pedagogical Grammar'. In J. I. Liontas, T. International Association & M. DelliCarpini (Eds.), *The TESOL encyclopedia of english language teaching*. https://doi.org/10.1002/9781118784235.eelt0089

Williams, J., & Evans, J. (1998). What kind of focus and on which forms? In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 139–155). Cambridge University Press.

Willis, D., & Willis, J. (2007). *Doing task-based language teaching*. Oxford University Press.

Chapter 10
Designing and Using Tasks to Foster Metaphoric Competence Among Learners in Indian Contexts

Baburam Upadhaya and N. P. Sudharshana

Abstract Recently, it has been argued that metaphoric competence is an essential component of general communicative competence. However, the way metaphoric expressions are dealt with in Indian ESL context is far from satisfactory as we discovered from an analysis of popular textbooks at secondary level. Our review showed that activities in textbooks are based mostly on the traditional views which treat metaphoric expressions as largely arbitrary linguistic constructs. Such an approach encourages memory-based learning. The theory of conceptual metaphor (CM) in cognitive semantics, on the other hand, argues that metaphors are not embellishments used in literary works; rather, human thinking is fundamentally metaphorical in nature. We always perceive an abstract domain (e.g. time) in terms of a more concrete domain (e.g. money) as is evident in expressions such as 'spend time with family'. There have been some studies in applying principles of CM into language pedagogy. Using largely discrete exercises and metalinguistic explanation, such studies aim to raise learners' metaphorical awareness. In this paper, we report a study on designing and using tasks to facilitate the development of metaphoric competence among intermediate-level learners of English as a second language in Indian contexts. We first present details on designing tasks and then discuss insights from a small-scale intervention study using those tasks in an actual classroom.

Keywords Metaphoric competence · Indian ESL contexts · Conceptual metaphor · Task-based language teaching

Introduction

When people hear or see the term 'metaphor', they often think of it as a figure of speech. Popular high school grammar and composition books explain 'metaphor' along with 'simile'—the latter makes an explicit comparison between two entities (e.g. this piece of cloth is as white as snow), whereas the former is said to make an

B. Upadhaya (✉) · N. P. Sudharshana
Department of Humanities and Social Sciences, IIT Kanpur, Kanpur, India

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_10

implicit comparison without any comparative adverbs (e.g. like, as). For instance, the expression 'the camel is the ship of desert' is a metaphor in that the camel is compared with the ship implicitly. Very similar views are prevalent in ESL/EFL contexts. Metaphor is introduced primarily in context of literary works. The assumption that textbook writers and teachers generally have is that teaching and learning of metaphor are part of advanced proficiency. As a result, metaphor is introduced late and does not receive due attention. Though theories such as 'conceptual metaphor' argue that metaphor is ubiquitous because human thinking is metaphorical in nature (Lakoff & Johnson, 1980, 1999), metaphors have not received much attention from material developers and practising teachers.

However, of late metaphoric competence has been argued to be an essential component of communicative competence (Littlemore & Low, 2006b). This necessitates a more focused approach for teaching and learning of metaphor in ESL/EFL contexts. In this regard, the theory of conceptual metaphor is useful in that it provides the theoretical basis for analysis and interpretation of metaphor. However, discrete exercises and metalinguistic explanations may prove to be more complex for basic and intermediate-level learners. They are more likely to benefit if form focus is integrated with meaning focus in the classroom. This paper, therefore, presents a task-based framework to foster metaphoric competence among intermediate-level learners in Indian ESL contexts. The paper has two main parts: in the first part, we discuss the importance of developing metaphoric competence among learners and how metaphor was taught using the traditional and conceptual metaphor approach. In the second part, we show how form-focused tasks can be designed to foster metaphoric competence among learners and subsequently offer insights on using such tasks based on a small-scale intervention study with a group of intermediate-level learners.

Metaphors in Everyday Usage and Metaphoric Competence

Metaphor in Everyday Life

Metaphoric expressions are part of everyday language use in English (see Littlemore & Low, 2006a Chap. 1 for several examples). Metaphor is linguistically manifested as idioms (e.g. kick the bucket), phrasal verbs (e.g. the prices have gone up), verb + preposition combinations (e.g. the bill is on me), proverbs (e.g. A bird in the hand is worth two in the bush) and literary metaphors (e.g. 'All the world is a stage and all the men and women players'), among others. They are so pervasive in spoken and written English (Pollio, Barlow, Fine, & Pollio as cited in Cooper, 1999) that one could hardly think of avoiding these vocabulary items in their language use.

Metaphoric Competence as an Essential Part of Communicative Competence

Metaphors are found to be present in every aspect of language that 'learners need to use, understand or learn' (Littlemore & Low, 2006b, p. 269). For successful comprehension and production of metaphor, learners need to be able to understand one entity or concept in terms of another, which requires learners to infer the abstract underlying relationship(s) between the two entities or concepts. This ability to infer the meaning of metaphors effectively and to use them is termed as 'metaphoric competence'.

According to Littlemore and Low (2006b), metaphoric competence is not something that is 'recondite', but is very much part of communicative competence. Taking Bachman's (1990) general model of communicative language ability, they argue that metaphor plays an important role in all the areas of competence in the model, which includes sociolinguistic competence, illocutionary competence, textual competence, grammatical competence and strategic competence. According to them, metaphoric competence may help learners learn words and expressions that may not require active processing on the part of the native speakers. Citing Littlemore's (2004) work where a group of advanced language learners were engaged to understand the basic sense of the expression 'skirt around', they hold that spending class time on metaphoric thinking may facilitate L2 learning and use in general. Therefore, they stress the need for developing metaphoric competence of learners in second-language teaching and learning context.

Teaching and Learning of Metaphor in Classroom Contexts

Metaphoric expressions are regularly taught in the ESL/EFL classroom under the cover term 'idioms'. However, the amount of time devoted to them and the choice of instructional strategies to present these items in the classroom have been an issue. In most cases, teachers and course designers rely on the traditional approaches that hold the view that idioms are non-compositional since the meaning of these expressions cannot be deciphered from its constituent words. This perceived arbitrariness in the form and meaning has led to the conclusion that the only way to learn these expressions is to memorize them as separate linguistic units (Boers et al., 2007).

Traditional Approaches to Teaching and Learning of Metaphor

According to the traditional view, an idiom (a metaphoric expression in this case) is any multiword expression whose meaning cannot be determined by its constituent words. For instance, the meaning of the expression *kick the bucket* (which means 'to die') cannot be deciphered simply by combining the meaning of its constituent

words. This perceived arbitrariness in the form and meaning has led to the conclusion that metaphoric expressions are not a necessary part of everyday language use but are used mostly by proficient speakers to embellish their speech and writing.

The traditional view has been largely shaped by the structural and the generative theories of language, whereas the structural linguistics views language as structural arrangements based on abstract rules, the generative linguistics holds that human beings possess an innate language faculty, which allows them to construct multiple novel utterances with minimal number of structures. In this way, both the views accord importance to rules in language acquisition, albeit in different semblance. But since metaphoric expressions do not follow any such conventional language rules in the true sense of the term, they have been termed as 'type of formulaic language' (O'Dell & McCarthy, 2010, p. 6). For instance, the combination of words in the metaphoric expression *kick the bucket* is fixed, and it will lose its metaphoric meaning if it is changed into the passive form. It is because of such ambiguity, the traditional view holds metaphoric expressions as 'extra grammatical', and as a result, it made the mainstream linguists pay scanty attention to these lexical items.

Metaphoric expressions often present a difficulty for low proficiency language learners (Cooper, 1999). Despite this, the teaching of metaphoric expressions in ESL or EFL contexts has not received its due attention. There has been a collective belief among many second-language (L2) teachers that teaching metaphoric expressions is not as important as that of teaching grammatical rules and structures (Liontas, 2017). Traditional views (e.g. Aristotle's theory) considered metaphoric expressions as a list of exceptions whose purpose is merely to embellish the language (Gibbs, 1994), and such views could be responsible for lack of attention to metaphoric expressions in ESL classrooms and limited research in the language teaching materials (Hoang, 2014). Teachers in ESL contexts have mostly employed teaching strategies such as translation into learners' L1, guessing the meaning of expressions from context, working on the meaning of random expressions, using a dictionary and so on. So, the focus has been on itemwise learning of metaphoric expressions and with an expectation that learners would memorize such expressions and learn to use them (Boers et al., 2007).

It must be noted here that there have been some attempts to explain metaphoric expressions along pragmatic or functional lines. For example, McCarthy and O'Dell (2002), in their book *English Idioms in Use*, group expressions according to (i) functional lines, such as health, success and failure, praise and criticism; (ii) topic area, such as colour, games and sports, food and (iii) keywords, such as head, long, ground. By doing so, the authors are trying to teach metaphoric expressions through a systematic approach. But presenting idioms to learners by grouping them along pragmatic or functional lines does not provide any logical explanation as to the inner working mechanism of these expressions.

Teaching and Learning of Metaphoric Expressions in Indian Contexts

In Indian secondary schools, metaphoric expressions are generally found in the grammar section under the heading 'idioms and proverbs'. They are generally taught separately and not integrated with other classroom lessons. In an informal conversation with some English teachers, the first author of this paper found that metaphoric expressions hardly get more than two–three classes in an entire academic session. In these classes, the teacher introduces around 30–50 metaphoric expressions along with their meaning and example sentences. This is done either through oral dictation or by writing on the board. Learners are then asked to note them down and mechanically recall these expressions during exams. The way these expressions are taught gives learners the impression that these lexical items are meant only to embellish the language and are not so important to be integrated in everyday communication.

However, the teachers themselves are not to be blamed for this because the reference books they use for teaching purposes present these lexical items usually in the form of a long list, and it makes the teachers think that this is how these expressions are to be taught. One of the books popular among teachers of English and learners in India is Wren and Martin's *New Edition High School English Grammar and Composition* (2000). This book contains information on grammatical rules and composition, and there are exercises to practice those grammatical and compositional features in a traditional structural approach to language teaching. For idioms, there are two chapters dedicated. The first chapter is titled *Idioms*, and it contains metaphoric phrasal verbs grouped in alphabetical order. The phrasal verbs are introduced in discrete sentences with their meaning in parenthesis. The second chapter *Idioms (Continued)* contains quite a number of metaphoric expressions grouped randomly under numerical headings, 1, 2, 3 and so on. There are 39 such numerical headings, and each numerical heading consists of ten random metaphoric expressions, except the last one, which contains only seven. Surprisingly, there are no language activities in the book to practice these vocabulary items. So, teachers who use this book would be prompted to teach these expressions as they are presented in the book and assess them likewise.

However, this does not mean that there has not been any attempt to improve the learning of these vocabulary items. A look at the language activity section on metaphoric expressions of a couple of English textbooks published by National Council of Educational Research and Training (NCERT), an autonomous government organization for qualitative improvement in school education and which also prepares and publishes textbooks for primary and secondary schools, shows that there have been some attempts made to improve the learning of these lexical items. This attempt could be seen in the two language activities on metaphoric expressions in Chap. 4 of class VIII English textbook titled *Honeydew: Textbook in English for class VIII* (2014), whereas activity 1 is meant to develop learners' dictionary reference skills, and activity 2 is aimed at teaching metaphoric expressions by making the learners work in groups. Activity 1 contains some random idioms from the lesson and shows

learners the steps to look these expressions up in a dictionary. Activity 2 contains some discrete, jumbled sentences, and each of these sentences contains one metaphoric expression. Learners are first required to rearrange these words in a correct sequence, then underline the idioms and finally write their meaning by using a dictionary.

Another activity on metaphoric expressions in this textbook is found in the language activity section of Chap. 6. This activity is meant to teach idioms and phrasal verbs to learners by making them work in groups. In this activity, learners are given some words from the lesson, and they are required to order these words as would appear in a dictionary. Learners are then required to write down some idioms and phrasal verbs with these words by using a dictionary.

On a closer look, one can see that these activities in the NCERT book on metaphoric expressions are communicative in nature and are designed to facilitate interaction as the learners work through the items in pairs and groups. It must be acknowledged that the motive behind these activities is praiseworthy considering the opportunities they provide for learner interaction. However, the approach to deal with the teaching and learning of metaphoric expressions is not very different from the traditional approaches discussed earlier. An issue, for example, was that there was no rationale for selecting some metaphoric expressions and leaving out the rest. Another issue was that the activities did not provide any meaningful insights into the semantic content of the metaphoric expressions. In summary, what was found to be lacking in all these activities is an approach, which involves a systematic schema to understand the mechanism underlying the workings of metaphoric expressions.

Conceptual Metaphor Theory

Lakoff and Johnson (1980) in their now seminal work, 'Metaphors We Live By', argue that metaphors, rather than being an ornament to language, reflect the way how we think, talk and act. They emphasize that abstract human thought is mostly metaphoric and has a bodily basis (Lakoff & Johnson, 1999). In other words, metaphoric expressions are simply the linguistic manifestation of how we experience, perceive and understand the physical world around us. Our recurrent exposure to basic bodily experiences of the physical world helps in the gradual formation of a cognitive structure. This cognitive structure may be found either in the form of a bare image schema (e.g. UP-DOWN of a vertical axis, IN- - OUT of a container, NEAR- FAR of a distance) or in some specific metaphoric concepts (e.g. ANGER IS HEAT, TIME IS MONEY, ARGUMENT IS WAR).

These metaphoric concepts, which Lakoff and Johnson (1980) termed as 'conceptual metaphor', are defined as understanding one conceptual domain in terms of another. Generally, it is a less clearly demarcated domain that is understood in terms of a more clearly demarcated domain. The former is known as the 'target domain' (which is usually abstract) and the latter as the 'source domain' (which is relatively more concrete). For example, metaphoric expressions, such as *simmering with anger*, *boiling with anger*, *blow one's top*, *hit the ceiling* and *cool down* which at first seem

unrelated to one another are in fact motivated by the same conceptual metaphor ANGER IS A HEATED FLUID IN A CONTAINER.[1] Here, the abstract domain of anger is conceptualized by the imagery of a more concrete domain of a heated fluid in a container. Lakoff and Johnson (1999) argue that conceptual metaphors are grounded in our bodily experience about the physical world. For example, the conceptual metaphor ANGER IS A HEATED FLUID IN A CONTAINER is grounded in the experience of feeling hot by the human body when one gets angry. The experience of feeling angry is correlated with the bodily experience of feeling hot. This bodily experience serves as the basis for the conceptual metaphor ANGER IS A HEATED FLUID IN A CLOSED CONTAINER, where the body is the container for anger, the heated fluid is the emotion of anger, the pressure created by the temperature of the fluid is the intensity of anger, and the spilling over of the fluid due to pressure of the temperature is the outburst of anger. Thus, the resulting linguistic instantiations are a result of our conceptual knowledge about the corresponding elements between the source domain of heated fluid in a closed container and the target domain of anger (Kövecses, 2010).

Further studies in cognitive semantics are based on these principles. Kövecses (2001) claims that metaphoric expressions are 'conceptually motivated by the underlying conceptual metaphors and metonymies' (p.88). Metaphoric expressions, rather than being arbitrary linguistic phenomenon, are in fact conceptually structured (e.g. Kövecses & Szabo, 1996). According to Gibbs et al. (1997), 'idioms do not exist as separate semantic units within the lexicon, but actually reflect coherent systems of metaphorical concepts' (p. 142). The semantic motivation is a result of the conceptual mappings between the elements of one conceptual domain with that of the other (Kövecses, 2010).

Conceptual Metaphor in Classroom

How can one use the theory of conceptual metaphor in the classroom? Can it help foster 'metaphoric competence'? Cameron and Deignan (2006) argue that second/foreign-language learners, in order to acquire metaphoric competence, need three kinds of knowledge: (i) knowledge of linguistic expression which may vary slightly in various contexts; (ii) knowledge of conceptual meaning which forms the basis of the linguistic expression and (iii) knowledge of pragmatics that includes details of highly culture- and context-specific information. Out of these, the second type of knowledge pertains to conceptual basis and here the theory of CM assumes greater significance.

Second-language instruction that uses that the CM approach teaches metaphoric expressions by grouping them according to their conceptual metaphor meaning or by referring to their literal, original meaning. Hence, metaphoric expressions, which are

[1] We will be using small capitals to refer to conceptual metaphors in this paper, while individual instances of metaphor are in small letters and italicized (e.g. *simmering with anger*).

often considered arbitrary, become more transparent if they are explained in terms of conceptual metaphors. Findings from experimental studies suggest that if learners' metaphoric awareness is raised, it would facilitate comprehension and retention of these vocabulary items. This has been explored by a series of experimental studies by the likes of Kövecses and Szabo (1996), Boers (2000), Csábi (2004), Skoufaki (2005), Beréndi et al. (2008), Yasuda (2010) and many more. Some of these studies are briefly described below to show the efficacy of the cognitive semantic approach to teaching metaphoric expressions.

Kövecses and Szabo (1996) carried out an experiment with 20 unknown metaphoric phrasal verbs having adverbial particles *up* and *down* in them. Of the 20 phrasal verbs, ten were presented to Hungarian learners and the rest ten were not. The learners were divided into two groups. One group of learners was taught by writing the phrasal verbs on the blackboard along with their Hungarian L1 equivalents. The meanings of the phrasal verbs were explained to the learners who were then given 15 min to memorize. The same procedure was followed for the second group of learners with one exception. The phrasal verbs, written on the blackboard, were grouped according to conceptual metaphors, such as HAPPY IS UP (e.g. *feel up*, *cheer up*, etc.) and MORE IS UP (e.g. *speak up*, *turn up*, etc.). Apart from the ones presented to the learners of the first group, several other phrasal verbs that manifest the conceptual metaphors were also written on the board as linguistic examples. Fifteen min time was provided to the learners for memorization and explanation of the phrasal verbs. Once the intervention was over, a post-test was conducted to test the effectiveness of the two methods of instruction. The result showed that the group taught using conceptual metaphor performed better than the other in giving correct responses to the meaning of ten phrasal verbs that were introduced to them and also the other ten that were not introduced. This indicates that making learners' aware of conceptual metaphors underlying metaphoric phrasal verbs helps them perform better in recall tasks.

In another experiment, Boers (2000) presented 118 Dutch-speaking secondary school learners with a list of 18 expressions, such as 'Her comments *added fuel to the fire*', 'She *blew up* at me' and 'Don't *bite my head off*'. These expressions were put under conceptual metaphors as—ANGER IS FIRE, ANGER IS A HOT FLUID IN A CONTAINER and ANGRY PEOPLE ARE DANGEROUS ANIMALS. The same expressions were presented to a different group of learners but were put under functional heads which describe a slow build-up of anger, sudden anger and angry personalities. In both the cases, the vocabulary input had the same degree of lexical organization. The learners were given 10 min time to go over the metaphoric expressions and to ask for clarification if any. Following this, the learners were engaged in a 15-min guided classroom discussion about anger and conflicts. After that, the vocabulary items were removed and the participants were given an immediate post-test. The result of the post-test suggested that teaching metaphoric expressions along metaphorical themes facilitated retention as compared to teaching the same items along functional lines.

Likewise, Yasuda (2010) presented 21 phrasal verbs to Japanese learners having adverbial particles *up*, *down*, *into*, out and *off*. One group was told the meaning of each of the phrasal verbs by simply translating it in Japanese. The learners were

then given a checklist in which the phrasal verbs were arranged alphabetically with their Japanese translations. After that, they were asked to memorize the phrasal verbs. These phrasal verbs were explained to learners in terms of the orientation of the adverbial particles, such as COMPLETION IS UP (e.g. *dry up*, *use up*, etc.) and OFF IS DEPARTURE/SEPARATION (e.g. *get off*, *take off*, etc.). The emphasis of the instruction was on the manner of orientation of the adverbial particles that contributed to the overall meaning of the phrasal verb. The learners were asked to memorize the phrasal verbs grouped under different conceptual metaphors along with their Japanese translations. Learners were especially asked to pay attention to the conceptual metaphors. The entire procedure for both the groups lasted for 10 min before the participants were asked to take a post-test task. The post-test task consisted of filling up of 30 sentences with adverbial particles. However, the phrasal verbs of the first 15 sentences had been exposed to the learners of both the groups, and the phrasal verbs of the last 15 sentences had not been. The result of the post-test suggests that making learners aware of the orientation of the adverbial particles aids in the learning of novel idioms than mere memorization does.

As reviewed above, most of the studies on applying CM theory into classroom teaching–learning show that conceptual grouping of metaphoric expressions aids in comprehension and retention of these the vocabulary items. Learners are more likely to gain from instructions motivated by conceptual metaphors than from other forms of instructions.

A Task-Based Framework to Foster Metaphoric Competence

The primary purpose of all the studies discussed above was to make participants aware of the underlying conceptual metaphors in metaphoric expressions and to see whether such awareness is actually beneficial in comprehension, retention and production of the lexical items in the subsequent tests. In other words, the experiments were meant to study the effects of using the cognitive semantic approach in teaching metaphoric expressions. They were conducted in controlled settings guided by specific research objectives and research questions. However, neither the instructions nor the instructional materials used in the experimental treatments reflect normal classroom procedure. That means ESL/EFL teachers might not be able to use the same instructional procedure in their classroom. So, when it comes to applying this approach to actual instructional settings, teachers need to consider the classroom realities and as such adopt a language teaching and learning approach that is supported by current second-language acquisition research and theories. This is important because language acquisition is a whole process, and it involves multiple factors ranging from recognizing the sound system to understanding the context of use. It cannot be acquired in bits and pieces. We feel that TBLT in this regard can help bring together form focus and meaning focus. In the following subsections, we present a framework for bringing together CM and TBLT in classroom contexts. The framework would be discussed in the light of selecting the target expressions

for the tasks, designing and sequencing of the tasks and trialing the tasks in actual classroom.

Selection of Target Metaphoric Expressions

The first step in designing a task-based framework for fostering metaphoric competence is selecting the target expressions. While choosing target expressions, three main factors need to be considered: (i) age, cognitive ability and cultural background of learners, (ii) cross-linguistic similarities and differences and (iii) frequency and salience of metaphoric expressions. The first factor is important because if learners have not developed the cognitive ability to understand the source or the target domain, then teaching of these expressions would be meaningless. Research findings show that the cognitive ability to interpret complex metaphorical mappings is developed somewhere in between 9 and 12 years of age (Özçalişkan, 2007). In addition, it is also important to consider the cultural background of learners while designing the task because a task set in a different cultural context may not be relatable to learners, and as such, they may not be interested or able to perform it. The more the task is relatable to learners, the more they would be interested and able in performing it.

The second factor, which is the cross-linguistic similarities and differences between the source and the target language, plays an important role at the conceptualization level. This is because similarity between the two languages at the conceptual level helps learners in understanding the conceptualization process of the target idiomatic expressions and as such facilitates faster learning of these expressions. Similarity between the two languages at the conceptual level also saves the teacher's time and effort required to familiarize the learners with the conceptual dimension of the target metaphoric expressions.

The third factor which deals with the frequency and salience of target expressions is important because frequently used expressions or expressions which are salient in the target language have greater chances of becoming learners' part of language repertoire because chances of encountering such expressions in real-life communicative situation are quite high. This will further motivate the learners to learn because it would provide them with a sense of achievement as they have been able to use what they were taught.

After considering all the three factors, we have chosen metaphoric expressions from the domain of heat in this study to talk about anger. This was done for chiefly two reasons. First, this metaphor is 'central' in English (Kövecses, 2003). Second, the domain of heat is used in both Hindi (the L1 of target learners) and English to talk about anger, and learners are expected to be conceptually familiar with both the source and the target domains.

Following are the metaphoric expressions chosen for our study, grouped under different conceptual metaphors.

CM1: anger is heat

English—[hothead, burning with anger, hot temper, hot under the collar].

Hindi—[*garam dimaag*, *gusse se jalnaa*, *garam swabhaaw/mizaaj*].[2]

CM2: anger is fire

English—[spitting fire, flare up, adding fuel to fire, a fiery temper].

Hindi—[*aag ugalnaa*, *bhaD\ak jaanaa*, *aag mE ghi daalnaa*, *aag par tel chID\aknaa, ugra swabhaaw*].

CM3: anger is a heated fluid in a container

English—[simmering with anger, make one's blood boil, explode in anger, blow one's top].

Hindi—[*khuun khowlnaa/ubalnaa*, *gussaa phuuTnaa*].

Designing and Sequencing Tasks

This study draws on the use of form-focused tasks which, according to Ellis (2003), are tasks that aim to induce learners to use a particular linguistic feature in order to process it, either receptively or productively. This processing of linguistic feature must be in accordance with the key criteria of a task, which is to use the language pragmatically in order to achieve a non-linguistic outcome. Ellis (2003) lists two aims of form-focused tasks: (i) encourage communicative language use and (ii) target the use of a particular predetermined linguistic feature. So, accordingly, tasks can be designed in two main ways. The first way is to design tasks in such a manner that learner has no other option other than using the targeted feature. As designing such a task is not easy because learners can always find other strategies to communicate the intended message without using the targeted feature, the second way is to make the language itself the content of the task. Here, a language point is made the topic of the task. Learners are supplied language data, and they will try to work out a rule to describe how that particular targeted feature is used. This kind of task is called 'consciousness-raising (CR) task' (Ellis, 1997), and it, by all means, fulfils the criteria of a task (i.e. to use language pragmatically to achieve some non-linguistic outcomes) as it requires learners communicate in the target language by exchanging information and ideas while working out a rule.

Developing language tasks involves the actual writing of the tasks. While writing the tasks, the following points were taken into consideration in this study:

1. The primary focus of the tasks would be on meaning, although there would be provisions for attending to linguistic elements.

[2] Conventions for Hindi transliteration is attached in Appendix 2. This is adapted from Vijaya (2007).

2. Learners would be using their own linguistic resources to complete the tasks. However, learners would be induced to use the metaphoric expressions that have been provided to them in the form of input from various authentic sources.
3. In keeping with the cognitive development and language proficiency of the learners, the input material was selected and modified so that it is appropriate for their level (see Long, 2015 on input modification).

Willis (1996) advocates a three-phase framework for task-based teaching: Stage 1: pre-task (introduction to the topic and task, preparation); Stage 2: task cycle (task performance, planning and report) and Stage 3: language focus (language analysis, practice). We adapted this framework to suit our purposes. The final framework for this study is given below:

4. Pre-task: activating learners' existing knowledge on the topic, exposing them to the target metaphoric expressions and preparing them for the main task
5. Main task: task performance, planning and report
6. Post-task: consciousness/awareness raising (metaphor awareness and cross-linguistic awareness).

Tasks are sequenced along the three-phase TBLT framework—the pre-task, the main task and the post-task phase. We discuss a possible set of tasks in each phase below. See Appendix 1 for sample tasks and also Upadhaya and Sudharshana (2020). To begin with, learners' background knowledge about the target domain ANGER is first activated in the pre-task phase. It is done through questions the teacher asks to elicit responses from learners. It is a whole class discussion, and this task acts as a warm-up activity. Apart from this, a reading comprehension task is also provided. Here, instead of presenting the metaphoric expressions by grouping them according to their CM, these expressions are presented in the form of a reading task with general comprehension questions. Each metaphoric expression is presented in the context of a reading passage. Learners have to read the given passages and answer the comprehension questions. They are also instructed to underline the expressions related to the target domain ANGER. The reading task is interactive in nature and does not specifically focus on the target metaphoric expression. It is for the general understanding of the context in which the target metaphoric expression can occur. This task is basically meant for providing input of target metaphoric expressions to learners in an authentic context. The pre-task also includes a controlled production task to prepare learners for the main task. The controlled production task provides learners a passage which contains gaps that need to be filled with anger expressions they have underlined in the reading task. Students will do this task in pairs.

The main task involves a free production task. It is based on the controlled production task (except in the first lesson unit to show that it is not always necessary to do so). This task provides some hypothetical situations to students based on the context of the passage in the controlled production task. Students have to discuss with the friend sitting next to them and present their report before the class. In order to facilitate noticing of the target metaphoric expressions, students are asked to use any two expressions they have used in the controlled production task. Deciding which

expressions to use in the report is also a task in itself as students have to utilize their decision-making ability. This will in turn help them to decide which expressions are appropriate to use in the context of the report they are presenting and which expressions are not. This task allows students to use their mother tongue while discussing with their friends so that communication between them is not hindered due to the inhibition they experience while speaking in English with their friends. The idea is to create a relaxed and easy atmosphere in the classroom. Any form of imposition on students to use English may demotivate them and as such may backfire the learning process. However, learners will have to present the report in English. This will ensure that learners will have to think in English to a certain degree. The whole idea is to help learners develop a positive attitude towards English by not making English the only language in the classroom but by allowing them interact with each other in the language they find comfortable and work towards successful completion of tasks which has to be accomplished in English.

The post-task phase is meant for consciousness/awareness raising and consists of a metaphor awareness-raising task and a cross-linguistic awareness-raising task. The metaphor awareness-raising task requires learners to list the expressions related to anger and discuss with a friend how this emotion is talked about and what similarities do they see among these expressions. Based on these similarities, the researchers will try to elicit other expressions from the students related to that particular anger conceptual metaphor.

The cross-linguistic awareness-raising task requires learners to think about expressions in their mother tongue that are similar to the expressions they have listed in the metaphor awareness-raising task. Learners are required to discuss in pairs the similar expressions that are found in their mother tongue. If they do not find any such expressions in their MT, they are asked to think about expressions they use to talk similar things about anger. Later, they will present it before the class.

Trialing the Tasks in Actual Classroom

In this section, we present a brief report on trialing the tasks. This trialing was part of our large-scale study on using tasks for teaching and learning of metaphoric expressions (Upadhaya, work in progress). Since the tasks designed are intended for Grade VIII learners of the age group 12–14 years, learners are expected to be familiar with both the source and the target domain but not familiar with the way how the source domain of HEAT is used to talk about the target domain of ANGER. Moreover, it is highly likely that learners do not know many of the metaphoric expressions that have been derived from the source domain of HEAT to talk about the target domain of ANGER.

In order to check how the tasks would work with real students, a real classroom tryout was conducted by the first author. This was done to ensure that the implementation is consistent with the design, and the tasks would be able to meet the objectives of the proposed framework before being used for the main study.

The tryout was conducted on 27 Grade VIII Hindi-speaking learners of English in a real classroom in Kanpur, Uttar Pradesh. Their mean age was 13 years (12–14 years), and mean age of exposure to English was 8 years (7–9 years). Their English language proficiency was between early intermediate to intermediate. As per our estimation and going by their teachers' account, learners would be able to read and understand English texts on topics of general interest (to most extent) which do not require any specialized knowledge on the part of the learner. Talking about their socio-economic conditions, learners belonged to the broad category of 'middle class'.

In the first class, students had certain difficulty in following the instructions because they were not familiar with this type of instruction earlier. This was their first experience with TBLT. Never before had they been allowed to take charge of their own learning. Nevertheless, learners adjusted very soon and started performing the tasks by following the instructions correctly. Learners seemed to be more than excited to report their task outcome. However, silent learners needed more time and cues. The author's intervention was not that much required when learners were performing the reading task, the controlled production task and the free production task. The only thing the author had to do was to go around the class to monitor learners' progress and as such assist them wherever required for successful task completion.

However, in the metaphor awareness-raising task and cross-linguistic awareness-raising task, the author had to intervene quite often when learners were performing these two tasks. None of the learners could associate the anger expressions with any of the source domain on their own. The author had to explain how the metaphoric expressions are motivated by the various source domains. This confirms the findings that advocate the explicit instruction of CMs in language learning contexts. Moreover, learners could not come up with similar expressions in their mother tongue for the conceptual metaphor ANGER IS HEAT. This may be because the metaphoric expressions of this particular CM in Hindi [e.g. *garam dimaag* (hothead), *garam swabhaaw/mizaaj* (hot-tempered)] may not be widely used. What is to be noted here is that learners did not seem to be so sure about the presence of such expressions in their L1 even after they were given some example sentences. It could also be because of inhibition which they were not able to overcome completely.

However, in the case of the conceptual metaphor ANGER IS FIRE and ANGER IS A HEATED FLUID IN A CONTAINER, learners could say the similar expressions in Hindi. Many students could easily tell the Hindi equivalent expressions of the metaphoric phrases *spitting fire* (*aag ugalnaa*) and *adding fuel to the fire* (*aag mE ghi daalnaa, aag par tel chID\aknaa*). One student wrote *jale par namak chID\aknaa.* But *jale par namak chID\aknaa* has a different essence which has the following equivalent interpretation in English—*rub salt into the wound*. However, they could not think of similar expressions in Hindi for *flare up*. But they could tell the general meaning of it by guessing the meaning from the context. Even though learners understood the general meaning of the expression *flare up* with the help of the context, they could not tell its precise meaning. This may be because they did not know the literal meaning of the word 'flare' in *flare up*. This indicates the importance of making learners aware of the literal meaning of the word in the metaphoric expression. It also indicates that

guessing meaning from the context is not an effective method to teach vocabulary in general and metaphoric expressions in particular.

For ANGER IS A HEATED FLUID IN A CONTAINER, learners were able to find some connection between the expressions. They tried to convey the association by using the expressions 'heat', 'boil' and 'boiling water'. This may be because of the expression *make someone's blood boil* and not because of the imagery of 'a heated fluid in a container' as such. Moreover, learners were able to tell the meaning of these expressions not because of the conceptual metaphor but because *make someone's blood boil* has an identical expression in Hindi (*khuun khowlnaa/ubalnaa*). Again, learners were able to understand the meaning of the anger expressions *explode in anger* and *blow one's top* because of the words 'explode' and 'blow' in them as learners have a similar expression *gussaa phuuTnaa* in Hindi, where *gussaa* means 'anger' and *phuuTnaa* means 'explode' or 'blow off'. This shows how cross-linguistic similarities help in the understanding of an expression.

The reporting of the trialing of the tasks above shows continuous engagement and involvement of the learners with the form, meaning and context of the target expressions. This engagement and involvement happened at various depths and levels of processing, and it is believed that it would help learners for better comprehension and retention of the target expressions, thereby facilitating their production. In the meantime, the claims made above should not be treated as some kind of hard empirical data about the effectiveness of conceptual metaphor-based tasks but a general observation about how these tasks could provide not only meaningful learning opportunities to learners but also insights into the learning process involved. So, only a proper intervention with a pretest and a post-test can tell us the efficacy of these tasks in teaching metaphoric expressions.

Conclusions

Given the importance of teaching metaphoric expressions to L2 learners of English, the paper briefly examined how metaphoric expressions are taught in Indian secondary schools and proposed a methodology that combines conceptual metaphor and task-based language teaching. While conceptual metaphor is used to explain the inner working mechanisms of metaphoric expressions, TBLT provides a communication-oriented framework for teachers to give learners a meaningful language experience to acquire the metaphoric expressions. The study argues that teaching metaphoric expressions by integrating CM and TBLT, especially by using form-focused tasks, would help learners gain this vocabulary knowledge in an effective way in both breadth and depth. It must be noted that engaging learners in such tasks does not guarantee that these expressions would start appearing in their language use immediately. This is because production usually takes longer and the first effect is usually seen on comprehension. To make learners use the expressions, they must first be provided with opportunities to do so.

This was a preliminary attempt at developing tasks. Future studies need to consider several factors to make the teaching and learning more effective. For instance, tasks need to cater to learning style and preferences of learners in a heterogeneous class. Also, teacher training plays a significant role since many teachers may not be aware of CM theory and also how to design form- and meaning-focused tasks.

Appendix 1

Sample task

Unit 1

I. **Answer the following questions**

1. Imagine you are studying hard for exams and your brother/sister turns on the TV at a high volume. How do you feel?
2. List a few such situations which generally make you get angry.
3. When you get angry, do you shout at people or break things around you?
4. Have you observed how others behave when they get angry?

II. **Read the following paragraphs and answer the questions that follow. UNDERLINE the expressions related to ANGER. One has been done for you. For difficult words, look up a dictionary or ask the teacher**

1. Jitu is a brilliant football player. He has both skill and pace. Last year, he had scored 12 goals in the interschool tournament. But, this time, he has not been selected. This really made him **hot under the collar**, as he could not believe that the coaches did not select him. Some of his friends saw him shouting at the coaches. He thinks that the coaches have taken money from the other players who had been selected in his place, and that is why, they left him out.
   a. How did Jitu react when he knew that he had not been selected for the tournament? Was he happy for not getting selected?
2. As a kid, I remember I told my mother a lie. I was probably 5 or 6, and I had told her that our class teacher slapped me for not doing the homework. Upon hearing this, she started burning with anger and decided to meet the teacher the next day. Seeing her burn with anger, I could not keep up the lie for a long time and told her that I was lying. In fact, no homework had been given to me by the teacher. She then calmly called me near to her and warned me not to tell such lies again.
   a. How did the mother react when she heard the incident of slapping? Why could not the writer keep the lie for a long time?

3. Satyanath is the only child of his parents. He has a hot temper. Whenever he is not given what he asks for, he would immediately get angry and start screaming at his parents and banging things around with his fists. The parents have asked him many times not to behave in that manner, but he does not listen to them. This behaviour of Satyanath really upsets his parents, and they really do not know what to do to control his anger.
   a. What kind of person is Satyanath? Why are his parents upset with him?
4. Raman is a great dancer. He has won several awards by participating in dance competitions. But he has a bad reputation of being a hothead because he gets angry quickly and easily at little things. In the last competition, when his dancing partner missed some steps, he started yelling and shouting at her. The judges did not like his rude behaviour and asked him to leave the competition immediately.
   a. What is Raman infamous for? What does being a 'hothead' mean?

III. **Work in pairs**

Aman and Mohan are brothers, and they have a class test tomorrow. Aman wants to study but Mohan is jumping around. Even after telling him several times not to disturb him, Mohan does not listen to Aman. Look at the following conversation between them and fill in the blanks with the expressions that you have underlined in task II.

Aman: (in a sharp voice) See, I have an exam tomorrow, and you are not letting me study.

Mahan: I am not doing anything to disturb you.

Aman: But you are jumping around.

Mohan: So what? You just keep studying.

Aman: (in a raised voice) How can I study when you are jumping around the house and making noise?

Mohan: Why are you getting _____?

Aman: I am not. Unlike you, I am not a _____.

Mohan: You do not have to argue with me and start _____.

Aman: Sorry man. I do not have a _____. And I cannot argue with you. Can you please move away now?

Mohan: Calm down, my brother. You study. I am going to bed now. Good night.

IV. **Work in pairs**

Imagine a situation where one of your friends got a change of a hundred rupee note from a shopkeeper. The shopkeeper gave him 10 ten-rupee notes. On reaching home, his mother counted it and found ten rupees less. She then told him to go back and get that Rs. 10 from the shopkeeper. When he went back and told him about the whole thing, the shopkeeper got hot under the collar and said, 'Don't you try to trick me little boy. I remember very well that I gave you 10 ten-rupee notes. Now go away from here'. Upon hearing this, your friend started burning with anger.

i. Being his friend, what would you do to calm down the situation and help him in this problem?

a. Discuss it with the friend sitting next to you for 5 min. One of you from each pair will report it before the class.
b. Try to use at least two expressions you have used in task III.

Note: You can discuss it in your mother tongue but you will have to report in English before the class.

V **Work in pairs**

a. Go back to task II and list all the expressions that are related to anger.
b. Discuss and report how anger is being talked about here. Can you see any connection in all these expressions? What is that connection?

VI. **Work in pairs**

For the teacher
When *heat* is applied to a substance, the temperature of that substance increases. And when the temperature of that substance reaches the ignition point, it starts to burn. Apart from the actual use of heat, we also use heat to refer to anger in English, which you must have seen in the anger expressions above.

In task V, are there any similar expressions in your mother tongue? If yes, then list them alongside the English expressions. If no, what expressions do you use to talk similar things about anger? Discuss and report.

Appendix 2

Hindi Phonet/mic Chart

VOWELS

NON-NASAL	NASAL
a (nagar ‘town’)	A (Ank ‘marks’)
aa (aap ‘you’)	AA (Aakh ‘eye’)
I (idhar ‘here’)	I (Idu ‘name’)
i (niilaa ‘blue’)	II (IIdhan ‘fuel’)
u (udhar ‘there’)	U (Uglii ‘finger’)
uu (uupar ‘up’)	UU (UUT ‘camel’)
e (meraa ‘mine’)	E (mE ‘in’)
ae (paesaa ‘money’)	AE (mAE ‘I’)
o (or ‘direction’)	o (hoTh ‘lip’)
ou (ourat ‘woman’)	OU (sOUf ‘fennel seed’)
ow (kowaa ‘crow’)	OW (don’t know a word)
|e (h|e ‘is’)	E (h|E ‘are’)
ei (sound of “a” in English ‘wait’)	EI (nasalized counterpart of “ei”)

CONSONANTS

bilabial	p	b	ph	bh	m	
dental	t	d	th	dh	n	
palatal	c	j	ch	jh	ny	
retroflex	T	D	Th	Dh	N|	
velar	k	g	kh	gh	N	
LIQUIDS		y,	r,	l,	w	
FRICATIVES		s,	sh,	h,	f,	z
FLAPS		D| (peD| 'tree')			D|h (paD|h 'read')	

References

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Beréndi, M., Csábi, S., & Kövecses, Z. (2008). Using conceptual metaphors and metonymies in vocabulary teaching. In F. Boers & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 65–99). Mouton de Gruyter.

Boers, F. (2000). Metaphor awareness and vocabulary retention. *Applied Linguistics, 21*(4), 553–571.

Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research, 11*(1), 43–62.

Cameron, L., & Deignan, A. (2006). The emergence of metaphor in discourse. *Applied Linguistics, 27*(4), 671–690.

Cooper, T. C. (1999). Processing of idioms by L2 learners of English. *TESOL Quarterly, 33*(2), 233–262.

Csábi, S. (2004). A cognitive linguistic view of polysemy in English and its implications for teaching. In M. Achard & S. Niemer (Eds.), *Cognitive linguistics, second language acquisition, and foreign language teaching* (pp. 233–256). Mouton de Gruyter.

Ellis, R. (1997). *SLA research and language teaching*. Oxford University Press.

Ellis, R. (1999). *Learning a second language through interaction*. John Benjamins.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding*. Cambridge University Press.

Gibbs, R. W., Bogdanovich, J. M., Sykes, J. R., & Barr, D. J. (1997). Metaphor in idiom comprehension. *Journal of Memory and Language, 37*(2), 141–154.

Hoang, H. (2014). Metaphor and second language learning: The state of the field. *The Electronic Journal for English as a Second Language, 18*(2), 1–27.

Kövecses, Z. (2001). A cognitive linguistic view of learning idioms in an FLT context. In M. Putz, S. Niemeier, & R. Dirven (Eds.), *Applied cognitive linguistics II: Language pedagogy* (pp. 87–115). Mouton de Gruyter.

Kövecses, Z. (2003). *Metaphor and emotion: Language, culture, and body in human feeling*. Cambridge University Press.

Kövecses, Z. (2010). *Metaphor: A practical introduction* (2nd ed.). Oxford University Press.

Kövecses, Z., & Szabó, P. (1996). Idioms: A view from cognitive semantics. *Applied Linguistics, 17*(3), 326–355.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. The University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh*. Basic Books.

Littlemore, J., & Low, G. (2006a). *Figurative thinking and foreign language learning*. Palgrave MacMillan.

Littlemore, J., & Low, G. (2006b). Metaphoric competence, second language learning, and communicative language ability. *Applied Linguistics, 27*(2), 268–294.

Liontas, J. (2017). Why teach idioms? A Challenge to the Profession. *IJLTR, 5*(3), 5–25.

Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley Blackwell.

McCarthy, M., & O'Dell, F. (2002). *English idioms in use: Intermediate*. Cambridge University Press.

NCERT. (2014). *Honeydew: Textbook in English for class VIII*. National Council of Educational Research and Training.

O'Dell, F., & McCathy, M. (2010). *English idioms in use (Advanced)*. Cambridge University Press.

Özçalişkan, Ş. (2007). Metaphors we move by: Children's developing understanding of metaphorical motion in typologically distinct languages. *Metaphor & Symbol, 22* (2), 147–168. https://doi.org/10.1080/10926480701235429

Skoufaki, S. (2005). Use of conceptual metaphors: A strategy for the guessing of an idiom's meaning? In M. Mattheoudakis & A. Psaltou-Joycey (Eds.), *Selected papers on theoretical and applied linguistics from the 16th international symposium (April 11–13, 2003)* (pp. 542–556). Aristotle University of Thessaloniki.

Upadhaya, B. (work in progress). *Fostering metaphoric competence of ESL learners: An integrated approach using conceptual metaphor and task-based language teaching* (Unpublished doctoral dissertation). Indian Institute of Technology Kanpur, India.

Upadhaya, B., & Sudharshana, N. P. (2020). *Developing conceptual metaphor–based tasks in the teaching and learning of idiomatic expressions*. Advance Online Publication. https://doi.org/10.1002/tesj.560

Vijaya. (2007). *The 'noun advantage' in English as a second language: A study of the natural partitions hypothesis* (Unpublished doctoral thesis). Central Institute of English and Foreign Languages, Hyderabad, India.

Willis, J. (1996). *A framework for task-based learning*. Longman.

Wren, P. C., & Martin, H. (2000). *New edition high school English grammar and composition* (Rev. by N. D. V. P. Rao). S. Chand.

Yasuda, S. (2010). Learning phrasal verbs through conceptual metaphors: A case of Japanese EFL learners. *TESOL Quarterly, 44*(2), 250–273.

Part II
Task-Based Language Assessment

Chapter 11
Task-Based Language Assessment for Large-Scale and Classroom-Based Oral and Print Assessments

Lina Mukhopadhyay and N. P. Sudharshana

Abstract This chapter serves as an introduction to Part 2 of this book where a set of six chapters based on applications of task-based language assessment (TBLA) are included. Task-based language assessment emerged in the 1990s as an application of the task-based language approach where learners trained through this methodology would need to be assessed using a similar test/assessment construct. Thus, the use of the task-based construct emerged as an alternative mode of language assessment with a focus on authentic real-time tasks from which inferences could be drawn about test-takers' ability to use the target language to complete real-life tasks in a meaningful manner. At the turn of the twenty-first century, TBLA has seen a significant rise in its application for assessment for various purposes such as tasks as standards, tasks in proficiency assessment, tasks for employment certification and tasks for language educational assessment (Norris in Ann Rev Appl Linguist 36:230, 2016). TBLA has necessitated a linked use of task characteristics specification to build on difficulty or complexity, rating scales and the inferences assessors have drawn from such performances about test-takers ability to perform under real-life task conditions and their underlying general linguistic abilities. The chapters on TBLA in this volume present the large-scale use of task as a construct in oral proficiency assessment as well as tasks for language educational assessment, and more specifically tasks for classroom-based assessment in ESL oral and print proficiency.

Keywords TBLA · TLU domain · Content domain · Task difficulty · Triadic componential framework

L. Mukhopadhyay (✉)
Department of Training and Development, The EFL University, Hyderabad, India
e-mail: linamukhopadhyay@efluniversity.ac.in

N. P. Sudharshana
Department of Humanities and Social Sciences, IIT Kanpur, Kanpur, India

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_11

Task-Based Language Assessment (TBLA): Inception and Progress

Task, as an offshoot of second-language education as conceptualized by Prabhu (1987, 2019) and taken forward by other researchers to the development of syllabuses and its possible benefits for language learning outcome, has been well studied in the last five decades. Seminal works of researchers like Prabhu (1987), Skehan (1998), Robinson (2011), Ellis (2003) among many others have been to examine the impact of increasing task demands on promoting learner ability to use L2 in various target-like situations. Building on the argument that increasingly higher demands made on SL learners to perform in real-life situations by manipulating task characteristics, such battery of tasks have also been extended for learner assessments when they exit programmes (Baker, 1990). Thus, the task as a test construct is equally applicable to the context of assessment and therefore has given rise to the alternative branch of language assessment called Task-Based Language Assessment (TBLA) in the late 1980s. It has served as a complementary tool of TBLT, which focuses more on pedagogical context-based learning. TBLA refers to 'an assessment that utilizes *holistic tasks* involving either real-world behaviour or the kinds of language processing found in real-world activities' (Ellis, 2003: p. 285).

Although TBLA has its primary focus on the real-time performance of learners, the concept of Second-Language (SL) performance has had its inception in the 1970s with the perspective of direct testing of SL performance gaining popularity due to the work of Bachman (1990) and Macnamara (1996) along with other applied linguist-based research (Kramsch, 1986; Tarone, 1998). The addition of task-based construct to language assessment is a more recent one gaining impetus from the work of task-based language pedagogy by Long (1985), Candlin (1987), Skehan (1998) and making a foray into assessment through research by Robinson and Ross (1996) Bachman (2002), and more recently by Norris (2016) and Nakatshuhara et al. (2017, 2020).

Principles of TBLA

TBLA is a researched-based assessment model that can be used for both large-scale as well as classroom-based purposes. As a test construct, it prioritizes meaning-making based on task (or a battery of tasks with increasing demands) as the key construct and through the task performance development of interlanguage of learners and prepare them to perform in real-life situations. At its inception, TBLA had two prominent viewpoints about the inferences that can be drawn from performances: the first one is by Brindley (1998), Skehan (1998), and Macnamara (1996) that task-based performances would be able to estimate generalizable language abilities of learners. A more recent perspective by Brown et al. (2002) and Norris et al. (1998) is that task-based language performance would predict learners' ability to perform in

similar and specific task-based situations. Both the viewpoints remain valid in their interpretations derived from the TBLA-based performances. However, Bachman in 2002 has shared his reflections on the principles of TBLA and that the inferences would be dependent on the authenticity of tasks and what test-takers can do with the task features and an interaction between code complexity, cognitive complexity and communicative stress (Skehan, 1998: p. 88). So the predictions about learners' future use of the target language would be based on a set of generalizable task difficulty features arising out of the interaction between these three elements and the relatedness of the assessment tasks to the TLU domain of learners along with the content specificity of the real-life tasks. Bachman cautions that the predictions can be narrow in scope because of the content of the task construct and that the predictive validity can be confounded by what is defined as a task. So a way out of this theoretical problem, according to Bachman, would be to define a task-based and related language ability based construct, create task characteristics and identify possible areas of language abilities that can be assessed through a mix of analytic approaches both quantitative and qualitative to show the interactions between task types and characteristics and test-takers and their correlation to SLL (2002: pp. 468–471).

Norris et al. (1998, 2016) forward the viewpoint that tasks offer advantages and comprehensive frames for language assessment in alignment to teaching through the TBLT method. While discussing the interpretations, intended uses and designs in TBLA, Norris asserts that TBLA builds upon the foundations of communicative language teaching by taking into consideration the purpose of assessment, objectives of the assessment, task selection or design, and assessment criteria for the tasks. He maintains that the type and processes of inferences made from the assessments help educators for further considerations about learner performances, necessary feedback for learners and washback for teachers. He explains that in contrast to product-oriented discrete-point testing of knowledge about language, TBLA provides a choice to assess 'communicative goals with valued tasks' for both summative and formative assessment purposes in language education (pp. 239–241).

Norris presents four directions in which the application of TBLA has been practised to forward the principles of this model of assessment: tasks as standards, tasks in proficiency assessment, tasks for employment certification and tasks for language educational assessment (2016: 232–240). According to him, various stakeholders have been involved in bringing out the critical assessment demands that task as a test construct presents, namely in four directions: one, policy demands on what language ability can be for specific groups (e.g. nursing, teachers in TESOL programmes), two, interpretations of language ability of test-takers based on large-scale proficiency assessment (e.g. IELTS), three, use of job-related task-based assessment to predict employee's readiness to perform in the workplace using the target language, four, in raising teacher-learner awareness about language learning processes, outcomes and potential areas of feedback to support learning. Norris emphasizes that TBLA provides scope for formative classroom assessment, an important tool for language education and if introduced as innovations can impact curriculum and generate useful

data based on which public understanding of language ability and outcomes could be formed.

Researching TBLA

Research on TBLA in both, first- and second-language learning contexts, has been to examine teacher–learner awareness about language learning processes which can improve learning outcomes, and provide opportunities for constructive feedback (Norris, 2016; Shehadeh, 2018). In the past twenty years, task complexity factors have provided researchers with a rationale to apply Robinson's Triadic Componential Model, and they have assessed the effects of task demands on speech production (Robinson, 2011; Iwashita et al., 2001; Gilabert, 2007; Ishiwaka, 2008; Tamboli, 2017) or written production (Bygate & Samuda, 2005; Ortega, 2003) employing general measures of accuracy, fluency and complexity. The application of TBLA to assess English for academic purposes, especially reading, has been studied by Robinson and Ross (1996) to show that authentic task conditions that require the use of academic reading ability when used in combination with traditional system referenced tests become a clear predictor of success and a valid measurement tool of academic reading for selection purposes into graduation courses.

A key characteristic of formative assessment is that it foregrounds the use of assessment *for* learning rather than an assessment *of* learning (as in summative assessment), and TBLA precisely aims at assessment *for* learning within a framework of formative assessment. As Norris (2016) justifies, TBLA requires a structured process in designing and selecting relevant tasks to attend to language abilities and task performance abilities. In this collection, we propose that this structured process can be further aligned with the 'task complexity' feature of Robinson's cognition hypothesis (2001,2011). To support the hypothesis in a practical way, Robinson has provided language teachers and researchers with the Triadic Componential Framework (2001,2011), which distinguishes between *task complexity, task conditions* and *task difficulty*. Thus, these features can account for 'individual differences' in learner performances, which is valuable information for a teacher. In this respect, knowledge of the Cognition Hypothesis and using its guidelines would help the teachers in the selection of tasks, design of classroom assessments and student evaluation following the principles of TBLA.

Robinson introduces the possibility and requirement of reflecting upon the complexity factors in language assessment tasks and using it with various levels of cognitive demands to assess different levels of language proficiency (Robinson, 2011: p. 5). However, this argument does not further explain how complexity factors could be incorporated in second-language assessment. While Robinson asserts that there is a debate on whether linguistic complexity should be thought of as a continuum on a scale of complexity, he maintains that 'a text' aligned with these features does constitute a dimension of complexity and can be manipulated in an assessment task

to understand the underlying cognitive functions that are activated during the performance on a specific task. In an academic setup, learners at threshold level of proficiency (B1 level of CEFR: p. 25) are expected to exhibit certain age-appropriate cognitive parameters of learning. Learners' expected language proficiency and functional ability are defined for every level by the Common European Framework of Reference (CEFR, 2001: pp. 20–42). Hence, linguistically complex tasks for a particular level can pose an expected level of cognitive load on learners that can engage them in putting effort to accomplish the task goals and thereby experience a language growth in specific dimensions on complexity, accuracy and fluency.

The Part 2 of this book charts out specific features of oral and print tasks that can be selected based on the TCF dimensions and the same framework be furthermore use to design evaluation criteria.

Overview of Chapters in Part 2

The chapters in this section of the volume are presented with twin objectives: one, to show recent research and deliberations on the construct of TBLA; and two, to draw implications of this model for classroom-based assessment for practising teachers and young researchers. Thus, most of the chapters discuss how to apply the assessment ideas for classroom purposes by attending to the principles of authenticity, validity and reliability and provide a genuine learning opportunity and to give feedback to develop L2 performance. Like the chapters of Part 1, we hope practising teachers find the chapters in Part 2 useful and applicable. Each of the seven chapters in this section is summarized below.

Chapter 12 'Validation of a large-scale task-based test: Functional progression in dialogic speaking performance' by Chihiro Inoue and Fumiyo Nakatsuhara reports on a small-scale study on the revision of the Integrated Skills in English (ISE), an exam conducted at four levels (0, I, II and III) by Trinity College, London. The study had two aims: (i) validate the function lists in the specifications for the speaking tasks for ISE at all four levels and (ii) identify features of language elicited across the four levels and four score bands (Band A, B, C and D), leading to conversion of holistic rating scales into analytic ones. The authors examined 32 audio recordings of performances on the ISE spoken examination, across four ISE levels and four bands. The recordings were transcribed and subjected to both quantitative and qualitative analyses. The results reveal that most of the listed functions in ISE are captured in the actual performance of the test-takers validating the functions list in the ISE exam. However, some functions that were expected were not elicited (e.g. *expressing ability and inability expressing reservations*). Also, certain language functions were not expected, yet they were observed in the performance (e.g. *negotiating for meaning, asking for opinions*). The authors argue that the ISE specifications need to be modified accordingly. When some functions were observed across levels, learners at a higher level used the same language functions but with greater precision and more sophistication. The results

have implications for empirically validating function lists and for developing rating scales for a variety of tasks at varied proficiency levels.

Chapter 13 'Communicative Strategies as a Tool for Assessing Spoken Interactional Competence' by Pankaj Narke is situated in the classroom-based assessment of oral production of ESL learners at the tertiary level. The study proposes the use of communicative strategies (CS) as a tool to assess learners' interactional competence. The study used group discussions as interactional tasks with the contemporary argumentative topics in six rounds. The oral data of tertiary level ESL learners was audio-recorded and analysed qualitatively and quantitatively to identify the trends in the use of CSs across the six rounds of group discussions. The data analysis [following Dornyei and Scott's (1997) taxonomy] revealed that learners used eighteen CSs with varying frequency to solve basic and advanced communication problems during their performance. Learners showed progress in the frequency of using advanced CSs because of task familiarity and peer collaboration. The replacement of basic CSs with advanced CSs and change in the purpose of using CSs was also observed which led to the improvement in learners' speech performance. The study thus establishes CSs as a valid construct to assess interactional competence. The chapter concludes with a section on implications for assessment in ESL contexts.

Chapter 14 'Effects of Task-Structure and Interaction Conditions of Oral Performance' by Sajit Mathews and N. P. Sudharshana is a study in the oral assessment set in Skehan's framework focusing on tertiary level learners. The chapter the influence of task structure on task performance in terms of fluency, accuracy and complexity. Three kinds of tasks, Personal Information Exchange (PIE), Narration (NAR) and Decision-making (DM), were administered in two interlocutor settings, peer–examiner and peer–peer in two planning time conditions (no planning and three-minute planning) to a total of 56 tertiary-level Indian ESL learners. The researchers hypothesized that with a complex task, speech complexity of learners would increase; however, their accuracy and fluency will decrease. Audio recordings of the task performances were transcribed and coded. The study found that the complexity of oral performance has a strong positive correlation with task complexity. On the contrary, the accuracy of oral performance showed a negative correlation with task complexity. No consistency in the patterns of fluency of oral performance was found. The study offers multiple implications for ESL teachers and test designers which include the relevance of task structure and task output in terms of complexity, accuracy and fluency. The trade-off between complexity and accuracy can help the teachers to cater to the needs of the learners from different proficiency levels.

Chapter 15 'Using Task-Based Speaking Assessment to measure lexical and syntactic knowledge: Implications for ESL learning' by Vasim Tamboli is a study set in Robinson's Triadic Componential Framework for TBLA focusing on young learners. Tamboli begins by underlining the gaps in the oral assessment system in the Indian ESL contexts, particularly about the arbitrariness of topics and structures. The author advocates the Task-Based Speaking Assessment (TBSA) model which focuses on content, assessment objectives, learners' needs, preparation and authentic assessment situation. In order to investigate whether the task complexity

features affect language use in a series of task performance, two kinds of tasks, monologic (picture description simple and complex) and dialogic (information sharing and decision-making), were used. The subjects comprised 12 ESL learners (aged 13–14 years). Audio recordings of task performances were transcribed and analysed based on two micro measures—lexical variety and syntactic knowledge. The results indicated that task complexity feature affected lexical variety in learners' performance by exhibiting steady growth in the use of transitive and intransitive verbs. Further, learners' syntactic knowledge growth was captured in terms of the variety of verb phrases occurrence in their performance. The empirical findings presented in the study validate the effectiveness of TBSA to assess language growth among ESL learners. The chapter concludes with a set of detailed instructions for designing tasks and criteria for assessment of speaking.

Chapter 16 'Whole Text Reading Comprehension: An Application of Task-Based Language Assessment' by Vrishali K. Patil is an application of Robinson's triadic componential framework for the assessment of reading, primarily for classroom purposes. The chapter argues that the task-based construct of reading assessment is likely to help teachers identify the levels of complexity of texts, select texts for assessment purposes and prepare the criteria of assessment to measure understanding of the content of whole texts. The target group is learners which have a minimum of 6 to 7 years of exposure to academic English proficiency (approximately B1/B2 levels of CEFR). In this chapter, the author argues for a broader perspective for the assessment of reading skill by advocating the whole text comprehension approach. This can be achieved by employing free summary recalls wherein learners are required to understand the literal meaning, suggested meaning and mental representation of the text. The chapter offers a systematic process for designing a classroom-based assessment of whole text comprehension in multiple stages that can broadly be divided into four categories, selecting text type, analysing the internal rhetoric structure of the text, parsing the ideas to create summary propositions and validating the tool and applying it. The study concludes that unlike assessing individual sub-skills of reading comprehension, whole text comprehension assessment would provide a wider understanding of learners' text comprehension.

Chapter 17 'Tasks in language acquisition research: more than what meets the blinking eye' by Shruti Sircar and Lina Mukhopadhyay reviews research tasks that can be used for pedagogical and classroom assessment purposes for studying children's language acquisition with specific attention to task features like task design, task administration and methods of scoring. The chapter begins with a comprehensive definition of an ideal test—the one that does not allow methodological aspects to affect test-takers' performance. The authors subsequently examine six types of tasks, namely grammaticality judgement, truth value judgement, elicited imitation, Dictation, picture description and narrative retelling task. In methodological consideration, authors comment on the suitability of these six tasks in particular contexts, the kind of data to be elicited and the ideal method of task administering.

Overall, the seven chapters present in this section of the book offer different examples of TBLA. We hope the information provided, the perspectives shared and the evidence shown will inspire young researchers and ESL/EFL teachers to take

up further analyses of the use of a similar model in their local contexts to find out the validity of the inferences drawn from such task-based assessments in both quantitative and qualitative manner.

References

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.
Bachman, L. F. (2002). Some reflections on task-based language performance assessment. *Language Testing, 19*(4), 453–476.
Baker, D. (1990). *A guide to language testing*. Edward Arnold.
Brindley, G. (1998). Outcomes-based assessment and reporting in language learning programmes: A review of the issues. *Language Testing, 15*(1), 45–85.
Brown, J. D., Hudson, T. D., Norris, J. M., & Bonk, W. (2002). *Investigating task-based second language performance assessment*. University of Hawai's Press.
Bygate, M., & Samuda, V. (2005). *Integrative planning through the use of task repetition*.
Candlin, C. (1987). Towards task-based language learning. In C. Candlin & D. Murphy (Eds.), *Language learning tasks* (pp. 5–22). Prentice Hall.
Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment (CEFR)*. Cambridge University Press.
Dörnyei, Z., & Scott, M. L. (1997). Communication strategies in a second language: Definitions and taxonomies. *Language Learning, 47*(1), 173–210.
Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics in Language Teaching, 45*(3), 215–240.
Ishikawa, T. (2008). The effect of task demands of intentional reasoning on L2 speech performance. *Journal of Asia TEFL*, *5*(1).
Iwashita, N., McNamara, T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information-processing approach to task design. *Language Learning, 51*(3), 401–436.
Kramsch, C. (1986). From language proficiency to interactional competence. *The Modern Language Journal, 70*(4), 366–372.
Long, M. H. (1985). A role for instruction: Task-based language teaching. In M. Pienemann & K. Hyltenstam (Eds.), *Modeling and assessing second language acquisition* (pp. 77–99). Erlbaum.
McNamara, T. F., & Macnamara, T. J. (1996). *Measuring second language performance*. Longman Publishing Group.
Nakatsuhara, F., Inoue, C., & Taylor, L. (2020). Comparing rating modes: Analysing live, audio, and video ratings of IELTS Speaking Test performances. *Language Assessment Quarterly*, 1–24.
Nakatsuhara, F., Inoue, C., & Taylor, L. (2017). *An investigation into double-marking methods: comparing live, audio and video rating of performance on the IELTS Speaking Test*. The IELTS Partners: British Council, IDP: IELTS Australia and Cambridge English Language Assessment.
Norris, J. M. (2016). Current uses for task-based language assessment. *Annual Review of Applied Linguistics, 36*, 230.
Norris, J. M., Norris, J. M., Brown, J. D., Hudson, T., & Yoshioka, J. (1998). *Designing second language performance assessments* (No. 18). National Foreign Language Resource Centre.
Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*(4), 492–518.
Prabhu, N. S. (1987). *Second language pedagogy* (Vol. 20). Oxford University Press.
Prabhu, N. S. (2019). *Perceptions of Language pedagogy* (Ed). Orient Blackswan Private Limited.
Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27–57.

Robinson, P. (Ed.). (2011). *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (Vol. 2). John Benjamins Publishing.
Robinson, P., & Ross, S. (1996). The development of task-based assessment in English for academic purposes programs. *Applied Linguistics, 17*(4), 455–476.
Shehadeh, A. (2018). Task-based language assessment. *The TESOL Encyclopedia of English Language Teaching*, 1–6.
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
Tamboli, V. (2017). *The role of task complexity and task conditions in ESL speaking assessment*. [Unpublished doctoral dissertation]. The English and Foreign Languages University.
Tarone, E. (1998). Research on interlanguage variation: Implications for language testing. *Interfaces Between Second Language Acquisition and Language Testing Research*, 71–89.

Chapter 12
Validation of a Large-Scale Task-Based Test: Functional Progression in Dialogic Speaking Performance

Chihiro Inoue and Fumiyo Nakatsuhara

Abstract A list of language functions is usually included in task-based speaking test specifications as a useful tool to describe target output language of test-takers, to define target language use (TLU) domains and to specify task demands. Such lists are, however, often constructed intuitively, and they also tend to focus solely on the types of function to be elicited and ignore the ways in which each function is realized across different levels of proficiency (Green, 2012). The study reported in this chapter is a part of a larger-scale test revision project for Trinity's Integrated Skills in English (ISE) spoken examinations. Analyzing audio-recordings of 32 performances on the ISE spoken examination both quantitatively and qualitatively, the aims of this study are (a) to empirically validate lists of language functions in the test specifications of the operational, large-scale, task-based examinations, (b) to explore the usefulness and potential of function analysis as a test task validation method, and (c) to contribute to a better understanding of varied test-taker language that is used to generate language functions.

Keywords Task-based test · Language functions · Dialogic tasks · Proficiency levels · Speaking

Introduction

Challenges of Large-Scale TBLA

Much like in language teaching, the field of language testing has been strongly influenced by the rises of different theories of language learning (Nakatsuhara et al., 2021). The goal of a speaking test is to elicit and rate an appropriate sample of speech from test-takers in a short period of time, so as to make inferences about their ability to use the language in real life (Bachman, 1990). Seeing the advancement of the

C. Inoue (✉) · F. Nakatsuhara
University of Bedfordshire, Luton, UK
e-mail: chihiro.inoue@beds.ac.uk

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_12

communicative language teaching approach in the 1970s, language testers started to explore how language ability could be assessed in a more communicative way. Large-scale language tests are no exception, and a task-based language assessment (TBLA; Brindley, 1994; Wigglesworth & Frost, 2017) approach has been employed in various international tests today. Examples of such large-scale task-based speaking tests include ACTFL Oral Proficiency Interview, Trinity Graded Examinations in Spoken English, and speaking papers in LanguageCert International ESOL and Cambridge English Qualifications. These speaking tests use dialogic speaking tasks, such as role-play, presentation followed by a question-and-answer session, conversation and discussion, which are designed to elicit different speech acts and language functions through simulated real-life speaking activities.

As pointed out by Timpe-Laughlin (2018), the two major challenges in TBLA lie in designing the tasks that elicit sufficient amount and appropriate quality of language samples under a communicative situation, as well as developing rating scales to evaluate the elicited language to draw inferences on learners' ability to use the language communicatively. Addressing these two challenges, some TBLA researchers (e.g. Kuiken & Vedder, 2018; Youn, 2018) have developed well-designed communicative test tasks together with relevant rating scales to capture the degree of communicative success and salient features of the elicited language. While these recent studies have enhanced our understanding of a specific task format or task variation, these cannot be readily applied to large-scale operational tests, where multiple tasks are usually included to evaluate an overall spoken ability of learners and comparability across a number of test administrations is crucial (Weir & Wu, 2006). Test providers cannot keep using the same tasks because there is a shelf life for operational test tasks, after which they must be retired, in order to maintain test security. Therefore, it is of paramount importance for large-scale tests to keep developing different, but comparable, task and test versions, and it is enabled and ensured by detailed test specifications.

Task Validation: Expected Versus Elicited Functions

In large-scale language tests, tasks are developed based on test specifications, which specify the complexity of the prompts and expected performance in terms of the levels of vocabulary, grammar, functions, discourse type etc. (Taylor, 2011). The specifications are the crucial blueprint of a test, since they outline what kinds of learner language are expected to emerge at which stages (i.e. tasks) during test performance (Weir, 2005). As noted above, test specifications enable producing parallel task and test versions; i.e. different task and test versions that are designed to have comparable demands in order to elicit comparable samples of speech across test-takers and administrations (Inoue, 2013).

One of the effective ways to design a specification for task-based speaking tasks is to specify the target speech acts broadly and list relevant language functions that may be necessary to carry out the speech acts (O'Sullivan et al., 2002). By specifying

the required functions to perform the tasks in the test specifications, some variations among task versions (e.g. the roles assigned in role-play tasks) are allowed, yet the target aspect of test-takers' functional knowledge (Bachman & Palmer, 1996) can remain unchanged. Therefore, it can contribute to enhanced comparability of task and test versions, which warrants meaningful comparisons among different test administrations and ensures fairness to test-takers. Thus, focusing on the functions is a useful, widely applicable way to examine the validity of the test tasks (Green, 2012), especially for dialogic speaking tasks where a wider variety of functions are expected to be used.

After the tasks are designed based on test specifications, checks must be carried out on the match between what is expected (as stated in specifications) and what is actually elicited (used by test-takers), as part of a test's validation (Weir, 2005). Regarding the mapping of the expected and elicited functions in a task or test, there are two research gaps; the first one is that the mapping exercise is seldom reported in the literature (but see Nakatsuhara & Dunlea, 2020). The second gap is that the lack of empirical work on how the functions are actually realized using the language at different levels of speaking proficiency under examination conditions (Green, 2012). There is some evidence from developmental studies in pragmatic abilities, which focused on language functions that are related to certain speech acts such as apology and request (e.g. Blum-Kulka & Olshtain, 1986; Rose, 2000), refusal (e.g. Houck & Gass, 1996), suggestion and rejection (Bardovi-Harlig & Hartford, 1993), but they do not cover the wide range of functions often listed in the specifications for large-scale speaking tests.

Systematic Investigation into Expected and Elicited Functions

As noted earlier, a list of expected language functions is often included in speaking test specifications (Nakatsuhara et al., 2017) as a useful means to describe target output language of test-takers related to their functional knowledge (Bachman & Palmer, 1996), grammatical encoding abilities (Field, 2011), as well as the situations where learners use the target language that the tasks aim to simulate (Weir, 2005). The instrument that is used most often to validate a function list is arguably the observation checklist developed by O'Sullivan et al. (2002). Building on earlier work of Bygate (1987) and Weir (1993), O'Sullivan et al. reported the development, refinement, and successful application of the function checklist to the Cambridge Main Suite examinations. It was suggested that the language of a speaking test can be categorized into three broad types: *informational functions* (e.g. providing personal information and expressing opinions), *interactional functions* (e.g. agreeing and asking for information), and *interaction management functions* (e.g. initiating and changing topics). While the checklist was originally developed for analyzing the language functions elicited in paired speaking tasks of the Cambridge Main Suite examinations, its applicability for other speaking tests with different formats has been demonstrated, such as the IELTS Speaking Test (Brooks, 2003), the GESE

examinations (O'Sullivan et al., 2011) and the Test of English for Academic Purposes (Nakatsuhara & Dunlea, 2020).

Such function lists in test specifications, however, tend to focus only on the types of function to be elicited, and the validation reported in scarce literature also tends to be limited to mere box-ticking exercises (e.g. Zhou et al., 2018). Reports on test validation using a function checklist often neglect the ways in which each function is realized across different proficiency levels. As Green (2012) states, any functions can be realized in various manners, and many functions start to emerge at lower levels with basic language, but test-takers at higher levels may also use the same functions with higher linguistic precision and more sophistication. As such, sufficient elaboration on the types of language observed at different levels of proficiency would enhance the usefulness of function lists, more accurately describing the test tasks' potential to elicit and capture the features of performance for effective evaluation of test-takers' speaking ability. As Timpe-Laughlin (2018) states, looking into the actual language used to perform the functions would help to demonstrate that the test tasks elicit performance as intended, as well as to provide empirical evidence of the qualities of performance at different ability levels, which would enhance the validity of scoring of the test. With these research gaps in mind, this small-scale study aims to contribute to a better understanding of the functional progression across different levels of speaking proficiency, as well as of the methodology to conduct an empirical validation on the lists of language functions.

The Context of This Study

This study was part of a test revision project[1] that was carried out in 2013 on the Integrated Skills in English (ISE) exams, a suite of exams at 5 levels that is designed and managed by Trinity College London. The data in this study came from the specifications and transcripts of the oral interviews of the ISE prior to the revision.

The ISE interviews are conducted in a one-to-one format with a test-taker and a Trinity examiner. This study focused on the first 4 ISE exam levels: ISE 0,[2] I, II, and III, which correspond with CEFR A2, B1, B2, and C1, respectively (Papageorgiou, 2007). According to the exam information document (Trinity College London, 2013: 13), the ISE interviews are designed to 'replicate real-life exchanges in which the test-taker and the examiner pass on information, share ideas and opinions and debate topical issues'.

The structure of the ISE interview at each level prior to the revision is summarized in Table 12.1.

For all ISE levels, two tasks were included: *topic discussion*, in which a test-taker and an examiner discusses a topic that the test-taker had chosen and prepared

[1] This project was funded and supported by Trinity College London, and this chapter is based on unpublished project reports submitted to Trinity College London.

[2] After the revision, ISE 0 has been renamed as ISE Foundation.

Table 12.1 Structure and duration of the pre-revision ISE interview (Trinity College London, 2013)

ISE 0 (8 min)	ISE I (8 min)	ISE II (12 min)	ISE III (20 min)
			1. Formal topic presentation
			2. Topic discussion
		1. Topic discussion	3. Interactive task
1. Topic discussion	1. Topic discussion	2. Interactive task	4. Listening task
2. Conversation	2. Conversation	3. Conversation	5. Conversation

for beforehand, and *conversation*, where they talked about one (at ISE 0 and I) or two (at ISE II and III) subject areas chosen by the examiner from a list.[3] At ISE II and III, *interactive task*[4] was included, in which the examiner presented a dilemma or problem and the test-taker needed to proactively engage in the interaction by asking questions and discussing possible solutions with the examiner. At ISE III, further two tasks were added: a *formal topic presentation,* prepared by the test-taker beforehand, and a *listening task* (the listening task was not included in this study). Sample questions of the conversation and interactive tasks can be found in Appendix 1.

With the increasing variety and demands of the tasks at each exam level, test-takers were 'expected to display their ability to use the functional, grammatical, lexical, and phonological items specified for the level and the preceding levels (Trinity College London, 2013:13). Accordingly, the ISE interview specifications included a list of functions as well as those of grammar and vocabulary that were expected at each exam level, and this study focused on the lists of functions for ISE 0, I, II, and III.

The aim of this study was twofold. One was to validate the function lists in the specifications for the speaking tasks. The other was to explore features of language elicited across the four ISE level exams (0, I, II, and III) and across score bands (Bands A, B, C, and D) within respective ISE exams, so as to inform the revision of what was then holistic rating scales into analytic ones. While its immediate contributions were for the ISE's test specifications, examiner training materials, and task specifications, there are wider implications for the field of large-scale task-based testing, in terms of offering a better understanding of the functional and linguistic progression of learner language observed in dialogic speaking tasks.

Research Questions

- **RQ1**. To what extent does each exam level (ISE 0, I, II, and III) elicit intended language functions?

[3] The list of conversation themes is publicly available in the *Guide for Teachers* for each exam level on the ISE website.

[4] After the revision, interactive tasks have been renamed as *collaborative* tasks.

Table 12.2 Number of test-takers whose performance were analyzed at each Band across the four ISE levels

	ISE 0	ISE I	ISE II	ISE III
Band A (Pass with Distinction)	2	2	2	2
Band B (Pass with Merit)	2	2	2	2
Band C (Pass)	2	2	2	2
Band D (Fail)	2	2	2	2

- **RQ2**. Are there any differences observed in the ways in which these language functions are realized across different exam levels (ISE 0, I, II, and III) and across different band scores (Bands A, B, C, and D) within an exam level?

Methodology

Audio-recordings from 32 test-takers of the ISE spoken examination, comprising of eight test-takers from each of the four ISE levels (ISE 0, I, II, and III), were made available by Trinity. As shown in Table 12.2, the eight recordings within each ISE level included two test-takers each at Bands A, B, C, and D (i.e. Pass with Distinction, Pass with Merit, Pass, and Fail). In the ISE spoken test, test-takers' performances are rated separately on individual tasks, and the test-takers featured in this study were those who received straight As, Bs, Cs, or Ds at the exam level that they took (e.g. a Band C test-taker at ISE 0 received a C for each of the two tasks, and a Band B test-taker at ISE III received a B for each of the five tasks). For the purpose of this research, these samples were carefully selected by Trinity's senior examiner trainer to best represent the four bands across the four ISE levels.

The recordings were transcribed using a simplified version of Conversation Analysis (CA) notation (Atkinson & Heritage, 1984; see Appendix 2 for the CA symbols used in the study). After transcribing all the recordings, the research proceeded in four stages:

Stage 1: In order to systematically code observed language functions, it was first of all necessary to map the list of target language functions at each ISE level (Trinity College London, 2009) against O'Sullivan et al.'s (2002) function categories (see Table 12.3[5]). The aim of this exercise was threefold: (a) to better understand the nature of each intended function by redefining them in the light of a widely accepted, comprehensive framework, thus improving the reliability of coding, (b) to locate the individual lists of language functions for four exam levels in a single framework to

[5] The scarcity of intended functions at ISE 0 and ISE I in the categories of interactional functions and managing interactions in O'Sullivan et al.'s list was expected, as the tasks at these exam levels do not require the test-taker to use more advanced language functions such as persuading and reciprocating.

Table 12.3 Mapping Trinity's (2009) intended language functions onto O'Sullivan et al.'s function categories (with abbreviations for coding)

O'Sullivan et al. (2002)		Trinity's intended language function list			
Informational Functions		**ISE 0**	**ISE I**	**ISE II**	**ISE III**
Providing personal information	*Present circumstances* **(PP)**	• Giving personal information about present circumstances/activities **(PC)** • Describing routines **(DR)** • Expressing ability and inability **(ABL)**			
	Past experiences **(PE)**	• Giving personal information about past circumstances/activities **(PE)**	• Describing past actions in the indefinite and recent past (i) **(PE)**		
	Future plans **(FP)**	• Describing future plans **(FP)**	• Describing the future: informing + Expressing intention **(FPI)**		
Expressing opinions/ preferences/feelings **(EOP)**	*Express opinions/preferences?*	• Expressing likes and dislikes **(EOP)**	• Giving opinions and preferences **(EOP)**	• Expressing and expanding ideas and opinions **(EOP) (ELB)**	• Developing an argument **(ARG) (EOP) (ELB)**
Elaborating **(ELB)**	*Elaborate on, or modify an opinion*				
Justifying opinions **(JUS)**	*Express reasons for assertion s/he has made?*		• Giving reasons **(JUS)**		• Justifying an argument **(JUS)**

(continued)

Table 12.3 (continued)

O'Sullivan et al. (2002)		Trinity's intended language function list			
Comparing **(COM)**	*Compare things/people/events*	• Expressing simple comparisons **(COM)**		• Highlighting advantages and disadvantages **(A/D)**	
Speculating **(SPC)**	*Speculate?*		• Describing the future: predicting + Expressing certainty and uncertainty **(SPC)**	• Speculating **(SPC)**	• Deducing and inferring **(SPC)**
Staging **(STG)**	*Separate out or interpret the parts of an issue?*				
Describing **(DSC)**	*Describe a sequence of events/people/information*	• Describing people, objects and places **(DSC)**	• Describing past actions in the indefinite and recent past (ii) **(DSC)**		
Summarizing **(SUM)**	*Summarize what he/she has said?*				• Summarizing **(SUM)**
Suggesting **(SUG)**	*Suggest a particular idea?*			• Giving advice **(SUG)**	
N/A			• Expressing obligation **(OB)**		
N/A					• Evaluating options, past actions/course of events, different standpoints **(EVL)**

(continued)

Table 12.3 (continued)

O'Sullivan et al. (2002)		Trinity's intended language function list			
Interactional Functions		**ISE 0**	**ISE I**	**ISE II**	**ISE III**
Agreeing **(AG)**	*Agree with an assertion made by another speaker?*			• Expressing agreement **(AG)**	• Expressing empathy **(EMP)**
Disagreeing **(DIS)**	*Disagree with what another speaker says*			• Expressing disagreement **(DIS)**	
Modifying **(MOD)**	*Modifying arguments or comments made by another speaker*				• Expressing reservations **(RES)**
Asking for opinions **(AOP)**	*Asking for opinions*		• Asking for opinions **(AOP)**		
Asking for information **(AIN)**	*Asking for information?*	• Asking for information, e.g. simple questions about everyday life **(AIN)**	• Asking for information **(AIN)**	• Eliciting further information **(AIN)**	
Persuading **(PRS)**	*Persuading*			• Persuading and discouraging **(PRS)**	
Conversational repair **(CR)**	*Repair breakdowns in interaction*				
Negotiating meaning **(NEG)**	*Check understanding*			• Negotiating meaning **(NEG)**	
	Indicate understanding of points made by partner				
	Establish common ground/purpose or strategy				

(continued)

Table 12.3 (continued)

O'Sullivan et al. (2002)		Trinity's intended language function list			
	Ask for clarification when an utterance is misheard/misinterpreted				
	Correct an utterance made by other speaker which is perceived to be incorrect or inaccurate				
	Respond to requests for clarification				
N/A					• Expressing sympathy **(SYM)**
Managing Interaction Functions		**ISE 0**	**ISE I**	**ISE II**	**ISE III**
Initiating **(INIT)**	*Start interaction*				
Changing **(CHN)**	*Change topic*				
Reciprocating **(REC)**	*Share responsibility for developing interaction*				
Deciding **(DEC)**	*Come to a decision*				

allow for comparisons across the four levels and (c) to enable systematic identifications of the functions that were not listed in the Trinity list but were observed in test-taker performances. Trinity's senior examiner trainer and the two researchers, who were all familiar with O'Sullivan et al.'s function checklist, carefully examined and discussed Trinity's intended functions. The transcripts from the ISE recordings as well as example excerpts in O'Sullivan et al.'s categories were compared to help the mapping exercise. The scrutiny process also served to consolidate the researchers' common understanding of each function, helping to further build a coder consensus for the next coding stage.

Stage 2: Transcripts were arranged in Microsoft Excel spreadsheets, segmenting the entire examiner and test-taker interactions by turn. The columns of the spreadsheet were arranged by: line number, test-taker ID, band score, task phase, observed language functions, transcribed turn, and each line represented a single turn either by an examiner or by a test-taker. While only test-taker turns were to be coded in this study, it was important to include examiner turns there, so that the function(s) of each test-taker turn could be assessed appropriately in the given interactional sequence. Since one turn can serve more than one language function, coding of a turn involved identifying and entering all functions served by each test-taker turn (see Table 12.3 for the abbreviations used in coding). All transcripts were double-coded. First, the two researchers coded a separate batch of transcripts, and the coded transcripts were then swapped to review the codes entered by the other researcher. Any queries, inconsistencies, and suggested changes in coding were flagged in red, and discussed between the two researchers until complete agreement was obtained.

Stage 3: To answer RQ1, the number of turns that involved each language function in the four levels of the test was counted, and divided by the number of test-takers at each level. This was to calculate how many turns were on average produced by test-takers to realize each language function across all the tasks for each exam level. The analysis was carried out only for the entire test at each level. Due to the small sample size available, only descriptive statistics were used to report the results.

Stage 4: The ways in which language functions were realized across the four exam levels were then qualitatively analyzed to address RQ2. Test-taker turns as well as relevant examiner turns were extracted for each language function, and they were compared across different ISE levels, as well as across different score bands within each ISE level. Task-specific language features were also described as part of the qualitative analysis. As a preliminary analysis for RQ2, we also investigated the amount of overall production, lexical variety, and lexical sophistication per exam level using the corpus-based frequency lists that cover the British National Corpus (BNC) and Corpus of Contemporary American English (COCA) (up to K20 lists) as well as Academic Word List (AWL) using the Web-based Text Inspector programme (WebLingua, n.d.).

Results and Discussion

Preliminary Analysis Per Task (RQ1)

Before presenting the results for RQ1, Table 12.4 shows the mean number of words produced on each task at each exam level and band. Although each band at each exam level consists of a very small number of test-takers (i.e. two test-takers, as shown earlier in Table 12.2), Table 12.4 offers an overall indication of the amount of language in which the functions were elicited.

In Table 12.4, the mean number of words for the topic discussion task for ISE 0, I, and II shows a steady increase from Band D (Fail) to Band A (Pass with Distinction), although some SDs are very large (e.g. 137.9 words for Band D at ISE II), indicating the amount of language on this task varied among individuals. It can also be observed that test-takers at higher bands did not always produce a larger amount of language than those at lower bands (e.g. Band C test-takers for ISE III produced an average of 521.5 words on the topic presentation task while Band A test-takers on average produced 363 words).

Language Functions Observed at Each ISE Level (RQ1)

Table 12.5 presents the target language functions in each ISE level and the average number of turns in which each function was produced per test-taker across the four levels of the test. In the 'Target' column, the target functions found in the function lists are specified using the abbreviations included in Table 12.3.

In the 'Elicited' columns, only those functions with an average realization rate of 0.7 turns or above per test-taker are ticked (✓), based on the project team's agreement on the threshold for identifying the main functions elicited in the test. The threshold is also in line with other recent test task validation studies such as Nakatsuhara and Dunlea (2020) and Zhou et al. (2018). Additionally, a triangle sign (△) indicates those functions with an average realization rate of 0.50–0.69 turns.

Therefore, the first function of the table, *giving personal information: present*, can be interpreted as follows; it was observed at ISE 0 in the form of *giving personal information about present circumstances/activities* (PC) and *describing routines* (DR) with average turns of 8.25 and 2.38, respectively (shown with a tick in the 'Elicited' column), while there were not sufficient utterances for *expressing ability and inability* (ABL) (shown with a triangle in the 'Elicited' column). The *giving personal information: present* function was not targeted at the remaining ISE levels, but average turns of 4.25, 2.88 and 2.13 were observed at ISE I, ISE II, and ISE III, respectively.

As expected from the number and types of tasks as well as the target proficiency level and allowed timings for each ISE level, the results demonstrate a clear progression across the four levels of the test in terms of the number of main language

Table 12.4 Descriptive statistics of total words produced per task at each band and exam level

Band		ISE 0		ISE I		ISE II			ISE III			
		Topic Discussion	Conversation	Topic Discussion	Conversation	Topic Discussion	Interactive	Conversation	Topic Presentation	Topic Discussion	Interactive	Conversation
A	M	243	108.5	196.5	127	541.5	294	477	363	408.5	311.5	311
	(SD)	(18.4)	(30.4)	(12.0)	(7.1)	(47.4)	(130.1)	(55.2)	(80.6)	(316.1)	(67.2)	(50.9)
B	M	196	55.5	369.5	150.5	415.5	300	441	453	254	304	233
	(SD)	(18.4)	(24.7)	(17.7)	(44.5)	(41.7)	(52.3)	(5.7)	(301.2)	(45.3)	(26.9)	(84.9)
C	M	223.5	97	247.5	116.5	320.5	168.5	250.5	521.5	268.5	270.5	169
	(SD)	(27.6)	(35.4)	(84.1)	(51.6)	(116.7)	(9.2)	(26.2)	(36.1)	(65.8)	(24.7)	(63.6)
D	M	102.5	53.5	239.5	89.5	280.5	106	204	280.5	345	237.5	197
	(SD)	(3.5)	(27.6)	(102.5)	(24.7)	(137.9)	(50.9)	(56.6)	(58.7)	(15.6)	(4.9)	(29.7)

Table 12.5 Language functions targeted and elicited across the four ISE levels

Informational functions	ISE 0		ISE I		ISE II		ISE III	
	Target	Elicited	Target	Elicited	Target	Elicited	Target	Elicited
Giving personal info: present (PP)	PC	✓(8.25)		✓(4.25)		✓(2.88)		✓(2.13)
	DR	✓(2.38)						
	ABL							
Giving personal info: past (PE)	PE	✓(6.63)	PE	✓(3.25)		✓(6.88)		✓(2.63)
Giving personal info: future (FP)	FP	✓(3.00)	FPI	✓(2.13)		✓(1.38)		
Expressing opinions/ preferences/feelings (EOP)	EOP	✓(5.13)	EOP	✓(6.50)	EOP	✓(18.13)	EOP	✓(15.38)
							ARG	✓(17.88)
Elaborating (ELB)		✓(1.13)		✓(5.88)	ELB	✓(6.75)	ELB	✓(11.75)
Justifying opinions (JUS)		✓(2.38)	JUS	✓(4.50)		✓(2.63)	JUS	✓(9.88)
Comparing (COM)	COM	✓(2.00)		✓(1.25)		✓(1.25)		✓(1.50)
					A/D			
Speculating (SPC)			SPC	✓(1.38)	SPC	✓(1.75)	SPC	✓(4.00)
Staging (STG)				△(0.50)		✓(0.88)		✓(1.63)
Describing (DSC)	DSC	✓(4.38)	DSC	✓(5.50)		✓(11.88)		✓(20.50)
Summarizing (SUM)							SUM	✓(3.25)
Suggesting (SUG)					SUG	✓(4.38)		
Expressing obligation (OB)			OB	✓(2.25)				
Evaluating (EVL)							EVL	
Interactional functions								
Agreeing (AG)				✓(1.63)	AG	✓(1.88)	AG	✓(5.50)
							EMP	
Disagreeing (DIS)					DIS			
Modifying (MOD)								✓(1.63)
							RES	
Asking for opinions (AOP)			AOP	△(0.63)		✓(2.88)		✓(3.25)
Asking for info (AIN)	AIN	✓(1.00)	AIN	✓(0.88)	AIN	✓(4.63)		✓(2.00)
Persuading (PRS)					PRS	△ (0.63)		✓(0.88)
Conversational repair (CR)								

(continued)

Table 12.5 (continued)

Informational functions	ISE 0		ISE I		ISE II		ISE III	
	Target	Elicited	Target	Elicited	Target	Elicited	Target	Elicited
Negotiating meaning (NEG)		✓(2.38)		✓(4.38)	NEG	✓(7.63)		✓(11.75)
Expressing sympathy (SYM)							SYM	
Managing interaction functions								
Initiating (INIT)								
Changing (CHN)						✓(1.50)		
Reciprocating (RES)						✓(2.38)		✓(2.00)
Deciding (DEC)								

functions elicited and the number of turns that were produced to express those functions.

The analysis confirms that almost all intended language functions at each level were sufficiently elicited, validating the function lists in the ISE test specifications. There were also some cases where a similar or broader function was observed, even when the exact function as specified in Trinity's list was not observed. All four levels also elicited many additional functions that were not explicitly targeted in the test specifications. The information presented in Table 12.4 can be summarized under the four categories:

a. target elicited,
b. target not sufficiently elicited but a similar/broader function elicited,
c. target not sufficiently elicited, and
d. non-target elicited (i.e. The function was not on the list but was observed in actual performance).

The classification is useful to verify the existing intended function list and/or suggest modifications to the function list. As noted earlier, we used function names specified in Trinity's target function list, and where target functions were absent, O'Sullivan et al.'s (2002) function categories were applied.

ISE 0

a. **Target elicited**: Giving personal information about present circumstances/activities; describing routines; giving personal information about past circumstances/activities; describing future plans; expressing simple comparisons; describing people, objects and places; asking for information
b. **Target not sufficiently elicited but a similar/broader function elicited**: Expressing ability and inability; expressing likes and dislikes
c. **Target not sufficiently elicited**: N/A

d. **Non-target elicited**: Elaborating; justifying opinions; negotiating meaning.

ISE I

a. **Target elicited**: Describing past actions in the indefinite and recent past; describing the future: informing + expressing intention; giving opinions and preferences, giving reasons, describing the future: predicting + expressing certainty and uncertainty; expressing obligation; asking for information
b. **Target not sufficiently elicited but a similar/broader function elicited**: N/A
c. **Target not sufficiently elicited**: Asking for opinions
d. **Non-target elicited**: Providing personal information (present circumstances, past experiences), elaborating, comparing; negotiating meaning.

ISE II

a. **Target elicited**: Expressing and expanding ideas and opinions, justifying an argument; speculating; giving advice; expressing agreement; eliciting further information; negotiating meaning
b. **Target not sufficiently elicited but a similar/broader function elicited**: Highlighting advantages and disadvantages
c. **Target not sufficiently elicited**: Expressing disagreement, persuading
d. **Non-target elicited**: Providing personal information (present circumstances, past experiences, future plans), justifying opinions, comparing; staging, describing; agreeing; asking for opinions; changing; reciprocating.

ISE III

a. **Target elicited**: Developing an argument; deducing and inferring; summarizing
b. **Target not sufficiently elicited but a similar/broader function elicited**: Expressing empathy; expressing reservations
c. **Target not sufficiently elicited**: Evaluating options, past actions/course of events, different standpoints; expressing sympathy
d. **Non-target elicited**: Providing personal information (present circumstances, past experiences), comparing; staging; describing; agreeing; modifying; asking for opinions; asking for information; persuading; negotiating meaning; reciprocating.

The classification also indicates the course of action to be taken for improving the function lists in the ISE speaking tasks (Trinity College London, 2009). That is, those functions under Category (a) should remain as is. Those under Category (b) should also stay but their names need to be reconsidered to reflect the nature of the elicited functions more accurately. For Category (c), however, careful consideration has to be given, as they were intended but not sufficiently observed in the test-taker performances. Recommendations could be either removing these functions from the specifications or revisiting test tasks if they are deemed to play an important role in tapping the construct of the specific level of the ISE exams. Lastly, the functions classified in Category (d) need to be examined individually, since while some functions emerged

at higher levels (e.g. *Speculating, Changing topics*), other functions were consistently observed across the four levels (e.g. *Expressing opinions/preferences/feelings, Asking for information, Negotiating meaning*). While some of these functions were mentioned elsewhere in the exam guides (Trinity College London, 2013) such as in the 'task' descriptions or 'skills' required, they were not mentioned in the function lists. A suggestion was therefore made to include them in the function lists, so that coherent sets of information would be presented in the exam guides.

The findings for RQ1 suggest that some functions that are elicited across the four exam levels especially deserve scrutiny for RQ2 regarding how they are actually realized. For example, an informational function, *expressing opinions/preferences/feelings*, was elicited across the four levels and increasingly more as exam levels progressed; an average number of turns allocated for the function was 5.13 at ISE 0, 6.50 at ISE I, and 18.13 at ISE II. At ISE III, test-takers on average produced 17.88 turns of the *developing an argument* function in addition to 15.38 turns of *expressing opinions*. Similarly, an interactional function, *negotiating meaning*, is another example of this increasing pattern, where ISE 0 test-takers on average produced 2.38 turns, ISE I 4.38 turns, ISE II 7.63 turns, and ISE III 11.75 turns to perform this function.

Another interactional function, *asking for information*, was elicited throughout, although not showing a steady increase (1.00 turns at ISE 0, 0.88 turns at ISE I, 4.63 turns at ISE II, and 2.00 turns at ISE III). This function often co-occurred with *asking for opinions* at all three exam levels except for ISE 0 (0.63 turns at ISE I, 2.88 turns at ISE II, and 3.25 turns at ISE III). Because these two interactional functions are one of the key elements for successful performance, especially in ISE II and III, where test-takers led the interaction with the examiner in the interactive tasks, they will also be the focus of RQ2.

Preliminary Analysis Per Exam Level (RQ2)

Prior to reporting the qualitative analysis for RQ2 on the language elicited at each exam level, some descriptive statistics are first presented in order to show an overall picture, including indices of the amount of production, lexical variety, and lexical sophistication per exam level. Table 12.6 summarizes the descriptive statistics.

The mean total number of words uttered by test-takers during the test showed a steady increase with the exam levels (279.8 for ISE 0, 399.9 for ISE I, 973.1 for ISE II, and 1274.9 for ISE III). Similarly, the average VOCD, an index for lexical variety which is not affected by the text length, increases with the exam levels (58.2 for ISE 0 and I, 70.8 for ISE II, and 74.7 for ISE III). These two indices demonstrate that test-takers who took the higher-level ISE exams generally spoke more with a wider lexical variety.

Lexical sophistication according to the percentages of words found in the frequency-based vocabulary lists according to different corpora (i.e. BNC, COCA, and AWL) did not present a linear pattern across the exam levels. However, when the

Table 12.6 Descriptive statistics of overall production, lexical variety and sophistication per level

	ISE 0 ($n = 8$)		ISE I ($n = 8$)		ISE II ($n = 8$)		ISE III ($n = 8$)	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Total no. of words	279.8	(89.3)	399.9	(116.5)	973.1	(332.6)	1274.9	(274.0)
VOCD	58.2	(13.9)	58.2	(11.5)	70.8	(14.9)	74.7	(14.0)
BNC 1 K	59.1	(3.7)	63.4	(5.0)	59.4	(2.7)	58.4	(3.6)
BNC 2 K	7.6	(2.8)	7.8	(1.5)	8.0	(1.1)	10.7	(1.3)
BNC 3 K	6.5	(1.5)	5.3	(2.1)	6.3	(1.1)	6.0	(0.7)
BNC 4-20 K	15.0	(2.6)	12.8	(3.1)	13.5	(2.3)	16.5	(3.3)
BNC Off-list	9.7	(3.8)	9.0	(2.3)	10.3	(2.5)	5.5	(2.3)
COCA 1 K	58.7	(3.9)	63.7	(5.7)	60.3	(3.0)	59.6	(4.3)
COCA 2 K	10.1	(3.5)	8.9	(2.9)	8.9	(2.4)	10.7	(1.4)
COCA 3 K	4.4	(1.3)	4.5	(1.5)	4.1	(1.7)	6.0	(1.9)
COCA 4-20 K	14.5	(2.9)	11.8	(3.3)	13.4	(1.6)	15.2	(3.4)
COCA Off-list	9.7	(3.8)	9.0	(2.3)	10.3	(2.5)	5.5	(2.3)
AWL All lists	2.1	(1.3)	2.4	(1.2)	2.3	(1.2)	5.8	(0.9)

Notes All values for BNC, COCA, and AWL lists are percentages based on types. 1 K indicates the list of most frequent 1000 words in the relevant corpus

elicited language was qualitatively examined across not only the exam levels but also across bands within an exam level, there seemed to be some indicative, distinctive differences. The next section presents the results of the qualitative analysis for RQ2.

Ways in Which Functions Were Realized (RQ2)

Following the quantitative analysis of elicited functions for RQ1, we now present and discuss how selected functions were linguistically produced across four exam levels, namely *expressing opinions/preferences/feelings*, *negotiating meaning, asking for information*, and *asking for opinions*. The results are reported briefly at the beginning of each section, followed by some examples of test-takers' utterances at different bands. In the excerpts, speakers are identified as A, B, C, and D, representing test-takers at Bands A, B, C, and D, respectively, and E denotes the examiners' utterances. Each excerpt is presented with the task (in brackets) in which the utterance was observed, indicated as Presentation (i.e. topic presentation), Discussion (i.e. topic discussion), Interactive, and Conversation. Due to space limitation, not all tasks are discussed here.

Expressing Opinions/Preferences/Feelings

ISE 0

Most test-takers *expressed likes and dislikes* in both topic discussion and conversation tasks. At Bands A and B, test-takers were able to demonstrate the function in conjunction with the function of *justifying*, elaborating the reasons that they liked something. These participants were able to use justification as they used subordinate clauses with basic causal markers (e.g. 'because'), and the use of contrastive markers (e.g. 'but') was also notable. In contrast, Bands C and D test-takers tended to use simple clauses.

- **A (Discussion)**: jazz er classical (.) but I prefer er playing a jazz than playing classical (.) because classical for me is er boring (laughs) kinds of music
- **B (Discussion)**: erm I like it staying in a hotel because erm erm […] the people tidying your bedroom
- **C (Discussion)**: so erm we: like listening music together we: like er eating er er a lot of food erm we: when we stay with my sister and her cousins we (.) erm like watching horror film
- **D (Conversation)**: I like in the zoo erm animals (.) for example lions (.) tigers (.) monkeys (.) I'm (.) I like (.) animals

ISE I

All the ISE I test-takers used the two functions of *giving reasons* and *giving opinions and preferences*. These two functions often occurred within the same turn using subordinate clauses (using 'because'), particularly in the topic discussion task. Test-takers at Bands C and above specified reasons clearly. In the example below, the emerging use of relative pronouns, auxiliary verbs, comparative adjectives (e.g. 'the radio message that would be easier), and reflective pronouns (e.g. 'yourself') is noteworthy in contributing to the clarity of their messages. Band D test-takers, however, managed to give opinions and preferences in a simple way, but their reasons were not always clear or logical. Moreover, Band D test-takers sometimes misunderstood questions and gave off-the-point responses.

- **A (Discussion)**: I think that erm we are not alone in the universe but erm aliens don- don't want to communicate with us by making crop circles (.) because the patterns are very complicated then they haven't got a precise meaning (.) maybe aliens er would er prefer to erm send the radio message that would be easier
- **B (Discussion)**: erm:: yes er because er er it have erm: er it put erm it in touch with people because it is it is erm a a about common people (.) so you er compare yourself with this contestants (.) in fact there are romantic contestants (.) there are love story and so er this is the reason it is so successful
- **C (Discussion)**: the acting I think it's a great because (.) it's a really a good story (.) a great cast and er a really good music (.) Moon River is a (.) such a sweet song (.) and I really love it

- **D (Discussion)**: I wouldn't be happy because (.) football is my life (.) but it's difficult (.) very difficult

ISE II

Again, this function was observed more often in the topic discussion task. Test-takers at Bands A and B frequently *expressed* and *expanded their ideas and opinions*, resulting in longer turns. They also used more sophisticated vocabulary, such as 'choreography' (found in K6 list) and 'innovative' (K4). Band C test-takers expressed their ideas and opinions with some reasons, but did not elaborate as much as Bands A or B test-takers. Neither elaborated explanation nor reasoning was observed at Band D.

- **A (Discussion)**: and I think she's great she started she starting uh she started in her dance academy dancing then she started to to to dance with companies to dance with shows and for example in two thousand and seven she was in a show called Carmen and uh she created the choreography she designed the dresses she she made so many things
- **B (Discussion)**: my favourite favourite books are: 'I kill' and 'I am god' they talk about two: mad (.) two mad people and erm that have suffered in this life. I like it because erm for the tech- techniques erm is very innovative in fact (.) for this time you can see both the points of view (.) the points of view of the good that er:m want to erm capt- erm to catch the killer and erm the point of view of the killer that erm suffered in his life
- **C (Discussion)**: because the: academy is more professional and (.) there I had also er to study the music and I: can er write my own song
- **D (Discussion)**: erm erm I me (.) is the (1.0) erm is most important than Rafael Nadal erm win (.) this year erm

ISE III

All test-takers of this level *developed an argument* extensively with a lot of *elaboration* at the beginning of the topic presentation task, where they spoke about a topic of their choice and prepared beforehand as part of the task requirement. This is a task which requires extensive elaboration and is not given at lower exam levels, ISE 0, I or II. Performing this function often involved *justifying* their views and opinions. Again, Bands A and B test-takers used more sophisticated words, such as 'superficial' (K5) and 'esteem' (K6). More complex subordinate clauses (e.g. 'while I think that when you're growing…' at Band A) were very commonly used as well as relative clauses (e.g. 'how they treat them' at Band B). At Bands C and D, not as many subordinate clauses were used, and more hesitation markers like pauses (indicated by (.)) and false starts were observed.

- **A (Presentation)**: when people talk about young people that suffer from anorexia I think it's very superficial the way they approach it (.) because they think that it's just about wanting to look good or wanting people to think (.) oh my god he looks so thin (.) while I think that when you're growing old when you're becoming an

adult, you start shaping your personality and if there's something that is going wrong then it's something that you're going to carry your whole life

- **B (Presentation)**: the self-esteem is er it may change time to time because er it's not a static (.) it's a lifelong process (.) and the main characters in the self-esteem of the children are their parents (.) their parents erm give them the confidence by different conducts like erm how they treat them (.) how they give them the attention (.) so that way their their children learn how to act in during the life
- **C (Presentation)**: I think that (.) er everyone has the the ability of learning another language (.) so when you learn another language you increase your erm knowledge (.) you learn more things than others that just have only one also er learning a second language er gets you (.) more intellectual growth in everything (.) in vocabulary or experiences erm learning a: language also ca- helps you to communicate with other people everywhere in the world
- **D (Presentation)**: I think it's erm the the the television (.) and internet (.) increase in a lot of attitudes in the kids (.) and in the (.) in the young people (.) and for the other hand in education nowadays erm increase the another options that if you teach only with your book (.) I think it's boring for the kids and and you need to improve your (.) your skills like erm use computers

The same trend continued in the next three tasks in the oral interview at ISE III. Presented below are the examples from the conversation task, where the test-taker and examiner talked about two topic areas (of examiners' choice from a list). Test-takers at Bands A, B, and C *expressed their opinions* following the examiners' questions and comments adequately, speculating on issues at times. Although performing the same function, Band D test-takers tended to show less accurate use of verb tenses (e.g. 'four years ago I live in a village') and word inflection (e.g. 'more small').

- **A (Conversation)**: okay here we have er five years high school and er but in other countries we have four years because we have the fourth year that is er er a bit (.) useless because you do things that are not important for the university
- **B (Conversation)**: well I think that human rights have have been changing erm during the the during life (.) because I remember that er my grandmother telling me that (.) at her times they didn't use to vote
- **C (Conversation)**: yes I do it because I think that er if we we don't er re: recycling the rubbish will be everywhere
- **D (Conversation)**: erm I'm (.) I like live in a city (.) I eh eh four years ago (.) I live in a village (.) erm when you: erm (.) when you (.) when you (.) when you was small you like live in a village because it is more small and are more erm er:m er:m (.) erm people.

Negotiating Meaning

ISE 0

Although the ISE 0 function lists did not include it, language to negotiate meaning was often observed across different bands, in relation to *asking for clarification* (when an

utterance is misheard/misunderstood). Test-takers at Bands A, B, and C were able to ask for the question to be repeated in a simple way (e.g. 'can you repeat?'), as shown below. However, Band D test-takers sometimes expressed the need by signalling that they did not understand (e.g. 'ah: erm?'), rather than explicitly asking for repetition.

ISE I

At ISE I, test-takers at Bands A and B were observed to *respond to requests for clarification* from the examiners smoothly, solving communication problems promptly.

A (Conversation)

A: erm the original trains you have to pay about 35 Euros
E: 35 Euros?
→**A**: yes
E: it's (.) it's cheaper to buy a ticket isn't it?

B (Discussion)

B: the first language that we must speak (.) and then (.) in the street erm in erm in erm like (.)
when you're talking with a friend with this is.
E: really? so (.) and so Spanish was your second language (.)
→**B**: yes yes (.) it's a second language.

In the example below, a Band C test-taker did not understand the phrase 'stay fit' and asked for repetition. After repeating twice, the examiner had to paraphrase twice in order to achieve comprehension. Once understood, the test-taker was able to provide a relevant answer, despite some minor errors (e.g. use of 'very fruit' instead of 'lots of fruits').

C (Conversation)

E: and erm what kind of food do we have to eat just to stay fit?
→**C**: why- can you repeat?
E: what kind of food do we should we do we have to eat to stay fit
→**C**: study-?
E: what kind of food do we have to eat to stay fit
→**C**: study fit?
E: to keep fit
→**C**: erm (1.0)
E: if I want to be fit (.) can I go to McDonalds all the time or
→**C**: erm no (.) erm you must erm (2.0) erm you mustn't eat a lot of food (.) you must eat very fruit because it's very good.

At Band D, test-takers negotiated meaning, but often needing more turns to achieve comprehension. As presented below, a test-taker *responded to a clarification request* from the examiner, which had to be repeated due to unclear pronunciation.

D (Conversation)

D: I must eh sit in class (.) I mus:t eh do (exams)
E: must what?
→**D**: do exams
E: jumps?
→**D**: exams
E: exams
→**D**: yes exams
E: oh yes (.) uh huh

ISE II

At this exam level, test-takers at Bands A and B sometimes negotiated meaning mainly to *establish common ground* to continue further interaction on the topic. In such interactions, both the examiner and test-taker contributed to the negotiation process as exemplified below.

A (Discussion)

E: do you mean the city (.) or the the the countryside (.) when you say environment?
→**A**: oh er the environment of the school
E: oh okay
→**A**: er how you say [inside the school how do you feel
E: [yeah yeah yeah the surroundings the atmosphere

B (Interactive)

E: yeah that's true you can stay in contact that's for sure (.) yeah but I think they're worried about things like crime [you know?
→**B**: [ah rapers and something like that
E: yeah: (.) well it's a boy it's a nephew but
→**B**: ah it's a nephew
E: yes it's possible i guess

In contrast, many more instances of meaning negotiation were observed in interactions with test-takers at Bands C and D. Many of these instances appeared to aim to solve communication breakdowns caused by test-takers' misunderstanding or mishearing of what the examiner said.

C (Discussion)

E: okay so erm if you (.) had been (.) a character of this film what would you've done in that situation?
→**C**: what?
E: if you were a character in this film what would you do?
C: I I go to the police (1.0) and then on (.5) then (known) this (.5) this man

It is worth noting that while Band C test-takers seemed to be capable of resolving breakdowns (although their clarification requests were not always elegantly made), Band D test-takers had sometimes continued a long sequence before a misunderstanding was spotted and breakdowns were not always successfully resolved.

D (Interactive)

E: well as I said I-I think it might be a very good experience (2.0) for him (.) at his age
D: what do you think is one of the greatest thing that hmm we can see all the places hmm we can know what type of culture they are going and we can know more traditional they are doing (.) so it is a good thing
E: well I- I agree with you but my sister needs convincing
D: I think it's not the best thing erm (2.0) but what do you think about this?
E: well as I said, I think it's I think it could be a very positive experience bu:t we need to persuade my sister
D: yeah (.) I see (5.0) I think it maybe the right thing (1.0) what you sister thinks
E: I thought you agreed with me?
(1.0)
E: I thought you agreed with me
→**D**: I do? I don't know.
E: I thought you agreed with me?
→**D**: yes.

ISE III

Test-takers *indicated understanding of the points* made by the examiner more clearly at the ISE III level. Higher band test-takers often went on to extend the suggested points and/or their arguments.

- **A (Interactive):** y:eah I think (.) I think that too (.) I agree on that (.) but I think that in this erm particular era it's very di- it's very different from what has been er I think the last decades (.) even because of the crisis don't you think it's very tough to to move on and at least in I'm thinking about you know job problems and there is a lot of stress in that as well
- **B (Discussion):** yeah (.) because they feel frustrated and they start like maybe er behaving or doing this bad conduct because they have they want to express their frustration yeah so I definitely agree with you
- **C (Discussion):** I agree on that (.) because I had the exp- experience of going over in United States (.) and they didn't care if I if I wanted to express I say (.) you have to talk in my language
- **D (Interactive):** yeah I'm sure (laughs) I'm sure is (.) is so bad for them (.) er: (.) erm (.) in your country erm the young people erm (.) prefer erm to look for a job or (.) or (.) or study?

Moreover, test-takers who passed ISE III were facilitating interactions using phrases to *establish common ground* with the examiner at times, such as 'like you told me' and 'like I said before'. This is another sub-function in negotiating meaning that helps the flow of the interaction, so it is recommended to be included in the revised function table. Although this function was also observed in ISE II exams, ISE III test-takers seemed to be able to perform this function at their own initiative, rather than being prompted by the examiner.

- **A (Interactive):** I think that sometimes er you know complaining is the way to let of the steam don't you think (.) it's a way of kind of =
- **B (Conversation):** main (.) main topics like I said political or maybe erm (.) the employments
- **C (Discussion):** I think that it's it has to because (.) I know that United States (.) only cares about English (.) or like you told me that in London also they're very lazy

Asking for Information and Asking for Opinion

Since the ways in which *asking for information* and *asking for opinion* were explicated interrelate with each other, these functions are presented and discussed together.

ISE 0

Most of the test-takers, regardless of the Bands, *asked a simple question* when prompted by the examiner during the interview (e.g. 'Do you want to ask me any questions?').

- **A (Conversation)**: yes (.) do you travel a lot (.) for your job?
- **B (Discussion)**: er yes (.) are you married?
- **C (Conversation)**: have you ever travelled to Latin America?
- **D (Discussion)**: erm: (.) do you (.) play any sports?

ISE I

All the test-takers except one at Band D *asked questions* to the examiners, particularly on the conversation task. Higher band test-takers, especially at Band A, were able to *ask for opinions* and *reciprocate* according to the examiner's answers, but at lower bands, conversations did not usually expand, and Band D test-takers did not ask the examiner for their opinions.

- **A (Conversation)**: erm: what do you think (.) are they made by aliens or erm
- **B (Conversation)**: yes (.) erm (.) it's important for your earning money or save money?
- **C (Conversation)**: I prefer running and you?
- **D (Conversation)**: erm (3.0) do you know (.) some Catalan people that is important (.) or no?

ISE II

Test-takers at Bands A, B, and C often *asked the examiner for further information*, in particular during the interactive task performance, where test-takers needed to take the initiative to maintain the conversation on a dilemma or problem displayed by the examiners. Test-takers at Bands A, B, and C responded to examiners' utterances adequately with relevant questions to elicit more information.

- **A (Interactive)**: mm have you ever talked with them with her about er this
- **A (Interactive)**: um so it's a matter of money does she has a problem of money?
- **B (Interactive)**: I want to ask erm erm they only go to the cinema's and McDonald's do they do anything else?
- **C (Interactive)**: and where exactly?

Band D test-takers, however, sometimes struggled with formulating accurate interrogative forms when attempting to solicit further information from the examiner, as exemplified below.

→**D(Interactive)**: hmm (.) you: (.) erm has (.) er has a take taxi?
E: I'm sorry?
→**D**: erm eh you are (.) eh have eh eh take taxi? or erm
E: no (.) no he will drive me there (.) in h-h- you know ho-h-he will come and get me and take me in his car

ISE III

Test-takers at all bands asked questions to the examiners in order to *elicit more information and their opinions*. However, as observed at the lower exam levels, test-takers at Bands A, B, and C appeared to have asked more relevant questions which delves further into the examiners' views and opinions than Band D test-takers. It is also notable that they are successfully doing so on conceptually more demanding topics that are likely to require more sophisticated lexis (e.g. 'rule (the world)' (6 K), 'radiation' (4 K)).

- **A (Discussion)**: could we say that in 100 years' time (.) women will rule the world? (1.0) (laughs)
- **B (Interactive)**: so (.) er (.) but do you think nuclear power as such is good for us (.) er er without the radiation?
- **C (Interactive)**: why are you thinking that? (.) I mean ebooks are very cheap and err (.) very err: (.) er: very easy going
- **D (Interactive)**: and do you like your life (.) or you prefer to live alone?

Conclusion

Exploring the usefulness and potential of language function analysis, this study reported in this chapter attempted to validate existing language function lists included in the test specifications of an operational, large-scale, task-based spoken examination.

To summarize the findings for RQ1, a number of functions listed in the ISE specifications were sufficiently observed in the actual performance by the test-takers. However, there were functions that were expected but not elicited (e.g. *expressing ability and inability* in ISE 0; *expressing reservations* in ISE III). For these functions, the function lists would need to be revised accordingly. Several functions were found to have not been expected but elicited in test-taker performance (e.g. *negotiating meaning* for ISE 0 and I, *asking for opinions* for ISE II and III). Among them, several functions were elicited across all the four exam levels, which supports Green (2012), who argued that many functions start to emerge at a lower level and can still be found at higher levels, as well as Trinity's stipulation that test-takers should be able to use the functions listed for the preceding ISE levels.

The observations for RQ2 focused on such functions. The analysis revealed the common findings across the four exam levels that, when *expressing opinions/preferences/feelings*, Bands A and B test-takers tended to elaborate more, using more sophisticated vocabulary and more accurate and varied structures. As Bands A or B mean a passing grade with distinction or merit, so the test-takers' performances are generally thought to be stronger. Similarly, for interactional functions, test-takers with Bands C and above *asked* more relevant *questions* and successfully *negotiated meaning* more effectively than Band D test-takers (who failed the exams). These findings again support Green (2012) that at higher levels, test-takers use the same functions with higher linguistic precision and more sophistication. These findings prove the appropriate grouping of the test-takers according to the bands, which further evidences the ISE's scoring validity.

It is worth noting that the use of more advanced functions and more precise and sophisticated language at the higher exam levels is related to not only the test-takers' proficiency levels but also the increased task demands. For example, the *topic presentation* task, given only at ISE III, involves a formal presentation (prepared beforehand by the test-taker) that requires precision in building an argument with justifications using relevant examples, leading to a larger amount of production consisting of more varied vocabulary (as presented in Section 'Preliminary analysis per task RQ1' above). Similarly, for the *conversation* task, the topic areas become more abstract as the exam levels go up (as shown in Appendix 1): e.g. learning a foreign language for ISE I, public figures for ISE II, and independence for ISE III. Even though test-takers may perform the same function of *giving an opinion*, the more abstract topics require the use of more advanced vocabulary.

Despite being small scale, this study analyzed the representative test-taker performances at each band at each ISE exam level and demonstrated the usefulness of the methodology that can be used in validating the function lists, which are important

part of the specifications in dialogic speaking tests. While the lists of language functions in operational spoken tests are often intuitively developed by test designers (O'Sullivan et al., 2002), empirical studies are necessary to verify (or modify) the intended language function list against the actual language elicited in the test in order to better reflect the construct measured. Once the function lists are empirically validated, they are useful to represent the nature of the speaking tasks, offering helpful information to language teachers and learners, as well as contributing to examiner training. As noted earlier, this study was part of a larger project that revised tasks and rating sales of the ISE examinations.

Starting the task development process from a speech act (or a function) has widely been practised in TBLA, involving dialogic speaking tasks. However, as stated earlier, the task design and rating scales tend to be task-specific and they are often not readily applicable to other task types and formats (e.g. Youn, 2018). With the use of a cross-sectional function list of O'Sullivan et al. (2002) and Trinty's intended function list (Trinity College London, 2009), this study offers the potential for producing comparable tasks and tests for not only in large-scale TBLA but also in TBLT, where a number of task administrations may be needed within an academic year or across different years, and where different sets of multiple tasks are used to target different levels of proficiency. That is, a teacher may be teaching multiple classes and years at the same time and are required to prepare and assess students on a number of occasions. Looking at the types of language function and how they are achieved by the students can help track their progress over the duration of the courses and beyond, across different types of task. The sample ISE task questions shown in Appendix 1 would be useful in grasping the prompts and topic areas that are considered suitable for students with different levels of proficiency.

By demonstrating that the same functions can be elicited by different task types and that how they are performed can be differentiated across different band scores, this study has presented useful features for constructing rating scales that are more generally applicable to performances on various tasks, rather than on just one task type or one speech act. The analysis of test-taker language in this study indeed not only offered empirical validation evidence for the test specifications and the assignment of the bands in the ISE exams, but also laid empirical foundation for the revision of the original holistic rating scales into analytic ones, together with other evidence from a series of trials of draft rating scales and examiner feedback.[6]

Appendix 1: Sample ISE Task Questions

Conversation Task

ISE 0 Let's talk about the subject of shopping. What way do you usually go shopping?

[6] For the new rating scales, see https://www.trinitycollege.com/qualifications/english-language/ISE/ISE-results-and-certificates/ISE-rating-scales.

ISE I Let's talk about learning a foreign language. I'm thinking of coming to live in Spain. If I want to speak Spanish really well, what do you think I need to do?

ISE II Let's talk about public figures. Tell me: is there anyone at the moment in the public eye who you particularly admire or who you think is a good role model?

ISE III I'd like to talk about independence. I read in the newspaper in the UK. A recent survey showed that in the UK young people aspire to be financially independent by the age of 23. How do you feel about that?

Interactive Task

ISE II My friends' children are not allowed to use computers or mobile phones. I've been trying to persuade her that she is wrong.

ISE III A lot of people say wisdom comes with age. But it seems to me that the opposite is often true.

Note. The questions were taken from the sample videos published for exam preparation on the ISE website:

ISE 0: https://www.trinitycollege.com/qualifications/english-language/ISE/ISE-Foundation-A2-resources/ISE-Foundation-A2-videos

ISE I: https://www.trinitycollege.com/qualifications/english-language/ISE/ISE-I-B1-resources/ISE-I-B1-videos

ISE II: https://www.trinitycollege.com/qualifications/english-language/ISE/ISE-II-B2-resources/ISE-II-B2-videos

ISE III: https://www.trinitycollege.com/qualifications/english-language/ISE/ISE-III-C1-resources/ISE-III-C1-videos

Appendix 2: Transcription Notation

Unfilled pauses or gaps	Periods of silence. Micro-pause (less than 0.2 s) is shown as (.); longer pauses appear as a time within parenthesis. For example, (0.5) represents five tenths of a second
Colon (:)	A lengthened sound or syllable; more colon prolong the stretch
Dash (–)	A cut off, usually a glottal stop
.hhh	Inhalation
Hhh	Exhalation
Hah, huh, heh	Laughter
(h)	Breathiness within a word

(continued)

(continued)

Punctuation	Intonation rather than a clausal structure; a full stop (.) is falling intonation, a question mark (?) is rising intonation, a comma is (,) is continuing intonation
Equal sign (=)	A latched utterance, no interval between utterances
Opening bracket ([)	Beginning of overlapping utterances
Percent signs (% %)	Quiet talk
Asterisks (* *)	Creaky voice
Empty parentheses ( )	Words within parentheses are doubtful or uncertain
Double parentheses (( ))	Non-vocal action, details of scene
Arrows (><)	The talks speeds up
Arrows (<>)	The talk slows down
Underlining	A word or sound is emphasized
Psk	A lip smack
Tch	A tongue click
Arrow (→)	A feature of interest to the analyst

References

Atkinson, J. M., & Heritage, J. (1984). *Structures of social action.* Cambridge University Press.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford University Press.

Bardovi-Harlig, K., & Hartford, B. S. (1993). Refining the DCT: Comparing open questionnaires and dialogue completion tasks. *Pragmatics and Language Learning, 4*, 143–165.

Blum-Kulka, S., & Olshtain, E. (1986). Too many words: Length of utterance and pragmatic failure. *Studies in Second Language Acquisition, 8*, 47–61.

Brindley, G. (1994). Task-centred assessment in language learning: The promise and challenges. In N. Bird, P. Falvey, A. Tsui, D. Allison, & A. McNeill (Eds.), *Language and learning: Papers presented at the Annual International Language in Education Conference* (pp. 73–94). Hong Kong Education Department.

Brooks, L. (2003). Converting an observation checklist for use with the IELTS speaking test. *Cambridge ESOL Research Notes, 11*, 20–21.

Bygate, M. (1987). *Speaking*. Oxford University Press.

Field, J. (2011). Cognitive validity. In L. Taylor (Ed.), *Examining speaking* (pp. 65–111). UCLES/Cambridge University Press.

Green, A. (2012). *Language functions revisited*. Cambridge University Press.

Houck, N., & Gass, S. M. (1996). Non-native refusal: A methodological perspective. In S. M. Gass & J. Neu (Eds.), *Speech acts across cultures: Challenges to communication in a second language* (pp. 45–64). Mouton de Gruyter.

Inoue, C. (2013). *Task equivalence in speaking tests.* Peter Lang.

Kuiken, F., & Vedder, I. (2018). Assessing functional adequacy of L2 performance in a task-based approach. In N. Taguchi & Y. Kim (Eds.), *Task-based approaches to teaching and assessing pragmatics* (pp. 266–285). John Benjamins.

Nakatsuhara, F., & Dunlea, J. (2020). Applying the socio-cognitive framework: Gathering validity evidence during the development of a speaking test. In L. Taylor & N. Saville (Eds.), *Lessons and legacy: A tribute to Professor Cyril J Weir (1950–2018)* (pp. 132–158). UCLES/Cambridge University Press.

Nakatsuhara, F., Inoue, C., Berry, V., & Galaczi, E. (2017). Exploring the use of VC technology in the assessment of spoken language: A mixed-methods study. *Language Assessment Quarterly, 14*(1), 1–18.

Nakatsuhara, F., Inoue, C., & Khabbazbashi, N. (2021). Measuring L2 speaking. In P. Winke & T. Brunfaut (Eds.), *The Routledge handbook of second language acquisition and language testing* (pp. 285–294). Routledge.

O'Sullivan, B., Weir, C., & Saville, N. (2002). Using observation checklists to validate speaking-test tasks. *Language Testing, 19*(1), 33–56.

O'Sullivan, B., Taylor, C., & Wall, D. (2011). *Establishing evidence of construct: A case study*. Paper presented at the 8th annual EALTA conference, Siena, Italy.

Papageorgiou, S. (2007). *Relating the Trinity College London GESE and ISE examinations to the Common European Framework of Reference.* https://www.trinitycollege.com/about-us/recognition/english-language/cefr-alignment

Rose, K. R. (2000). An exploratory cross-sectional study of interlanguage pragmatic development. *Studies in Second Language Acquisition, 22*(1), 27–67.

Taylor, L. (2011). Introduction. In L. Taylor (Ed.), *Examining speaking: Research and practice in assessing second language speaking* (pp. 1–35). UCLES/Cambridge University Press.

Timpe-Laughlin, V. (2018). Pragmatics in task-based language assessment: Opportunities and challenges. In N. Taguchi & Y. Kim (Eds.), *Task-based approaches to teaching and assessing pragmatics* (pp. 287–304). John Benjamins.

Trinity College London. (2009). *Integrated Skills in English (ISE) examinations Syllabus—from 1 February 2010*. Trinity College London.

Trinity College London. (2013). *Exam information: Integrated Skills in English (ISE).* Trinity College London.

WebLingua. (n.d.). *Text Inspector*. https://textinspector.com/

Weir, C. J. (1993). *Understanding and developing language tests.* Prentice Hall.

Weir, C. J. (2005). *Language testing and validation: An evidence-based approach.* Palgrave Macmillan.

Weir, C. J., & Wu, R. W. (2006). Establishing test form and individual task comparability: A case study of a semi-direct speaking test. *Language Testing, 23*(2), 167–197.

Wigglesworth, G., & Frost, K. (2017). Task and performance-based assessment. In E. Shohamy, I. Or, & S. May (Eds.), *Language testing and assessment: Encyclopedia of language and education* (3rd ed., pp. 1–13). Springer.

Youn, S. J. (2018). Task design and validity evidence for assessment of L2 pragmatics in interaction. In N. Taguchi & Y. Kim (Eds.), *Task-based approaches to teaching and assessing pragmatics* (pp. 217–246). John Benjamins.

Zhou, Y., Dunlea, J., Negishi, M., & Yoshitomi, A. (2018). *Localisation of an international speaking test for Japanese university admission.* Oral presentation at the 6th British Council New Directions in English Language Assessment Conference, Kuala Lumpur, Malaysia.

Chapter 13
Communicative Strategies as a Tool for Assessing Spoken Interactional Competence

Pankaj Narke

Abstract Although there are several methods of assessing speech performance of ESL learners, assessing their interactional competence is a relatively lesser explored area. Interactional speech tasks require learners to consistently 'modify' their output to fit the dynamic context of the conversation, thereby creating long pauses, conversational breakdowns and speech overlap. Consequently, the quality of such performances cannot be evaluated by temporal features or grammatical accuracy factors alone. Learners' ability to negotiate for meaning, decision-making (choice of language and content) and ability to modify their output by considering the dynamic nature of the task play a vital role in the successful completion of the task. This study explores the use of communicative strategies (CSs) as a tool to assess interactional speech performance of English as a second language (ESL) learners on a series of group discussion tasks by adopting task-based language assessment approach (TBLA). Dornyei and Scott's (Language Learning 47:173–210, 1997) taxonomy of CSs was employed to identify the different CSs in the data, which were further categorized into basic (type 1) and advanced (type 2) strategies. Based on a mixed method of analysis, learners were found to be using type 1 CSs such as message abandonment, use of meaningless words as fillers, self-correction at syntactic level and rudimentary use of interpretive summary in the early stages of the study. As the study progressed, growth in speech performance was seen as a result of increase in the use of type 2 CSs such as use of restructuring, use of meaningful expressions as fillers, self-correction at sematic level and improved uses of interpretive summary. Learners' choices of CSs to solve different communicative problems during the interactional tasks indicated their increased level of interactional competence. The findings have implications for developing interactional competence in ESL learners by training them to use a range of communication strategies.

Keywords Task based language assessment · Interactional competence · Group discussion tasks · Communicative strategies · Interactional speech analysis

P. Narke (✉)
Department of Languages, VIT-AP University, Amaravati, India

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_13

Introduction

Proficiency of second-language learners in the target language can be affected by two major variables—their linguistic knowledge and the ability to use the knowledge appropriately across several communicative contexts. Among the multiple approaches adopted to determine learners' language proficiency, one of the most popular was skills-and-elements approach proposed by Lado (1960) and Carroll (1961), which advocated that learners' proficiency in the target language can be determined by measuring their ability in each skill—in this context, speaking skill. Tasks often preferred for assessing speaking skills are either unidirectional or bidirectional. In unidirectional tasks, there is only one speaker who speaks for a particular time period on a predetermined topic. Bidirectional tasks involve two speakers who take turns in a systematic way and cooperate with each other to complete the task. One of the most popularly used bidirectional tasks is interview, where the interviewer (examiner) asks questions and prompts learners if required. The spoken data produced through an interview-based task is usually linear and standardized. In other words, such data does not have unexpected interruptions, overlaps, awkward pauses and conversational breakdowns. Even if there are long pauses, it is easier for the examiner to attribute them to a particular speaker. Therefore, researchers and assessors often use *fluency*, *accuracy* and *complexity* as static measures for assessing speech performances.

On the other hand, researchers in the field of language education have found conversational tasks more useful for their language learning potential (Courtney, 1996; James, 1994). In interactional tasks, there are more than two participants who speak spontaneously. These tasks require learners to 'negotiate' for meaning to arrive at a commonly acceptable conclusion, mostly by finding a solution to the problem. As such tasks simulate authentic situations, the speech produced during these tasks is often featured with overlaps, interruptions, repetitions and false starts. Participants are required to make quick decisions to overcome these problems to complete the task. During conversations as learners are required to process a large chunk of information, their speech tends to become unstructured. It becomes difficult to assess such speech with the static measures that are applied for any systematic, structured and linear speech data (*contra* interview-based speech data). TBLA approach, as asserted by researchers in the field of language acquisition (Kuiken & Vedder, 2018; Youn, 2020), enables to elicit learners' oral data in real-life conditions and assess by applying appropriate analysis methods, such as conversation analysis. To serve this purpose, interactional tasks like group discussions and conversations can be used as they provide opportunities for learners to produce natural language to communicate with an authentic audience by 'negotiating' for meaning. Hence, it is of utmost importance to define learners' or test-takers' dynamic ability to perform in natural or natural-like speech situations.

So to understand how to measure and/or assess ESL/EFL learners' proficiency in natural speech situations, we now look at a brief overview of existing models of

communication strategies as propounded by researchers from the cognitive paradigm of language learning over the last five decades.

One of the primary approaches to assess interactional speech is to apply Canale and Swain's (1980) seminal model of *strategic competence* for defining spoken language proficiency. The researchers defined strategic competence as the ability of using communicative strategies (CSs) to compensate for communication breakdowns caused by performance variables or insufficient linguistics resources. Strategic competence is also viewed as the problem-solving ability during an ongoing communication. Researchers have considered the use of CSs as a feature of ESL/EFL learners' language production since long. For instance, Tarone (1977) conducted a study with ESL learners by administering picture description tasks. She analysed the verbal data qualitatively and identified five types of CSs—*paraphrase, transfer, appeal for assistance, mime* and *avoidance*. Further, Færch and Kasper (1983) identified the same set of CSs and classified them into two categories—*achievement* and *reduction* strategies. The former refers to the act of achieving the communicative goals by utilizing the existing resources or urging for assistance, whereas the latter means modifying communicative goals to avoid communicative problems.

Poulisse (1993) adopted a psycholinguistic approach to describe the use of CSs at three different stages—*conceptualization, formulation* and *articulation*—based on Levelt's (1989) speaking model. Levelt described the natural spontaneous language production of adults at three stages. His model distinguishes between declarative knowledge (conceptual and lexical) and procedural knowledge (application) about language. Poulisse restated the stages of Levelt's speech production model in the following manner: during *conceptualization*, the speaker gathers the content of the message based on the context and worldview. At the *formulation* stage, the speakers choose the appropriate lexical items and syntactic structures that would deliver the message effectively. Finally, the phonetic execution of the message happens in the *articulation* stage.

Dornyei's (1995) taxonomy of CSs is considered to be more inclusive and consistently referred to in modern research on second-language acquisition. He redefined the CSs by adding the interactive notion to it. From his perspective, the role of CSs is to promote mutual understanding among speakers. Dornyei classified some CSs as indirect strategies which do not necessarily serve the purpose of problem-solving, but enhance the effectiveness of the message. Nakatani's (2006) is one of the more recent and popular research on CSs, in which she viewed the functions of CSs beyond the skill of problem-solving. Unlike the previous researchers, Nakatani looked at the use of CSs from the productive as well as receptive side by adding non-verbal behaviour of the speakers.

Experts in the domain of language testing and assessment have proposed the communicative language ability (CLA) model of language assessment (Bachman, 1990; Bachman & Palmer, 1996; Fulcher, 2003; Purpura, 2013). The proponents of CLA believe in two major components of language proficiency: *language competence* and *strategic competence*. While Bachman (1990) and Bachman and Palmer (1996) conceptualized strategic competence into different fragments, such as planning, execution and assessment, Fulcher (2003) proposed the term *strategic capacity*,

referring to the use of strategies to solve communicative problems. His model was specifically focused on speaking ability that included language competence, textual knowledge, pragmatic knowledge and sociolinguistic knowledge. Thus, the CLA proponents have paid due attention to the significance of strategic competence in defining language ability of a speaker.

Communicative Strategies and Second Language Acquisition

There are two widely differing perspectives on the use of communicative strategies (CSs) in learners' language production. Skehan (1998) believed that learners use CSs at the cost of their linguistic competence. In his view, learners compromise with the linguistic challenges that they face while speaking in natural contexts. They tend to use certain CSs that help them bypass or avoid these linguistic challenges. This process of using CSs restricts the development of their linguistic competence.

A differential view by Kasper and Kellerman (1997) suggests that CSs play a major role in learners' target language development. According to them, firstly since learners use CSs to maintain the flow of their conversation, it provides them exposure to target language input as well as gives them the impetus for output. Secondly, it provides learners opportunities for problem solving and using the target language in interactive ways as they use CSs for meaning negotiation. Learners often use their existing language knowledge to employ different CSs, and therefore, they gain mastery over it. Thirdly, when learners use different CSs to solve conversational problems, they get positive feedback that helps them gain confidence. Lastly, learners use CSs when they are obligated to participate in a conversation. This pushes them to produce the output leading to a learning experience (Swain, 1985).

Canale (1983) categorized CSs into two types: in the first type, there are strategies that are generally used by learners as problem-solving devices during communication disruption, whereas in the second type, there are strategies that learners use to enhance the effectiveness of communication with their interlocutors.

The overall description of the utility of CSs in the process of language learning advocates that the use of CSs is a defining factor in second-language acquisition. They can also be used as descriptors for ascertaining the level of speech competency of language users.

Communicative Strategies as a Construct for Testing Interactional Speech Competence

Researchers in ESL/EFL contexts have considered *strategic ability* as one of the measures to decide growth in language learners' interactive oral proficiency (Canale & Swain, 1980). The traditionally held measures of fluency, accuracy and complexity

do not provide a comprehensive picture of learners' oral proficiency for they are static constructs and therefore difficult use as measures to apply interactional speech, which is a far more dynamic construct. In this respect, strategic competence is one of the most suitable measures to determine growth in learners' interactive oral proficiency.

Hasselgreen (2004) defines strategic ability as the ability to maintain the flow of conversation in threatening and conversational breakdown conditions. He identified three such conversational breakdown conditions: (1) when the speaker is confused or does not know which structure or words/phrases to use to express the intended meaning in a conversation, (2) when the speaker says something and means something else and (3) when the speaker says something but the interlocutor understands it differently. To handle such difficult conversational situations, speakers switch to their zone of knowledge that they are familiar with and tide over the conversational gaps using CSs. Dornyei and Scott (1997) also identified three major reasons for which learners use CSs: firstly, they do it when they find problems with their own language performance. When learners realize their mistakes during a conversation, they use CSs (e.g. *self-repair* or *comprehension check*) to compensate. Secondly, learners use CSs when they find problems with other's performance. When learners find the interlocutors incorrect or unclear, they use CSs like *clarification check* or *asking for repetition.* The third instance is when learners experience processing time pressure. In a conversation when learners are expected to respond instantly (mostly while in conversation with more proficient interlocutors) but they cannot produce what they want to, they use certain CSs such as the *use of fillers* or *paraphrasing*.

Therefore, from the discussion above it can be understood that CSs constitute a major linguistic device that learners use when they encounter a conversational problem. The context in which they use CSs, as described above, is mostly conversational and interactive. In the present study, an attempt will be made to understand which CSs do Indian ESL learners use to complete an interactive task like group discussion. Interactional competence is of crucial importance for adult Indian ESL learners as it is an important component of recruitment test that usually takes place in their final year of graduation. While summative assessments that the adult learners go through have a limited share of oral proficiency, their real-life professional needs necessitate the development of interactional competence. According to Young (2000), interactional competence is rarely highlighted though it is one of the most desired employability skills. To address this gap, it is important to raise teacher awareness to teach and assess interactional competence of ESL learners using suitable tasks and appropriate assessment measures. This is what the present paper aims to achieve by reporting a study on measuring the interactive spoken proficiency of a group of adult Indian ESL learners.

In the study reported in this paper, the communicative strategies classified in the Dornyei and Scott's taxonomy (1997) are the framework chosen to understand young adult ESL learners' interactional competence in a communicative task like group discussion. This taxonomy is used as this is an inclusive and well-classified taxonomy on CSs available till date.

The Study

Considering the vital role of CSs in defining spoken proficiency of ESL learners, this study was conducted to answer two research questions:

RQ1: *Do ESL learners use different types of communication strategies to complete group discussion tasks?*

RQ2: *Do the use of CSs show a positive change in use over time in ESL learners' speech performance in the group discussion tasks?*

Participants

Fourteen junior 'Computer Science and Engineering' learners (ten male; four female) with a mean age of 19 years participated in the study. All the learners were studying at Aurora's Technological and Research Institution in Hyderabad, India. All the participants had a minimum of 10 years of formal exposure to English language. The selection of the participants was based on their performance on a baseline proficiency test administered to understand the participants' overall proficiency in English language. The test was adopted from the course of English proficiency programme designed at The English and Foreign Languages University, Hyderabad. The course had a multiskill approach with tasks to develop reading, writing, indirect speaking, grammar and vocabulary within B1 to B2 range of the Common European Framework of Reference (CEFR) descriptors (p. 5). The participants were found to be at B2 level based on their scores in the proficiency test. They had sufficient vocabulary and grammatical knowledge to understand expository texts and respond to the questions based on such texts. Similarly, in the indirect speaking test, made of discourse completion items, the learners could fill in appropriate sentences and expressions to complete the discourses meaningfully.

Tasks

Group discussion tasks with argumentative topics on contemporary social issues have been found appropriate to facilitate natural interaction among participants. As claimed by Brooks (2009), in group tasks learners are required to produce natural language and negotiate for meaning. Lee (2008) argued that in the process of performing on group tasks, learners employ a range of CSs to negotiate for meaning. According to Leaper and Riazi (2014) in group discussion tasks, learners produce more language through peer collaboration. Group tasks allow learners to use more 'open moves' (with responses related to the prompts or the topic of discussion) to make their output spontaneous. In her recent study, Youn (2020) used role-play interaction tasks to extract distinct sequential organizations and interactional features

as critical validity evidence for assessing interactional competence of ESL learners across proficiency levels. In this study, she explored conversation analysis' feature of indicating the markers of interactional competence in terms of proposal sequence, signalling the shift between actions, responding to the proposal and closing the proposal sequence (p. 93).

In the present study, group discussion tasks were used to see if learners use different CSs to complete the tasks and if there was any development in their use of CSs as their experience in the group discussion tasks increased over time. To ensure sufficient exposure to group discussion tasks, six rounds of the tasks were conducted with a gap of 11–12 days between every two rounds. Each group discussion task was given a time limit of ten minutes which encouraged the participants to participate in the task and make their contribution, as directly relevant to the topic of discussion. Learners were given the topics for discussion in advance and some reference materials to read/watch to help them gather content knowledge prior to the group tasks. Each group discussion was preceded by an online synchronous discussion (OSD) on a small-scale topic that were thematically similar to group discussion topic, intended to activate learners' content and linguistic schema about the given topic. After every round of group discussions, the video recordings were uploaded on the Facebook page for the participants to view. Learners were asked to reflect on their performance and encouraged to give peer feedback. The researcher also provided feedback on technical aspects of participants' performances such as *voice audibility*, *long pauses*, *overlaps* and *turn taking*. Some participants were also given individual feedback on their concerns such as anxiety, control and body language (Table 13.1).

Every face-to-face group discussion was for a period of 10 min with 4–5 participants in each group. However, the online discussions were not timed as there were pre-speaking tasks done outside the classroom.

Table 13.1 Topics for online synchronous discussions and face-to-face group discussions

Round	Online synchronous discussion topics	Face-to-Face discussion topics
1	For marriage systems what are your views?	Which one is better—love marriage or arranged marriage?
2	Is practical education more important than theoretical education?	Does Indian education system fail to provide employability skills?
3	Will gender-based reservation help in achieving women empowerment in India?	Is gender equality just a dream in India?
4	Are private hospitals better than government hospitals?	Does the private sector assure more quality services than the government sector?
5	Is living in the urban area better than living in the rural area?	Should Indian government focus only on rural areas for the development of the country?
6	Is reservation necessary?	Reservation brings social equality - truth or myth?

Method of Data Analysis

Participants' performances during the group discussion tasks were video recorded and later transcribed to count for the frequency of occurrences of each type of CS use per participant across six rounds. The compilation of three CSs taxonomies constructed by Dornyei and Scott (1997) was used to do the frequency analysis of the presence of CS types in participants' speech during the six rounds of group discussion performances. The data was analysed to identify the frequency of types of use of CSs, growth in the use of CSs and the link between use of CS and the quality of interactional speech performance. To protect learners' identity, pseudonyms, such as MK, KLM and SHNTI, were used. A mixed method of analysis was attempted as is presented in the section on the findings.

The Findings

An in-depth analysis of the transcripts showed that learners used a range of communicative strategies to negotiate for meaning vis-à-vis knowledge co-construction. While categorizing the CSs for a descriptive analysis, Canale's (1983) definition of two types of CSs was used to categorize the sub-types of CSs chosen from Dornyei and Scott's taxonomy (1997): **Type 1** CSs are those which learners use to *solve* communicative problems, whereas **Type 2** CSs are used to *increase the effectiveness* of the message. This classification was done to understand the frequency of use of CSs at two levels—basic and advanced—of interactional speech performance and to understand if their frequency of use changes over time. In the following section, the findings of the study are presented to answer the two research questions through a frequency-based quantitative analysis of the speech data.

RQ1: Do ESL learners use different types of communication strategies to complete group discussion tasks?

Tables 13.2 and 13.3 show the overall presence of type 1 and type 2 CSs during the six rounds of GD across all the participants.

The first 14 CSs in Tables 13.2 and 13.3 are from Dornyei and Scott's (1997) taxonomy; however, the last four CSs have emerged from this study—*Reorganizing, Predictive clarification, Self-referencing*, and *Cross-referencing.* Overall, the use of a range of 18 sub-types of CSs was found across all the 14 participants. This indicates that they found these strategies for communication useful. Thus, the occurrence of 18 sub-types of CSs in Tables 13.2 and 13.3 proves RQ. One to be true that *ESL learners can use both varieties of CSs to solve interactional challenges.* These form the base with which their interactional competence to complete group discussion tasks can be understood systematically.

In Fig. 13.1, a per cent frequency of each strategy across the six rounds of group discussion tasks is given. The pie chart suggests that participants used the CSs like

Table 13.2 Frequency of Type 1 CSs used, definition and examples from the data

No	Strategy	Definition	Example	Mean frequency
1	Restructuring	Speakers abandon their original speech plan as they face linguistic or content-related difficulties. They prefer to restart in cases of a false start	If you do not … there will not be. I think there will be no possibility of developing something (VIJ, GD2, topic—Indian education system fails to provide employability skills: fact or fiction?)	3.8
2	Asking for clarification	Asking for explanation in case of unfamiliar language or content/idea used	What do you mean by…? Can you explain? (KLM, GD 3, topic—gender equality, a reality or dream?)	1.5
3	Messages abandonment	Leaving the message/turn incomplete due to language difficulty	Education system has to be changed because we have already … subjects all aa…(MK, GD2, topic—Indian education system fails to provide employability skills: fact or fiction?)	5
4	Use of fillers	Using routine words/ phrases in order to fill the pause that causes due to insufficient linguistic or content-related resources. Pauses create a space for cognitive processing	I mean, you see, like, okay, yeah (MK, GD 4, Topic Which sector assures better quality service: private or public?)	1.66
5	Self-correction	Correcting oneself after recognizing the linguistic or semantic error in the output	First women has to… women need to be… (RAJ, GD3, topic—gender equality, a reality or dream?)	5.66

(continued)

Table 13.2 (continued)

No	Strategy	Definition	Example	Mean frequency
6	Other correction	Correcting the interlocutor for making mistakes (grammatical or semantic) in speech	MK: I would like to make a comment here, as you have said rural areas should be developed more than urban areas// RAJ: no, giving importance to that (KLM, GD 5, for development of India which areas should government prioritize: rural or urban?)	1.4
7	Confirmation check	Confirming the understanding of the interlocutor's output in order to avoid misunderstanding of the point	Did you say? …You mean? (RAJ, GD 3, topic—gender equality, a reality or dream?)	1.66
8	Approximation	Using a similar word to replace the target word in case of not knowing the exact target word or not remembering it	..It is very good though the payment is quite high in Air India (instead of charges or ticket prices) (MK, GD 4, topic Which sector assures better quality service: private or public?)	1.25
9	Self-repetition	The speaker repeats what was said in order to cope with an unexpected pause or to emphasize on a particular point	Here the mediators…mediators … who are the mediators? (MNK, GD2, topic—Indian education system fails to provide employability skills: fact or fiction?)	1.4
10	Paraphrasing	Describing a term/object/action in alternative words in case of inadequate linguistic resources	Nowadays, we can see many of the girls, they are killing because of girls… (instead of female foeticide) (VIJ, GD 3, topic—gender equality, a reality or dream?)	1.5
11	Word coinage	Creating non-existent words in L2 by applying L2 rules	Agriculturing instead of farming (MK, GD 5, for development of India which areas should government prioritize: rural or urban?)	1

Table 13.3 Frequency of Type 2 CSs used, definition and examples from the data

No	Strategy	Definition	Example	Mean frequency
12	Interpretive summary	Summarizing what the other person has said earlier and then adding to it or putting the next comment. This strategy helps to contextualize the output. Also to gain a start	Summarizing the earlier person's comment	1.33
13	Exemplification	Using examples in order to elaborate a point. This is done in order to adjust one's output by considering the understanding level and response of the interlocutor	For example… Let's take an example… For instance…	4.25
14	Q uestioning	Asking questions to others in order to emphasize on a particular point. In some cases, speakers ask questions to themselves to get a good start	I want to raise one more point here that why this development is necessary in the rural areas? Because (MK, GD 5, for development of India which areas should government prioritize: rural or urban?)	1.66
15	Reorganizing	Reorganizing ideas in order to produce a coherent output. In such strategies, speakers abandon the initial idea and then include it in the latter part of the message	Like they can be on any… like that if one person is improved, he can help his family members to educate, to get jobs, for any other reasons, what I feel. (RAJ, GD 6, topic—reservation brings social equality—truth or myth?)	4.6

(continued)

Table 13.3 (continued)

No	Strategy	Definition	Example	Mean frequency
16	Predictive clarification	Providing clarification on the terms/ideas by predicting that the interlocutor will have problem in understanding that output. Speakers employ this strategy mostly when they are not sure about the use of the linguistic expression	The money drought occurs… as I say money drought occurs at the govt. side. See if we give money to the high caste people so money gets shifted (MK, GD 5, for development of India which areas should government prioritize: rural or urban?)	3.25
17	Self-referencing	Referring one's own speech in order to re-establish the conversational context and maintaining the stand	I would like to just clarify my point that I said that vegetables are becoming hygienic. (MNK, GD 5, for development of India which areas should government prioritize: rural or urban?)	1.8
18	Cross-referencing	Referring or the interlocutors' speech in order to: establish context, present a counterargument, take support to substantiate the argument	…as VIJ said that reservation because to uplift the lower caste people… (SRM, GD 6, reservation brings social equality—truth or myth?)	2.5

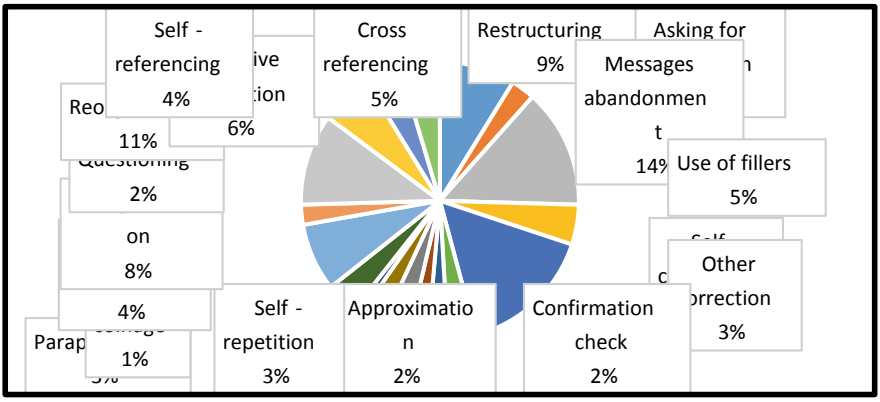


Fig. 13.1 CSs types across six rounds of group discussion tasks (in %)

message abandonment (14%), reorganizing (11%) quite dominantly throughout the study. The use of these strategies clearly suggests that the participants encountered conversational problems (*message abandonment*) and were able to solve them by using CS like *reorganizing* the output.

The use of CSs like *paraphrasing (3%), reorganizing message (11%), restructuring(9%)* and *use of fillers (5%)* highlights participants deliberate attempt to sustain the group discussion task despite having shortage of linguistic or content resources. Furthermore, the use of strategy of *self-correction (16%), predictive clarification (6%), exemplification (8%) and interpretive summary (4%)* suggests that the participants were aware of their output and the conversational contexts (Dornyei & Scott, 1997).

The presence of type 2 CSs is an evidence of participants' conscious efforts of making their output more effective (Canale & Swain, 1980). CSs like *asking for clarification (3%), confirmation check (2%)* and *questioning (2%)* are used for negotiation of meaning that leads to maintaining the flow of the discussion and keeping the track of the discussion in the context of the task (Lee, 2008). Although smaller in size, the use of CSs like *approximation (2%), cross-referencing (5%)* and *self-referencing (4%)* indicates that the participants were not only concerned about the quantity of their participation (in terms of total time) but also the relevance of it. Their awareness of the context and ability to modify their output qualifies the participants to have *interactional competence* as defined by Hall et al. (2011)—it is the speaker's ability to understand social context-specific communicative events, ability to deploy and recognize context-specific patterns and take consequent actions to repair the problems in maintaining the shared understanding of the interactional task, to be accomplished by a group.

RQ2: Does the use of CSs show a change in use across the six rounds of group discussion and thereby growth in ESL learners' speech performance in interactional tasks?

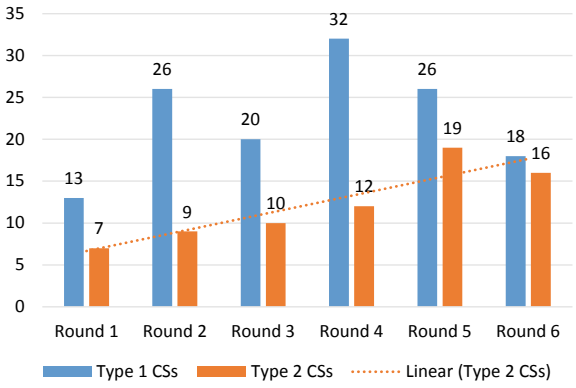


Fig. 13.2 Presence of CSs type 1 and type 2 across the six rounds of group discussions

We now look at an analysis of performance due to the use of two sub-types of CSs across six rounds of group discussions or over the entire duration of the study. Figure 13.2 shows the types of CSs used by all the participants for each round:

In Fig. 13.2, comparing the bars of Type 1 and Type 2, CSs indicate that the learners found a higher use of type 1 CSs over type 2 CSs during group discussion tasks. So they were more engaged to solve and tackle interaction-related issues, thus making their conversation more effective. An overall growth in the use of type 2 CSs from round 1 till round 6 is seen. This is interesting as it implies that the participants were able to *increase the effectiveness* of their output over time (Canale, 1983). The trend line of type 2 CSs thus suggests that the participants' concern for making their speech effective shifted from the basic language-related issues such as grammatical accuracy, oral fluency (type 1 CSs) to advanced communicative concerns like contextual appropriacy and intelligibility of the output (type 2 CSs). The increased use of type 2 CSs also underlines the improvement in learners' *interactional competence* and provides evidence for validating RQ two.

Tracking the Development in the Use of CSs: A Qualitative Analysis

Let us now look at a more micro-level into the speech data. This forms the qualitative analysis of the data to show growth in use of CS types and purposes over time. A participant-wise analysis showed that the use of CSs evolved as the study progressed. The following changes were observed in performance of the participants, which can be considered as indicators of development in their speech as there was a clear progression in the use of CSs from basic to advanced:

- Replacement of message abandonment with restructuring;
- Shift from CS use for grammatical problem-solving to semantic problem-solving;
- Change in the purpose of use of CS as interpretive summary;

- Change in the nature of CS use as fillers.

To begin with in the first three rounds, group discussions participants used the *message abandonment* strategy as a face-saving technique. In this situation, participants abandoned the entire message due to the lack of linguistic or content resources to express it. To avoid making a mistake, participants abandoned their turn.

For example,

Round 2. SHNT: Govt. should also provide the opportunities for the..aaa. (long pause) (GD 2, topic—Does Indian education system fail to provide employability skills?)
Next turn Round 2. SHNT: My suggestion is if you take any branch, like we are taking CSE we have, the subjects have to be included based on it…

In this example, the participant did not have the required linguistic resources, therefore she gave up her turn. This affected her overall participation time as well as the speech fluency (due to the long pause). However, in the latter round, she replaced her *message abandonment* with *restructuring*.

For example,

Round 5. SHNT: in the previous years… till today like (pause) we have communities in villages like sarpanch, so I think the youngster has to be. have to be given opportunity to further course so that they can develop their villages and we—total villages can be developed and we can be developed. (GD 5, topic: For development of India which areas should government prioritize: rural or urban?)

In this example, the participant had a false start due to the lack of linguistic resources. However, instead of abandoning her message, she restructured her output and continued her turn to convey her thoughts on the topic. This replacement of CS of *restructuring* improved her speech length as well as the fluency as her turn subsequently does not show any pause.

The trend lines in Fig. 13.3 suggest that once there is a gradual increase in the *restructuring* strategy, subsequently this is characterized by a decrease in *message abandonment*. This trend is seen as the group discussion rounds progress over time.

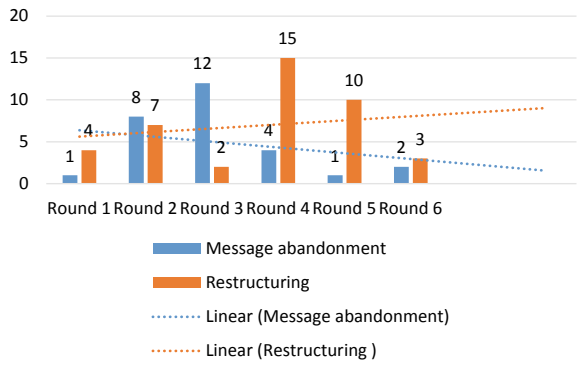


Fig. 13.3 Replacement of 'message abandonment' with 'restructuring' strategy over time

As discussed earlier, this replacement is a relevant indicator of positive growth in the participants' interactional speech competence.

Self-correction: From Grammatical Accuracy to Semantic Appropriacy

At the initial phase of the study, most participants are observed to employ the *self-correction* strategy to improve grammatical accuracy of their output. This was because their impression of being a good speaker was to be able to speak grammatically correct English. However, as the study progressed, it was observed that the participants got more comfortable with taking grammatical accuracy risks as their focus shifted to semantic appropriacy.

For example

Round 1. MK: so they have to start a new life, there is no backup to (pause) backup for them, no supports for them … (GD1, topic: What is better—Love marriage or arranged marriage?)
Round 3. MK: So reservations cannot be provided there. So I think reservations can be provided in some restricted areas…. (GD3, topic: Is gender equality just a dream in India?)

These excerpts from the group discussion represent the difference in the use of self-correction strategy at two different phases during the study. In round 1, MK self-corrected for being grammatically accurate in terms of appropriate use of prepositions, whereas, in round 3, he used the same strategy for correcting the meaning aspect of his output ('there' is replaced by 'in some restricted areas'). This suggests that the learner has shifted his focus from being grammatically accurate to being semantically acceptable.

Change in the Use of Interpretive Summary

The strategy of summarizing is normally used to highlight the context of the discussion and make one's argument relevant. Learner performance shows that they used the strategy of summarizing in two different ways through the study. In the beginning, they summarized other participants' speech to indicate that they have understood the point, and they would like to respond. Summarizing happened at the beginning of the turn. However, in the latter part of the study, summarizing happened towards the end of the turn (a *turn* is the amount of speech produced by a speaker in a conversation in a single take).

For example,

Round 2. KLM: no, as you said that it depends on the employment but I don't think that the stream require for the employment because if you do anything you should be good at that thing
Round 6. KLM: yes, I agree with you VMS, but the thing here is like you are giving the reservation like to a particular caste…(continued)
…But as you said, it is correct (pause) for the higher rank also they are not getting good colleges than the caste. So that is the disadvantage of that thing.

In this example, KLM referred to the earlier person twice, in the beginning and in the middle. In this example, he used *interpretive summarizing* strategy for two reasons: (a) to start his comment by contextualizing it and taking his stance in round 2 and (b) to present the counterview on his comment as expressed by another participant earlier and then making an inclusive conclusion of his turn in round 6. In the later instance, the advanced use of *interpretive summary* strategy by which the participant is making a counterargument is an indicator of his/her increased cognitive processing while speaking in an interactional task. The ability to employ summarizing skills for a range of functions to maintain speech fluency and express argumentation serves as an indicator of the speaker's increased interactional competence.

Change in the Nature of CSs: Use of Fillers

Fillers were often used when the participants required extra time to organize their speech or content during discussion tasks. These fillers helped them maintain their turns as well as speech fluency. At the initial stage of the study, they used certain words as fillers. However, the use of types of filler words changed towards the end stage. For instance, in the initial rounds, some learners used non-meaningful words as fillers.

For example,

Round 1. MK: so as KLM said, arrange marriages are good, it is a prediction. Aaa if aa couple is aaa doing a love marriage then aaa the parents may be anti.

In this speech sample, MK used '*aaa*' (phonetic transcription would be **/ɑː/**) as an expression for filler in order to maintain his turn and to give himself the time to think. However, a filler of this kind does not have a meaning of its own. However, the frequent use of such meaningless words makes the speech full of disturbance and may not sustain the listeners' attention.

In the latter rounds of discussions, the participants replaced these meaningless fillers with meaningful expressions which are normally used as fillers.

For example

Round 6. MK… As we today, most of the farmers are being (pause) I mean no one are interested in doing agriculture, …
KLM: The the one is given, the whole family is (pause) like will be quite improved.

VIJ: And the reservation when the (pause) actually I feel is, when the food is given to all the people, there won't be a necessity for some people to give more.

In these speech samples during the group discussion round 6, the participants used some formulaic chunks—'*I mean*', '*like*', '*actually I feel is*' as fillers instead of the non-meaningful phonetic expressions. The benefit of using these formulaic chunks is that the listeners find them meaningful, and they may not suffer from the loss of meaning. Simultaneously, as these chunks are meaningful, they also help learners in producing the relevant output. In the third example, VIJ faced a fluency issue due to the inadequate linguistic resources. He used the formulaic chunk '*actually I feel is*' and continued his speech in the same structure. Therefore, using formulaic chunks instead of non-meaningful expressions was an instance of growth in the use of fillers, adding to the speaker's interactional competence to maintain conversational fluency.

The micro-qualitative analysis of interactive speech data presented above as four features of speech maintenance—(i) restructuring messages, (ii) making semantic corrections, (iii) using interpretive summaries to place a counterargument and (iv) using lexical chunks as fillers—are indicators of improved interactional competence of the participants over time, thereby giving additional evidence of RQ two to be true.

Discussion: Implications for Assessment

The findings of the study attempt to answer both the research questions with evidence from the real-time performances of the 14 participants across the six rounds of group discussion tasks. We have observed that the participants have used 18 types of CSs, further classified into two categories: basic (type 1) and advanced (type 2), based on their functionalities to *solve* communication problems and *raise the effectiveness* of communication, respectively. The presence of type 1 CSs in learners' speech samples indicates that they are able to use the strategies to solve communication disruptions and convey their thoughts with increased fluency. So at such a stage, the primary concern is to be grammatically accurate and fluent in speech.

As their experience in interactive speech tasks increases, learners' comfort level with the task type also increases during the later rounds of the group tasks. Now they are able to handle the basic concerns, such as conversational breakdowns due the shortage of linguistic and content resources, with ease. So now their primary concern shifts to higher levels of thinking, and they strive to become contextually and semantically acceptable. Thus, the learners use advanced CSs to make their output effective by being contextual (use of summarization to express a counterargument) and more meaningful (lexical chunks as fillers, semantic appropriacy).

Lastly, identifying the presence of two categories of CSs in learners' interactive speech has helped us understand how to measure growth in interactional competence in adult ESL learners across time. It becomes evident from the findings that when ESL learners use CSs, they are guided by their linguistic resources at their disposal, while

the interactive challenges in the tasks lead them to use CSs to solve the problems at an advanced level. Thus, they are *pushed* to use a higher range of CSs to improve the communicative content of their oral output, a desirable employability skill that can be used for effective official communication.

Assessment Implications for ESL Learners

Interactional competence constitutes an important aspect of oral language assessment in ESL contexts. It is of crucial importance to closely analyse the naturally occurring discourse and social interaction which reveal the real standards of a speaker's ability of using language in particular contexts (McNamara, 1997). Furthermore, as implicated by Canale and Swain (1980), that observing learners' ability of using CSs (strategic competence) can enable the assessors to systematically evaluate learners' language ability. To achieve this goal, TBLA facilitates the suitable environment. Making use of interactive tasks, such as group discussions, helps assessors to provide a conducive environment for learners to produce speech in natural contexts (Lee, 2008). As seen in the present study, the use of group discussion tasks to encourage the production of natural interaction and the use of CSs are possible. Thereafter, the construct of CSs can be employed to assess learners' interactional competence add another layer to the oral proficiency assessment of ESL learners. For the reference of teachers–assessors, one entire group discussion data is presented as a sample with the communication strategy terms tagged onto the data as a separate column. This is provided in the Appendix 1. A similar template was used in this study for analysis. This template is included as this would help teachers understand the process of analysing interactional data to identify CSs and assess learners for their development of use of a variety of CSs according to task requirements.

The study has also underlined the potential of group discussion as an oral interactional task type and has demonstrated that it deserves an important place in oral language assessment. Regular exposure to group discussion tasks is likely to help learners develop their own mechanism for solving conversational problems. This study supports the group of researchers who believed that group discussions/group tasks allow learners to produce natural language and use CSs for a range of purposes (Brooks, 2009; Lee, 2008; Leaper & Riazi, 2014). Teacher can maintain individual learners' CSs use profile and observe the changes through formative assessment. While the use of type 1 CSs would suggest the conversational difficulties that the learners face during an interactional task and how they get over such problems, use of type 2 CSs will give a sense of learners' increased ability to raise speech effectiveness and thereby show a growth in interactional competence. Thus, interactional competence can be included as a measure to systematically determine learners' oral language proficiency within the scope of classroom-based formative assessment.

Conclusion

In this paper, we have explored the implications of TBLA approach in assessing interactional competence of adult ESL learners. Group discussions as interactional tasks serve to encourage the production of natural language among ESL learners. This paper supports the alternative method of measuring oral proficiency of adult ESL learners in the form of fine-tuned assessment of interactional competence. Use of CSs as a construct to assess interactional competence of a group of adult ESL learners forms the central idea of the paper. The paper advocates multiple rounds of assessment against the traditional product-based model of one-time oral assessment. While highlighting the importance of group tasks, the paper encourages second-language teachers–assessors to track learners' growth periodically and maintain a progress record. In this process, learners can be made more aware of their interactional competence and get feedback on the directions in which they can work on the construct to improve it.

Acknowledgements I would like to express my sincere appreciation to the participants of this study, who graciously agreed to be a part of it. My deepest gratitude goes to Dr. Kshema Jose who guided me during my PhD to understand the data better. Special thanks to Dr. Lina Mukhopadhyay who gave her valuable feedback to make this paper a better version of itself.

Appendix 1

Topic: Should Indian government focus only on rural areas for the development of the country?

Group Discussion Task _ Round 5_ Group 1 (participants = 5).

Note: The presented text here is the verbatim of participants' group discussion performance.

Total duration: 11.28 min.

Conversation exchange	Communicative strategy used
MNK: Good afternoon everyone, today's our topic is about is govt. sorry is our country should develop only rural areas more on rural areas. I agree with completely that rural area should be developed. Because even if the rural area is developed, no one will be … migrating will not be there… migrating will be not there. And they feel urban same as rural. So, they can live… where the situation… wherever they need. And one more thing is if rural is developed, obviously according to the situations, because why rural? Urban is already developed than rural so it is called as urban. So, if rural is developed I think the country will be developed. (pause)	Restructuring: topic is about is govt. sorry is our country should Reorganizing: no one will be … migrating will not be there… migrating will be not there
SHRVN: I agree with MNK. As she is saying that only the… as she is saying that the rural area should be developed, but one thing is that the rural area should be developed in the way they are affected to the nature, because we see that in the urban areas the people, they are like machines, they are working like machines. They are not affected to the nature. But the urban development should be … they are modernized also, the connection with the nature should also be there. Not like in the rural… in the urban area. (pause)	
KLM: yes, I agree with you SHRVN, but the development only can be done by destroying the nature. I think you cannot develop anything without disturbing the nature. If you want to develop something, you first need to move the thing… the natural thing from their original position and try something new in that. If you do not … there will not be… I think there will be no possibility of developing something. If you changing the nature, if you change something, that is called development. So, without disturbing the nature I think we cannot develop anything. So, is it possible to develop something without disturbing? As you said development should be there in the rural areas, then how is it possible to develop without destroying or destroying the nature?	Restructuring: If you do not … there will not be… I think there will be no possibility of developing something Questioning: So, is it possible to develop something without disturbing?

(continued)

(continued)

Conversation exchange	Communicative strategy used
VIJ: yes, KLM, what I feel is we can develop without disturbing the nature. See when we go back to some 50 to 60/70 years, aa beyond (prompt) … ago… education was under trees. Schools, hospitals, sanitation, so many things were under control of the nature. People try to modernize in such a way that they follow western culture and they call it modernization. When we follow our culture and use the techniques of the western culture, there will be easily… (interruption) development and we will never disturb the nature. It is like development is nothing… now-a-days development is nothing but constructing buildings, trying to make all the modern facilities, it is like following everything new. We… development is making lives… making sustainability should be … making sustainability)… like (fumble) development should make life better, but we are running backwards without noticing it. We thing that we are developing ourselves but somehow, we are becoming part of global warming. It is definitely just because of this development. We think development taking everything from aaa… when we go to rural areas and when we compare urban and rural, the temperature is very much different. Because so much of trees, so much of natural … everything is nature over there. Here everything is of manmade. We believe this as a development and we follow this. That is why we are being the victims of this natural disasters. Until the man doesn't stop destroying, there is never going to be end and development is never be ended. Every time we follow a new system, then we will be shifting to that. Every… now we are … we may not run behind the nature, but a time will definitely come that this urban will definitely go back to rural	Exemplification: See when we go back to some 50 to 60/70 years, aa beyond (prompt) … ago… education was under trees Repetition: development is making lives… making sustainability should be … making sustainability Restructuring: making sustainability should be … making sustainability) … like (fumble) development should make life better, but we are running backwards without noticing it Message abandonment: We think development taking everything from aaa… when we go to rural areas and when we compare urban and rural, the temperature is very much different

(continued)

(continued)

Conversation exchange	Communicative strategy used
VIJ: and this development is happening just because of the …aa what … the number of population increasing day by day. We are constructing a building… this apartment culture has come to this urban places only because of the population. The people … cannot… the independent house cannot accommodate so many people as apartment accommodation. That is why these people are encouraging this apartment culture. And even when the apartments are not aa even not being enough and independents are also not enough, this development is increasing day by day. Now look at this Hyderabad surroundings, they were not even just touched with this development before some years but now it is going on increasing in the surroundings of Hyderabad	Filler: of the …aa what … the number of population increasing day by day Message Abandonment: The people … cannot… the independent house cannot accommodate so many people as apartment accommodation
SHANT: in the previous years… till today like… we have communities in villages like sarpanch, so I think the youngster has to be… have to be given opportunity to further course so that they can develop their villages and we … total villages can be developed and we can be developed. So I think the youngsters has to be… have to be given opportunity to for their like sarpanch for these posts	Self-correction: so, I think the youngster has to be… have to be given opportunity to further course so that they can develop their villages and we…
VIJ: great point by SHRVN, I really appreciate it	
MNK: apart from all these, we should… (FUMBLE) why should we develop, we should be satisfied with whatever we are. If we go on developing, already needs will go in increasing. We should … each and everything will not be possible. First…	Reorganizing: apart from all these, we should… (FUMBLE) why should we develop, we should be satisfied with whatever we are

(continued)

(continued)

Conversation exchange	Communicative strategy used
VIJ: see, I would like to raise a point here, the most imp thing, when we compare with each other. Now we look at country to country wise. We say that we are developed by comparing to other country, or we talk to ourselves. When we look at ourselves, we always run behind something. Some or the other thing. Now, see now we will never be satisfied with what we have. We will somehow want to be developed. Somehow, we want to be… we don't even want to work for anything. Everything should be handmade and should come to our plates. This is what we feel and this is what … and this is what called as development is and when we stop comparing with others and … it should be definitely according to nature, and we should always in the path of it. Never destroy nature… Yes, it should come from within and we shouldn't be comparing to others…	Reorganizing: Somehow, we want to be… we don't even want to work for anything
KLM: yes, I agree with all of your points but I want to raise a one more point here that why this development is necessary in the rural areas? Because I think the rural areas now much developed so that they can do their works easily. See if we think different things, that mostly happen in the rural areas, agriculturing and I think agriculturing is the only thing that is the … done in the rural areas. If we …education is aaa… quite good for the people who live in the rural areas. the people who live in that area, educate the … themselves like how much they want. They can do and they can do agriculturing also… they focus on migrating to the urban areas and educate they who will provide the agriculturing? The people who live in the urban areas, can't bear the problems that occur in those areas. So … I don't think that it is a very much necessary for the development of those areas. They already use a lot of technology so I don't think further more development should be there … I agree with you and what I want to say is, development aa should not be done like industrialization and what you are saying is about agriculture. Rural areas only need agriculturing aaa development. No need for the development of industrialization. The only thing they should provide in the agriculture	Approximation: that mostly happen in the rural areas, agriculturing and I think agriculturing is the only thing that is the … Reorganizing: If we …education is aaa… quite good for the people who live in the rural areas

(continued)

(continued)

Conversation exchange	Communicative strategy used
VIJ: the rural areas, only need basic needs. They want electricity, water and agricultural facilities in sanitation, transportation. They want to sell their good to urban area. They should definitely come to … or these people want they will go there, but somewhat they should be linked up with other people and the basic needs should be like… the urban people they are like most beneficial people more than any other people. So, the rural people should be developed in such a way that they can only earn their own (prompt) development. See now for something they come to the urban area. If it is a handmade, not industrial, rural population can definitely work with handmade goods	Filler: should be like… the urban people
SHRVN: aa one thing I want to say is not only concentrating on the development of the rural area, the development of the urban area should also take place. Because if we go on developing the rural areas, the urban areas will be back… so there would be a connection between the urban people and the rural people. And the govt. is also taking some majors like digital India, where it wants everyone to connect with other people using the facilities like internet and everything. I think they are also developing the mobile app. For the farmers, so the urban people should help the rural people … m sorry … urban people should help the rural people in the development also, and they should also make their own developments	Self-correction: For the farmers, so the urban people should help the rural people … m sorry … urban people should help the rural people in the development also, and they should also make their own developments

References

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1996). Describing language ability: Language used in tests. *Language Testing in Practice* (pp. 61–79). Oxford University Press.

Brooks, L. (2009). Interacting in pairs in a test of oral proficiency: Co-constructing a better performance. *Language Testing, 26*(3), 341–366.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richard, & R. W. Schmidt (Eds.), Language and communication (pp. 2–14). Longman.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47.

Carroll, J. B. (1961). Fundamental considerations in testing for English language proficiency of foreign students. In H. B. Allen & R. N. Campbell (Eds.), *Teaching English as a second language: A book of readings* (2nd ed., pp. 313–321). McGraw-Hill.

Courtney, M. (1996). Talking to Learn: Selecting and using peer group oral tasks. *ELT Journal, 50*(4), 318–325.
Dörnyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly, 29*, 55–85.
Dörnyei, Z., & Scott, M. (1997). Communication strategies in a second language: Definitions and taxonomies. *Language Learning, 47*, 173–210.
Faerch, C., & Kasper, G. (1983). *Strategies in Interlanguage Communication.* Longman.
Fulcher, G. (2003). *Testing second language speaking*. Pearson Longman.
Hall, J. K., Hellerman, J., & Pekarek Doehler, S. (2011). *L2 interactional competence and development*. Multilingual Matters.
Hasselgreen, A. (2004). *Testing the spoken English of young Norwegians: A study of test validity and the role of "small words" in contributing to pupils' fluency* (p. 2004). Cambridge University Press.
James, G. (1994). Oral testing at tertiary level. In J. Boyle & P. Falvey (Eds.), *English language testing in Hong-Kong* (pp. 183–196). Chinese University Press.
Kasper, G., & Kellerman, E. (1997). *Communication strategies*. Longman.
Kuiken, F. & Vedder, I. (2018). Assessing functional adequacy of L2 performance in a task- based approach. In N. Taguchi & Y. Kim (Eds.), Task-based approaches to teaching and assessing pragmatics (pp. 266–285). John Benjamins.
Lado, R. (1960). *Language testing: The construction and use of foreign language tests.* Longman.
Lee, L. (2008). Enhancing learners' communication skills through synchronous electronic interaction and task-based instruction. *Foreign Language Annals, 35*(1), 16–24.
Leaper, D. A., & Riazi, M. (2014). The influence of prompt on group oral tests. *Language Testing, 31*(2), 177–204.
Levelt, W. (1989). *Speaking: From intention to articulation.* Bradford/MIT Press.
McNamara, T. (1997). "Interaction" in second language performance assessment: Whose Performance. *Applied Linguistics, 18*, 446–466.
Nakatani, Y. (2006). Developing an oral communication strategy inventory. *Modern Language Journal, 90*, 151–168.
Poulisse, N. (1993). A theoretical account of lexical communication strategies. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon.* John Benjamins; Levelt, W. (1989). *Speaking: From intention to articulation.* The MIT Press.
Purpura, J. E. (2013). Language learner strategies and styles. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (4th ed.). Heinle ELT.
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass, & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.
Tarone, E. (1977). Conscious communication strategies in interlanguage: A progress report. In H. D. Brown, C. A. Yorio, & R. C. Crymes (Eds.), *On TESOL '77* (pp. 194–203). TESOL.
Young, R. F. (2000). Interactional competence: Challenges for validity. Paper presented at a joint symposium on "Interdisciplinary Interfaces with Language Testing" held at the annual meeting of the American Association for Applied Linguistics and the Language Testing Research Colloquium, March 11, 2000, Vancouver, British Columbia, Canada.
Youn, S. J. (2020). Managing proposal sequences in role-play assessment: Validity evidence of interactional competence across levels. *Language Testing, 37*(1), 76–106.

Chapter 14
Effect of Task Structure and Interaction Conditions on Oral Performance

Sajit M. Mathews and N. P. Sudharshana

Abstract This paper investigates the influence of task complexity/structure and monologic versus dialogic interaction on the complexity, accuracy and fluency of 56 tertiary-level Indian ESL learners' speech in Skehan's limited attention capacity model. A narrative (NAR) and a personal information exchange (PIE) task of higher and lower task complexity/structure, respectively, are used. Robinson's concept of monologic versus dialogic interaction is used to predict interactive performance. It was hypothesized that speech complexity would increase with task complexity/structure while accuracy and fluency would decrease. Also, monologic interaction was expected to generate higher speech complexity than dialogic interaction. Task performances were recorded, transcribed, coded and analysed quantitatively using SPSS (version 23). The results show that NAR generated more complex and less accurate speech than PIE, and vice versa. Monologic interaction generated more complex speech than dialogues. These results confirm our hypotheses and show the operation of a complexity accuracy trade-off. However, Robinson's claim that complex tasks generate higher accuracy seems to be disproved. Regarding fluency, task complexity did not generate any consistent trend; however, PIE generated higher fluency than the NAR, but only in the dialogic condition. Task complexity as experienced by the test-taker/learner is heavily influenced by task-external implementation factors. Manipulation of task structure along with monologic versus dialogic interaction and complexity accuracy trade-off can be used to target-specific performance goals.

Keywords Task structure · Task complexity · Interaction conditions · Monologic and dialogic interaction · Oral performance · TBLT

S. M. Mathews (✉) · N. P. Sudharshana
Department of Humanities and Social Sciences, IIT Kanpur, Kanpur, India
e-mail: smmo@pm.me

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_14

Introduction

The centrality of tasks in current language pedagogy and assessment research is evident from an abundance of task-based research literature (Foster & Skehan, 1996; Long, 2015; Skehan & Foster, 1997; Vasylets et al., 2017). Tasks are increasingly being considered, both theoretically and practically, as effective tools in teaching and assessing a second language (see Ellis, 2009; Long, 2015; etc., for task-based language teaching and Brindley, 1994, 2000; Norris, 2002, 2016, Skehan, 1998, etc., for task-based language assessment). Regarding task-based assessment of oral proficiency, majority of studies till date have investigated the influence of task characteristics such as task structure and task complexity and implementation variables such as pre-task planning and interlocutor types on the complexity, accuracy and fluency of performance (e.g. Foster & Skehan, 1999; Skehan & Foster, 2001; Tavakoli & Foster, 2008).

This paper reports a part of an investigation of oral performance in Skehan's framework (Skehan, 1998; Skehan & Foster, 2001). The original study (Mathews, 2020) looked at how three independent variables, viz. task structure (realized as three kinds of tasks, viz. personal information exchange, narration and decision-making), pre-task planning time (3 min of pre-task planning compared with no-planning time condition) and interaction conditions (monologic vs. dialogic), influence oral performance in terms of three dependent variables, viz. complexity, accuracy and fluency of speech in Indian ESL contexts. In this paper, we report how task structure, implemented as two types of tasks, and nature of interaction, implemented as monologic versus dialogic performance, work in tandem in a complex manner affecting the independent variables when pre-task planning time is provided. We compare a cognitively more complex task ('Narration' or NAR) with a relatively simpler task ('Personal Information Exchange' or PIE) and show how these tasks resulted in a contrasting set of scores on complexity, accuracy and fluency parametres.

Previous Research

There are two main models regarding task-based oral performance assessment: Skehan's limited attentional capacity model (LACM) (Skehan, 1998; Skehan & Foster, 2001) and Robinson's Cognition Hypothesis (Robinson 2001a, 2001b). The LACM claims that human cognitive attention is limited in capacity; therefore, when cognitively complex tasks are being performed, second-language learner's attentional system will be forced to prioritize one of the performance goals over the others at a given time. In other words, during oral performance, L2 learners will focus on one of the aspects like complexity, accuacy or fluency, at the expense of others. The model argues that when performing complex tasks, learners generally focus on producing complex speech, and due to the limited availability of attentional resources, this may lead to lower levels of accuracy and fluency. The second model, cognition

hypothesis (Robinson 2001a), proposes a framework for determining the cognitive complexity of pedagogical tasks using three main variables, viz. task complexity (e.g. ± few elements), task conditions (e.g. one-way/two-way information split) and task difficulty (e.g. motivation).

Studies following the LACM, with the help of carefully designed tasks, intend to predict how learners prioritize attention and ultimately how learners score on complexity, accuracy and fluency. As previous research based on this model has shown, theoretically balanced task design and implementation have the potential to simultaneously promote complexity, accuracy and fluency (see Foster & Skehan, 1996; Skehan, 1996; Skehan & Foster, 1997). Task complexity or task structure is a widely discussed variable in this framework. Multiple studies have generated comparable results on how task complexity/structure influences speech as summarized below.

In a study that investigated how three task types and three planning time conditions influenced complexity, accuracy and fluency of performance, Foster and Skehan (1996) found that task types influenced fluency and complexity while there was a trade-off between complexity and accuracy. The three task types used were personal information exchange (PIE), narrative (NAR) and decision-making (DM) tasks which were performed under unplanned, non-detailed planning and detailed planning conditions. All tasks were performed dialogically in pairs by 32 pre-intermediate-level EFL learners of the age range 18–30. Tasks were designed in such a way that familiarity of information and predictability decreased from PIE to NAR to DM and therefore they were expected to have increasing levels of cognitive load in this order. By virtue of having a low cognitive demand and being familiar because of rehearsed information, PIE offered the possibility of greatest attention to form leading to higher accuracy scores. Since it required the use of simple linguistic forms, only a lower level of speech complexity was expected. NAR was conceptualized as cognitively more complex than PIE due to its visual-to-verbal information encoding requirement and the simultaneous necessity to use imagination to connect the given pictures. At the same time, the visuals, storyline and the plot of the story were expected to ease the cognitive burden to an extent. In comparison with PIE, the more complex NAR offered the scope for more complex language and therefore allowed only lesser attention to form and lower accuracy scores when compared to PIE. The third task DM, deemed cognitively the most demanding task, was a judgement task which required the participants to use and evaluate new information and formulate and defend their decisions/opinions. While unpredictable interactions during dialogic performance made it more complex, general familiarity with moral values may have eased its complexity. Such cognitive load warranted only minimal attention to form, but prompted the use of complex language. Summing up, the three tasks were arranged as shown in (1) below in terms of familiarity, cognitive load and overall difficulty:

(1)

Familiarity: PIE > DM > NAR;
Cognitive load: PIE < NAR < DM;
Overall difficulty: PIE < NAR < DM.

The mean scores showed that task complexity positively influenced speech complexity while it had an adverse effect on fluency. The simplest task PIE produced the most fluent speech; pauses increased from PIE to NAR to DM, while total silence time increased from PIE to DM to NAR. Both these measures were statistically significant for all three tasks, with high levels of significance for NAR and DM. Likewise, complexity scores—indexed as clauses/c-units—were also statistically significant for all task types. The most complex task DM achieved the highest complexity score, followed by NAR and PIE. NAR did not achieve statistical significance for complexity, while PIE and DM did. The interaction of planning and task complexity was analysed by contrasting the simplest task PIE with the more complex tasks, viz. NAR and DM. The trends were as given in (2) below:

(2)

Fluency: PIE > DM > NAR;
Complexity: PIE < DM < NAR;
Accuracy: mixed results, counter evidence.

Foster and Skehan (1996) conclude that task types were predisposed to certain kinds of performance. PIE promoted greater accuracy, but not complexity. NAR promoted greater complexity, but not accuracy. A trade-off between complexity and accuracy was observed. DM promoted useful levels of complexity and accuracy. Taking the advantages of such predispositions into consideration, it was therefore suggested that task types and planning time may be manipulated in order to promote complexity without sacrificing accuracy as a result of trade-off.

In a follow-up study, Skehan and Foster (1997) examined the influence of planning time and post-task activity on the complexity, accuracy and fluency of oral performance. Subjects were 40 EFL learners of pre-intermediate proficiency aged 18–25 from diverse L1 backgrounds. Three tasks, similar to the ones used in Foster and Skehan (1996), were performed dialogically in pairs in this study. PIE was deemed cognitively the least demanding task because it used familiar and probably rehearsed information; therefore, attentional resources would be available to focus on form. NAR required each subject to narrate a dialogue-less cartoon strip to the other in pairs. More cognitive effort was expected here than in PIE due to unfamiliarity of information and the need to weave a story out of the cartoon strips. Therefore, less attentional resources would be available to focus on form. DM—an agony-aunt task—required each pair of participants to decide on the best advice to be given to the problem mentioned in each of the given letters. This was deemed the most cognitively demanding task because it required the learners to comprehend, evaluate and use unfamiliar information and to decide on and defend the best advice possible. Two planning conditions and two post-task conditions were also used, which are not relevant to this study. Results showed that fluency indexed by number of pauses was the highest for NAR under both planning conditions, followed by DM and PIE. Complexity indexed as clauses per c-unit was the highest for DM, followed by PIE and NAR (where NAR was not statistically significant). Accuracy indexed as error-free clauses was the highest for DM (non-significant), followed by PIE and NAR. This

study showed that planning had significant impact on fluency and complexity while accuracy was less according to a trade-off effect. Another interesting observation in this study was how certain task characteristics tended to direct cognitive resources in certain directions. The authors suggested that tasks containing inherent structure tended to promote accuracy when planning time was provided. They suggested that the amount of transformations and on-line computation required and the complexity of outcome itself were positive influences on speech complexity.

In another study, Tavakoli and Skehan (2005) studied the influence of task structure, planning time and proficiency level on the complexity, accuracy and fluency of 80 EFL learners' speech during individual interviews with the first author of the paper. Four narrative tasks with increasing 'degrees of structure' were chosen with zero and five minutes of planning time. The proficiency and planning variables are not reported here as they are not relevant to this study. The hypotheses predicted that higher the degree of structure, greater would be the effect on fluency and accuracy of performance, and that there will be no effect on complexity. The results of the study showed that structured and unstructured tasks were significantly different from each other in case of fluency and accuracy of performance. Regarding complexity, unstructured tasks were found significantly different from one of the structured tasks used in the study, probably due to the effect of testing context. That is, though the results showed significant differences between structured and unstructured tasks, the concept of 'degree of structure' did not receive support.

In short, previous research suggests that the task structure has definite effects on speech. Speech complexity is generally found to be higher when task complexity is high. Accuracy has not given consistent results. The emergence of mixed results points to the need for further studies in this regard.

Our original study (Mathews, 2020) mainly follows the design framework of Foster and Skehan (1996), Skehan and Foster (1997), with an aim to investigate the effect of task structure on task-based speech performance in the Indian ESL context. This study also uses three kinds of tasks, viz. personal information exchange, narration and decision-making. In this paper, only the first two are reported. The second independent variable in the original study, namely interaction condition (monologic vs. dialogic interaction) is inspired by studies using Robinson's (e.g. 2001a, 2005) Cognition Hypothesis. Though interaction is studied in detail in the framework of Vygotsky's (1978) sociocultural theory of mind by scholars like Swain (see 2001, for example), not many studies comparing monologic and interactive interaction in task-based framework are available.

To make predictions about performance in relation to task complexity, Robinson (2001a, 2001b) proposed to distinguish between monologic and interactive tasks and explained how the interaction of task complexity and interaction variables influences performance. In the context of discussing the advantages of using increasingly complex tasks for effective second-language acquisition, Robinson (2001b, p. 34) argued that higher task complexity would lead to greater engagement of cognitive resources, greater attention to task input, greater incorporation of task input in working memory, greater attention to and modification of output and increased chances of noticing (Schmidt, 1990) and pushed output (Swain, 1995), leading to

enhanced interlanguage change. Robinson (2001a, 2001b) argued that when task complexity was higher, monologic tasks would generate more complex, more accurate and less fluent speech in comparison with simpler monologic tasks that tended to generate more fluent, less accurate and less complex speech. Likewise, for complex dialogic tasks, he predicted reduced speech complexity due to more negotiation of meaning, confirmation checks and clarification requests in comparison with simpler dialogic tasks.

According to Robinson (2001b), when complexity of monologic tasks is increased along variables like ± here and now, ± few elements and ± reasoning demands—which he calls resource-directing dimensions—attention will be directed to functional and linguistic language requirements. The reasons for these effects are threefold. Firstly, speech complexity develops in response to the complex demands of a task; secondly, greater accuracy results from greater communication demands and the resulting attention to speech, and thirdly, complex tasks result in greater cognitive efforts to produce language, stretch interlanguage and destabilize fossilized language (Robinson 2001b, p. 35). In contrast, when task complexity of monologic tasks is increased along ± planning time, ± prior knowledge and ± single task dimension—which he calls resource-dispersing dimensions—attentional resources are predicted to be depleted, resulting in reduced complexity, accuracy and fluency.

From this set of predictions, Robinson (2001b) extrapolates that when the complexity of dialogic tasks is increased, interaction would increase leading to greater negotiation for meaning, confirmation checks and clarification requests, which would reduce speech complexity considerably. This is because greater interaction in terms of shorter and elliptical turns would hinder the speakers' attempts at producing complex speech. However, accuracy is not expected to be affected by interaction.

In short, dialogic tasks are expected to generate less complex speech than monologic tasks, while accuracy and fluency may not be affected by interaction conditions. The interaction of task complexity with monologic and dialogic tasks is also predicted to have definite effects as we have seen in the review above.

Our Study

In the studies reviewed above, inherent task structure and interaction were found to be influential variables contributing to the differences in complexity, accuracy and fluency of performance. To investigate these effects, we started with two broad research questions:

1. How do task types with low and high task structure influence complexity, accuracy and fluency of speech?
2. How does monologic versus dialogic interaction influence complexity, accuracy and fluency of speech?

To investigate these questions, we formulated the following hypotheses.

1. The higher the task complexity, the higher the speech complexity will be. Therefore, PIE will generate lower complexity score than NAR.
2. The higher the task complexity, the lower the accuracy and fluency of speech will be. Therefore, PIE will score higher on accuracy and fluency than NAR.
3. Complexity scores in dialogic condition will be less than those in monologic condition.
4. There may not be any difference between accuracy and fluency scores of monologic and dialogic conditions.

Tasks

The two types of tasks reported in this paper are a personal information exchange task (PIE) and a narration task (NAR). They have different levels of task complexity as indexed by the predictability and familiarity of information and processes required, and the inherent structure and cognitive load of the tasks (Foster & Skehan, 1996). Task complexity is designed to increase from PIE to NAR. Both the tasks were trialled earlier with a comparable group, and necessary changes were made.

PIE requires the use of familiar information for task completion. The task used in this study is titled 'At Fresher's Party' in which the participants introduce themselves to a friend. Participants use information about themselves, their family, education and interests in this task. This is a 'required' information exchange task (Ellis, 2003) where task completion requires the exchange of information and elicits more consistent negotiation, than when the exchange is optional (Foster, 1998). Recall here that since PIE requires the use of familiar and probably rehearsed information, it is expected to be cognitively simpler (Skehan & Foster, 1997).

Narrative task is a commonly used tool in SLA studies for various reasons like the ease of manipulation and efficiency in assessing oral ability (see Foster & Skehan, 1996, 1999; Kawauchi, 2005; Pang & Skehan, 2014; Qian, 2014; Skehan & Foster, 1997, 1999; Skehan & Shum, 2014). NAR is cognitively more complex than PIE since it uses comparatively unfamiliar information and therefore tends to promote less attention to linguistic forms (Foster & Skehan, 1996: p. 307; Skehan & Foster, 1997: p. 193). It requires the use of background knowledge for task completion and produces greater complexity than simpler tasks (Skehan & Foster, 1997, 2008; Tavakoli & Skehan, 2005; Wang & Skehan, 2014). However, NAR is a non-negotiable task due to its input-based nature where the domination of the input provides only narrow scope for the speaker to modify or personalize the narration (Skehan, 2014). This makes NAR more structured than PIE. Moreover, the narrative structure of the task requires the speaker to adhere to the visual input and storyline for task completion. NAR used in this study required the participants to narrate the story in the given cartoon strip by Don Trachte which we titled 'Knock Knock'. The strip depicted a simple comical incident involving two boys.

Participants and Procedures of Data Collection[1]

The data for this study comes from 56 participants with an average age of 20 years, from 8 semi-urban degree colleges in the districts of Kottayam and Idukki, affiliated to Mahatma Gandhi University, Kerala, India, during February–August, 2018. Participants of the study were randomly selected from undergraduate batches of Commerce, Finance, Taxation, Cooperation, Marketing, Physics and English Literature. They came from middle class families with educated parents doing agricultural and professional jobs in public and private sectors. Within each institution, participants knew each other, but not across the participating institutions. Participants had scored either a 'B-Plus' or above on their Higher Secondary English paper and were judged as having lower intermediate to intermediate-level English language proficiency by their teachers.

Participants were instructed what to do orally as well as in written form prior to handing them the printed tasks. Each was given a pen and paper to be used for preparation if needed. Three minutes of planning time were given apart from 60 s of reading time, post which participants performed their tasks. For dialogic performance, participants were paired randomly. The oral responses were timed and recorded on a portable audio recorder as .mp3 files. Later, the participants filled a datasheet that collected their language-related information.

Coding, Transcription and Analysis

The recorded audio files were transcribed using a simple and accurate transcription convention adapted from Nitta and Nakatsuhara (2014). The transcribed data was then coded for specific measures of complexity, accuracy and fluency. To measure complexity, utterances were segmented into AS-units[2] and clauses. Non-clause units were included in the AS-unit count. For accuracy, specific syntactic, morphological and word order-related errors like erroneous tense forms of verbs, omitted/inappropriate auxiliary verbs, omitted verb or other major sentence components, erroneous subject–verb agreement, incorrect pronouns and prepositions, major inappropriate verb usage, incorrect word construction and glaring word order errors were marked. Repetitions of errors were counted as separate errors. For fluency, the durations of total speech time and total silence per participant were measured in milliseconds. All pauses longer than or equal to 1000 ms were marked and measured.

[1] The original study (Mathews, 2020) had 168 participants under two planning conditions (+ PL and – PL), two interaction conditions (monologic and dialogic) and three types of tasks (PIE, NAR and DM). This study reports only PIE and NAR task performance under + PL planning variable, performed under monologic and dialogic interaction conditions.

[2] AS-unit is defined as 'a single person's utterance consisting of an independent *clause or subclausal unit*, together with any *subordinate clause*(s) associated with either' (emphasis is original) (Foster et al., 2000, p. 365).

The total duration of performance, duration of examiner's utterances if any, total test-taker turn time (total length of each participant's turns), total silent time (each participant's silences in between utterances and between turns) and total length of utterances without silence were separately noted down in milliseconds, for each participant.

Complexity is defined as the ratio of the number of clauses to AS-units expressed as percentage and is calculated as given below:

$$\text{Complexity} = \left(\frac{\text{Number of Clauses}}{\text{Toal Number of AS Units}}\right)100 \tag{14.1}$$

Accuracy is defined as the ratio of the number of error-free AS-units to the total number of AS-units expressed as percentage and is calculated as:

$$\text{Accuracy} = \left(\frac{\text{Number of Error free AS Units}}{\text{Total Number of AS Units}}\right)100 \tag{14.2}$$

where the number of error-free AS-units were calculated by subtracting the number of AS-units with errors from the total number of AS-units.

Fluency is defined as the ratio of the total utterance/speech time to the total time taken by a test-taker (including silence) expressed as percentage and is calculated as follows:

$$\text{Fluency} = \left(\frac{\text{Total Utterance Time}}{\text{Total Test Taker Turn Time}}\right)100 \tag{14.3}$$

where total utterance time is the total time taken by a participant excluding silent pauses, and total test-taker turn time is time taken including silent pauses.

The coded data was then fed into SPSS for t-tests and multivariate analysis of variance (MANOVA). Results of the tests are discussed below.

Results

Table 14.1 presents the mean scores of three dependent variables. Figures in brackets indicate standard deviation.

Table 14.1 Mean values of dependent variables

	Complexity		Accuracy		Fluency	
	Monologic	Dialogic	Monologic	Dialogic	Monologic	Dialogic
PIE	116.45 (15.58)	82.32 (17.48)	66.32 (15.94)	84.26 (10.63)	79.45 (12.91)	92.31 (5.78)
NAR	134.71 (24.23)	127.87 (21.80)	47.59 (22.00)	48.79 (29.62)	82.31 (16.77)	71.71 (14.43)

From this table, it is clear that both the independent variables, viz. task structure and interaction conditions, had effects on dependent variables. A MANOVA showed that the differences in task structure and interaction conditions indeed had significant effects: Task structure—Pillai's trace = 0.582, $F(3,50) = 23.175$, $p = 0.000$; and interaction conditions—Pillai's trace = 0.264, $F(3,50) = 5.976$, $p = 0.001$.

We now discuss each of the three dependent variables in detail. Regarding complexity, recall here that Hypothesis 1 predicted that as task complexity increased, speech complexity would increase. The complexity scores from the table above confirm this hypothesis. The mean scores of NAR are higher than those of PIE under both monologic and dialogic interaction conditions (monologic: NAR = 134.71 and PIE = 116.45; dialogic: NAR = 127.87 and PIE = 82.32). Two independent samples t-tests were conducted separately to check if these differences are significant and the results showed that they are indeed significant: monologic—$t(26) = -2.372$, $p = 0.025$; and dialogic—$t(24.83) = -6.101$, $p = 0.000$. That is, the lower and higher task structure are significantly different from each other and affect speech complexity. Regarding interaction conditions, they had different effect on two different kinds of tasks. In case of the more complex NAR tasks, there was no significant difference between monologic and dialogic conditions: mean = 134.71 and 127.87, $t(26) = 0.785$, $p = 0.440$. In case of the less complex PIE task, there was significant difference between monologic and dialogic conditions: mean = 116.45 and 82.32; $t(26) = 5.455$, $p = 0.000$. These scores indicate that at higher task complexity, interaction conditions do not seem to exert significant influence on the complexity of performance, and task complexity/structure alone appears to be the most influencing factor.

Regarding accuracy, recall here that Hypothesis 2 predicted that with increasing task complexity, accuracy scores would decrease. The values in Table 14.1 confirm this hypothesis. Irrespective of the interaction conditions, the less complex PIE generated higher accuracy scores than NAR (monologic: PIE-66.32 and NAR-47.59; dialogic: PIE-84.26 and NAR-48.79). Two separate independent samples t-tests were conducted, and results showed that there was significant difference under both the interaction conditions: monologic—$t(26) = 2.581$, $p = 0.016$; dialogic—$t(26) = 4.218$, $p = 0.000$. We can see here that the effect is more pronounced under the dialogic condition. Looking at the interaction conditions, descriptive statistics in Table 14.1 shows that the two tasks behaved differently (PIE: monologic 66.32 and dialogic: 84.26; NAR: monologic 47.59 and dialogic: 48.79). The accuracy scores for PIE were higher under dialogic condition, whereas for NAR, the scores were higher under dialogic condition. Independent samples t-tests were conducted separately for PIE and NAR and results showed that for PIE there was significant difference between monologic and dialogic conditions: $t(26) = -3.502$, $p = 0.002$. In contrast, under NAR, there was no significant difference between interaction conditions: $t(26) = -0.122$, $p = 0.904$. This trend is similar to what was observed for complexity scores. The indication is that the influence of interaction conditions on accuracy is not significant in case of a complex task while in case of a simpler task, interaction conditions exert significant influence on accuracy. There is also a strong trade-off

between complexity and accuracy to accommodate the limited attentional resources available as seen in Table 14.1. Where accuracy for NAR is low (47.59 under monologic and 48.79 under dialogic), the corresponding complexity scores are higher (134.71 under monologic and 127.87 under dialogic). Since complexity is consistently high for NAR under monologic and dialogic conditions, the corresponding accuracy scores are lower due to trade-off effect.

Regarding fluency, recall here that Hypothesis 2 predicted that fluency would decrease with increasing task complexity. Table 14.1 indicates mixed results: fluency scores were significantly higher on PIE than NAR but only under dialogic condition: PIE = 92.31 and NAR = 71.71, $t(17.074) = 4.960$, $p = 0.000$. Under monologic condition, fluency scores on NAR were slightly higher than PIE, but the difference was not significant: NAR = 82.31 and PIE = 79.45, $t(26) = -0.505$, $p = 0.618$. This indicates a strong combined effect of interaction conditions and task complexity on fluency scores. Fluency scores under the two interaction conditions for two tasks differ significantly (PIE monologic: 79.45 and dialogic: 92.31; NAR monologic: 82.31 and dialogic: 71.71). To examine if these differences are significant, two independent samples t-tests were conducted separately. The results showed that for PIE, there was significant difference between interaction conditions: PIE: $t(26) = -3.401$, $p = 0.002$; but for NAR, there was no significant difference: NAR: $t(26) = 1.794$, $p = 0.084$. We can see here that these trends are again similar to what has been observed in case of complexity and accuracy. The difference between interaction conditions is more pronounced for the simpler task PIE, while for the more complex task NAR, the difference if at all significant is negligible.

Discussion

From the analysis in the previous section, the main findings can be summarized as given below:

- Complexity of performance
  - has a strong positive correlation with the task complexity: more complex task NAR scores higher than less complex PIE. This is in line with Skehan's predictions (e.g. Foster & Skehan, 1996; Skehan & Foster, 1997);
  - seem to be negatively affected by interaction conditions but only in case of PIE: monologic condition scored higher than dialogic condition. This is in line with Robinson's claims (Robinson, 2001b);
- Accuracy of performance
  - has a strong negative correlation with the task complexity: less complex task PIE scored higher than more complex NAR. This is in line with Skehan's predictions about a trade-off effect between complexity and accuracy (e.g. Foster & Skehan, 1996; Skehan & Foster, 1997);

  - seem to be affected by interaction conditions but only in case of PIE in a direction opposite to how NAR behaves: dialogic condition scored higher than monologic condition. This again indicates a trade-off between complexity and accuracy.

- Fluency of performance
  - does not show any consistent trend;
  - the less complex PIE scored higher than more complex NAR [as per Robinson's (2001b) claims] but only in dialogic condition; in monologic condition, the difference between them is not significant.

- Interaction conditions affect dependent variables in different ways:
  - complexity scores higher in monologic;
  - accuracy scores higher in dialogic mode;
  - regarding fluency, the results are mixed and indicate a strong effect of task structure.

These results confirm our Hypotheses 1 and 2 and Skehan's predictions that there will be a trade-off between complexity and accuracy. However, these seem to disprove Robinson's claim that more complex tasks will generate higher accuracy as a result of increased attention to form. Note here that the participants in this study were ESL learners of intermediate proficiency and it is clear that their attentional resources were hard-pressed during the performance in both monologic and dialogic conditions. Another interesting factor here is that the more complex task NAR resulted in higher speech complexity and lower accuracy whereas the less complex task PIE resulted in lower speech complexity and higher accuracy. This confirms Skehan's prediction that the complexity of the task and the speech complexity would be positively correlated.

Our results partially confirm Hypotheses 3 and 4. The results indicate that in case of complexity and accuracy scores, only task complexity—and not interaction conditions—seems to have a strong influence: in general, NAR scored higher on complexity, while PIE scored higher on accuracy. There was no significant difference between monologic and dialogic conditions under NAR on both the complexity and the accuracy scores. Robinson (2001b) contends that higher task complexity implies increased negotiation for meaning, confirmation checks and clarification requests which negatively affect participants' attempts at complex speech due to shorter turns and a higher number of elliptical responses. In other words, speech complexity under interactive tasks would suffer because of the nature of interaction itself. However, in our study, the interaction conditions affected only the simpler task PIE. This is interesting because this partially confirms Robinson's hypothesis and partially disproves it. As per our Hypothesis 3 and Robinson's hypothesis, dialogic condition should have resulted in less complex scores, but our results do not support this entirely.

Out of the three dependent variables, complexity and accuracy scores produce clear trends as discussed above. Regarding the third variable, fluency seems to have been affected by the combined effects of task structure and interaction conditions.

Recall here that in this study fluency was operationalized as the ratio of total utterance time (total oral performance time minus total pause time) to total turn time expressed as percentage. Generally, it is assumed that the more complex a task is, the less fluency scores would be. Accordingly, PIE should have had higher fluency scores; but our results only partially confirm this. Only in the dialogic condition did PIE produce higher fluency scores; in the monologic condition, there was no significant difference between the two tasks. In case of NAR, a more complex task coupled with dialogic condition produced lower fluency scores. Interestingly, under monologic conditions, there is no significant difference between PIE and NAR. This clearly indicates that in case of fluency, interaction conditions seem to be more important particularly when the task is complex, favouring monologic over the dialogic condition. This partially confirms Hypothesis 4.

We assume that in the case of PIE, personal familiarity with the information used in the task and scaffolding in terms of structures and vocabulary used by the interlocutor may have eased processing burden, thereby increasing fluency. Notice that while performing PIE under dialogic mode with higher fluency, accuracy is also higher, but complexity is lower than in the monologic condition. In other words, the lower cognitive demands of the task may have enabled participants to allot more resources to accuracy, while the familiarity of information and scaffolding provided by the interlocutor may have enabled higher fluency. We speculate that there is a complex interplay of task structure, proficiency levels of participants and familiarity with different conditions of interaction.

Regarding speech complexity, we need to note that Robinson (2001a, 2001b) argues that complex monologic tasks would generate more complex and more accurate speech than complex interactive tasks. The reason is that complex dialogic tasks would generate more interaction in terms of confirmation checks and clarification requests than their simpler counterparts (Robinson 2001b, p. 36) (also see Foster & Skehan, 1999). Such increased interaction may lead to frequent turn-taking and greater number of elliptical or short responses. This in turn mitigates speakers' attempts at complex speech. This could be one of the reasons why NAR—a complex task—in the dialogic mode generated lower complexity scores than monologic mode.

Regarding the fluency score of NAR (higher under monologic condition than dialogic mode), one possible explanation is that when there is higher number of confirmation checks and clarification requests each turn would naturally be short (see Foster & Skehan, 1999; Robinson 2001a, 2001b) and may contain longer and more number of pauses. This would affect fluency score, which is calculated as the ratio of total utterance time to total turn time expressed as percentage. The duration of pauses play a vital role in determining fluency score. If this is the case, the fluency score of complex dialogic task NAR would be lower than its monologic version. However, this assumption would require more research evidence for substantiation.

Implications

The most important findings of the study are as follows:

- Complex tasks significantly improve speech complexity.
- Simple tasks generate less accurate speech.
- Simple monologic tasks score higher speech complexity than simple dialogic tasks.
- Simple dialogic tasks score higher accuracy than monologic tasks.
- Simple dialogic tasks score higher fluency than complex tasks.
- Under monologic condition, interlocutor support has greater influence than task complexity.
- When task is complex, fluency is lower irrespective of interlocutor type.

These findings imply that interaction variables and task structure interact in predictable, principled and useful ways to promote performance and learning. We discuss the implications of our study below.

First, the study has showed that task structure is a complex construct that involves task-internal and task-external factors and their interaction. Though task complexity itself can be measured/controlled at the design stage using task design frameworks like those provided by Skehan (1998) and Robinson (2011), the actual complexity as experienced by learners and test-takers at the implementation stage cannot be determined merely by the measures of inherent task complexity. We have seen that the design characteristics of tasks (e.g. inherent structure and the materials used) that determine whether a task is simple or complex may interact with external variables—like interaction and the kind of information manipulation required—to alter their complexity as experienced by the learner.

The crucial question to ask is whether the task itself and its interaction with implementation variables support the goal(s) of performance set by the task designer. Therefore, in order to maximize learning and to elicit best performance out of the test-taker/learner, possible interactions with task-external variables have to be considered while designing and selecting tasks for various purposes.

Second, the observed trade-off between complexity and accuracy may be used to the advantage of the learner. In other words, the knowledge of trade-off enables us to design tasks that target particular performance goals for different proficiency levels. This emerges from the idea that at lower proficiency levels, learners require control over language, and at higher levels, they require control over complexity (Foster & Skehan, 1996). That is, we need to target accuracy at lower proficiency levels and complexity at higher levels. With the knowledge of trade-off, we may design tasks appropriate for different proficiency levels to promote or measure specific performance goals. For example, we found that highly complex tasks elicit higher speech complexity levels, but generate less accuracy. This finding allows us to use tasks that are moderately complex with low proficiency learners who would focus more on accuracy and improve it, thereby gaining confidence and control over their current level of language. Or, more proficient learners can be given comparatively

simpler tasks that would enable them to experiment with complex language, even at the expense of accuracy. This way, one could have specific performance targets to improve learners' proficiency levels or to elicit performance for assessment.

Third, the study shows that task complexity is conducive for increasing speech complexity and dialogic mode in general leads to higher accuracy and fluency scores. The teacher may design tasks accordingly. In assessment, particularly while assessing for specific purposes where one performance goal is preferred over another, these insights may be very useful. For instance, in the recruitment or assessment of a candidate for the post of a health professional, accuracy as well as fluency may be preferred over complexity.

Conclusion

The analysis in this study leads to three main conclusions. First, task complexity as experienced by the test-taker/learner is a construct that is heavily influenced by factors like various implementation variables that are external to task complexity. Second, monologic and dialogic interaction can be manipulated along with task structure to influence performance goals. Third, the constructs of task structure and the trade-off between complexity and accuracy can be manipulated in tandem with performance goals to design tasks and task sequences for specific groups of learners, keeping in mind the influences of task-external factors like interaction variables on task complexity.

These findings have potential applications in teaching as well as assessment. Using these findings, teachers may design appropriate tasks to match the existing language proficiency of the learner, to promote different performance goals independently and otherwise.

Nevertheless, further studies in the Indian context are required to corroborate and substantiate these findings. The efficacy of trade-off between complexity and accuracy as a tool in task design and implementation in the overall ecosystem of task-based teaching and assessment requires further research evidence. More investigations operationalizing dependent variables differently may give us further insights on performance, guiding better task design and implementation procedures. Additionally, studying the influence of different proficiency levels, kinds of interlocutor and their interaction could enhance our understanding of how these variables influence task-based oral performance. Since most task-based studies focus on foreign language contexts, ESL contexts like India where exposure to and opportunities for language use range from very high to very low, require more studies investigating the effects of task structure, interaction and other relevant variables like proficiency and language exposure to build and sustain a healthy learning-friendly environment.

References

Brindley, G. (1994). Task-centred assessment in language learning: The promise and the challenge (ED386045). In N. Bird, P. Falvey, A. Tsui, D. Allison, & A. McNeill (Eds.), *Language and learning: Papers presented at the Annual International Language in Education Conference (Hong Kong, 1993)* (pp. 73–94). Institute of Language in Education. ERIC. https://files.eric.ed.gov/fulltext/ED386045.pdf

Brindley, G. (Ed.). (2000). *Studies in immigrant English language assessment* (Vol. 1). National Centre for English Language Teaching and Research, Macquarie University.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics, 19*(3), 221–246. https://doi.org/10.1111/j.1473-4192.2009.00231.x

Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics, 19*(1), 1–23. https://doi.org/10.1093/applin/19.1.1

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299–323. https://doi.org/10.1017/S0272263100015047

Foster, P., & Skehan, P. (1999). The influence of source planning and focus of planning on task-based performance. *Language Teaching Research, 3*(3), 215–247. https://doi.org/10.1177/136216889900300303

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*(3), 354–375. https://doi.org/10.1093/applin/21.3.354

Kawauchi, C. (2005). The effects of strategic planning on the oral narratives of learners with low and high intermediate L2 proficiency. In R. Ellis (Ed.), *Language learning and lagnguage teaching: Vol. 11. Planning and task performance in a second language* (pp. 143–164). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.11

Long, M. (2015). *Second language acquisition and task-based language teaching*. Wiley Blackwell.

Mathews, S. M. (2020). *The effect of task-complexity, pre-task planning time and interlocutor on the complexity, accuracy and fluency of oral performance: A task-based empirical investigation.* [Manuscript in preparation]. Department of Humanities and Social Sciences, Indian Institute of Technology Kanpur.

Nitta, R., & Nakatsuhara, F. (2014). A multifaceted approach to investigating pre-task planning effects on paired oral test performance. *Language Testing, 31* (2), 147–175. https://doi.org/10.1177/0265532213514401

Norris, J. M. (2002). Interpretations, intended uses and designs in task-based language assessment. *LanguageTesting, 19* (4), 337–346. https://doi.org/10.1191/0265532202lt234ed

Norris, J. M. (2016). Current uses for task-based language assessment. *Annual Review of Applied Linguistics*, 230–244. https://doi.org/10.1017/S0267190516000027

Pang, F., & Skehan, P. (2014). The context for researching a processing perspective on task performance. In P. Skehan (Ed.), *Task-based language teaching: Issues, research and practice (TBLT): Vol 5. Processing perspectives on task performance* (pp. 95–128). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.5

Qian, L. (2014). Get it right in the end: The effects of post-task transcribing on learners' oral performance. In P. Skehan (Ed.), *Task-based language teaching: Issues, research and practice (TBLT): Vol 5. Processing perspectives in task performance* (pp. 129–154). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.5

Robinson, P. (2001a). Task complexity, cognitive resources and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *The Cambridge applied linguistics series: Cognition and second language instruction* (pp. 287–318). Cambridge University Press.

Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27–57. https://doi.org/10.1093/applin/22.1.27

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics, 43*(1), 1–32. https://doi.org/10.1515/iral.2005.43.1.1

Robinson, P. (2011). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In P. Robinson (Ed.), *Task-Based Language Teaching: Issues, Research and Practice (TBLT): Vol. 2. Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 3–38). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.2

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*(1), 38–62. https://doi.org/10.1093/applin/17.1.38

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skehan, P. (2014). The context for researching a processing perspective on task performance. In P. Skehan (Ed.), *Task-based language teaching: Issues, research and practice (TBLT): Vol 5. Processing perspectives on task performance* (pp. 1–26). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.5

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research, 1*(3), 185–211. https://doi.org/10.1177/136216889700100302

Skehan, P., & Foster, P. (2001). Cognition and Tasks. In P. Robinson (Ed.), *The Cambridge applied linguistics series: Cognition and second language instruction* (pp. 183–205). Cambridge University Press.

Skehan, P., & Foster, P. (2008). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the ealing research. In V. Daele, S. Housen, F. Kuiken, M. Pierrard, & I. Vedder (Eds.), *Complexity, accuracy and fluency in second language use, learning & teaching*, 29–30th March 2007 (pp. 201–222). Brussels, Belgium: Royal Flemish Academy of Belgium for Sciences and Arts. https://textarchive.ru/c-2002629-pall.html

Skehan, P., & Shum, S. (2014). Structure and processing condition in video-based narrative retelling. In P. Skehan (Ed.), *Task-based language teaching: Issues, research and practice (TBLT): Vol 5. Processing perspective on task performance* (pp. 187–210). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.5

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125–144). Oxford University Press.

Swain, M. (2001). Examining dialogue: Another approach to content specification and to validating inferences drawn from test scores. *Language Testing, 18*(3), 275–302. https://doi.org/10.1177/026553220101800302

Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 58*(2), 439–473. https://doi.org/10.1111/j.1467-9922.2011.00642.x

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Language learning and lagnguage teaching: Vol. 11. Planning and task performance in a second language* (pp. 239–273). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.11

Vasylets, O., Gilabert, R., & Manchòn, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning, 67*(2), 394–430. https://doi.org/10.1111/lang.12228

Vygotsky, L. S. (1978). Mind in Society: The development of higher psychological processes. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.). Harward University Press.

Wang, Z., & Skehan, P. (2014). Structure, lexis, and time perspective: Influences on task performance. In P. Skehan (Ed.), *Task-based language teaching: Issues, research and practice (TBLT): Vol 5. Processing perspectives on task performance* (pp. 155–186). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.5

Chapter 15
Using Task-Based Speaking Assessment to Measure Lexical and Syntactic Knowledge: Implications for ESL Learning

Vasim Tamboli

Abstract An enquiry into the present assessment practices in primary and secondary education in India shows that assessment of oral proficiency neither forms a part of certification exams nor is it regularly used for classroom purposes. In a few instances when assessments of speaking and listening are carried out, random topics are assigned to learners without checking for selection and/or gradation of task features using reliable conceptual frameworks, thereby limiting their assessment benefits. This paper presents an application of the Task-Based Speaking Assessment (TBSA) in the Indian ESL context by adapting Robinson's Cognition Hypothesis (Cognition and second language instruction. Cambridge University Press, 2001a) as a conceptual base of selecting 'task characteristics' to design and sequence a series of speaking tasks and assessing learner performance based on a mapping of task-specific criteria. The paper validates the benefits of applying the TBSA model with evidence from a recent empirical study done with Indian learners. It demonstrates the pedagogical benefits of selecting and sequencing speaking tasks in that it leads to a gradual improvement in oral performance, especially in two micro-aspects of language—lexical variety and syntactic knowledge. Thereafter, the steps of designing task-specific assessment criteria are discussed to help teachers understand how task-specific attributes of speech can be incorporated in descriptors to measure *linguistic* as well as *communicative attributes of speech* of SL learners and give them meaningful feedback.

Keywords Task-based speaking assessment · Cognition hypothesis · Assessment criteria · Lexical variety · Syntactic knowledge

V. Tamboli (✉)
EFL University, Hyderabad, India

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_15

Introduction

School education system in India has been biased with summative assessment across all stages of learning from primary and secondary to higher secondary levels. This attribute of 'examination oriented education' (Ramanathan, 2008, p. 124) is still retained in the current National Educational Policy with an introduction of *Census Examinations* (2020, p. 18) for Grades 3, 4, and 8 while restructuring the *Board Examinations* for Grades 10 and 12. But as a ray of hope, the policy now prioritizes assessment of basic learning and developing skills through core concepts and knowledge from local and national curricula and application to real-life situations (2020, p. 18).

In common parlance in India, summative exams are considered important in school education as they certify and promote learners to a next higher level. The summative examinations are conducted by two educational boards in India (national and state levels) for grades 10 and 12. These are external agencies conducting the exams, and in this assessment exercise, class teachers hardly get a chance to assess their learners. This has been critiqued in the works of Ramanathan (2008) and Ratnam and Tharu (2018) that teacher's perceived needs of what would benefit their learners finds no place in such externally driven assessment models.

Furthermore, in such examinations assessment of oral abilities is not a part of the proficiency construct. This is rather ironical to the fact that an inclusion of learners' communicative language use as a construct in language education was emergent as early as in the 1990s. In one instance, the Central Board of Secondary Education in 2004 made provision for oral assessment, but with very little weight and no official testing status. In another instance Maharashtra State Board in 2006 introduced oral assessment for year-end examination but revoked it soon after realizing its failure. These half-hearted and failed attempts project a general tendency that oral assessment in the Indian context has not been prioritized. It is true that oral assessment is genuinely difficult to administer due to the challenges arising from the principle of 'practicality' of time and resources. In addition to this, there is a lack of training for teachers to design, administer and evaluate speech performances (Brown & Abeywickrama, 2010). Similarly, inability to use the construct for classroom purposes results in because teachers report that they do not have sufficient time (Ramanathan, 2008, p. 123) or that due to lack of training they rely on controlled speaking tasks (e.g. predefined question answers); such tasks fail to track learners' communicative competence (Ramanathan, 2008, p. 121). In addition to this, large classes make it difficult to have one-to-one interactions required for authentic speaking assessments (Erling et al., 2016, p. 15).

Assessing Speaking and Listening (ASL)

A model that considers the inclusion of speaking skills in school education in India is Assessment of Speaking and Listening (ASL). That is practised in schools that come under Central Board of Secondary Education (CBSE). It requires teachers to conduct assessment of speaking skills using a holistic scale (CBSE, 2013, p. 1). This formal mode of summative assessment reserves 20% of scores from the end of year examination. A day before the examination, learners are generally given 20–30 topics to prepare their mini-presentation and pair interactions. The instruction given for *mini-presentations* is: (for student—B) 'describe something that you own and you value tremendously. You have 1 min to prepare and 2 min to present' and (for student—A) 'describe the most memorable journey you had' (CBSE, 2013, p. 26). In the *pair interaction task*, students A and B are given a visual stimulus (e.g. picture of a young village girl with the title 'child marriage') and asked to hold a mini-conversation for two minutes (CBSE, 2013, p. 27). After completing all the stages, the teacher–assessor awards marks to each learner.

An analysis of this speaking assessment model shows the presence of a *monologic* (one participant involved in the mini-presentation) and a *dialogic* (two participants involved in the pair interaction) task. However, the tasks have some limitations: the mini-presentation task does not have any detailed instructions and the pair interaction task does not have a prompt with cue questions to contextualize and make them more communicative. The dialogic task provides learners with only a picture to talk about a topic, and there are no separate cue/role cards available to create a rich and authentic context in which genuine interaction can take place. In addition, the task specifies only a generic version of criteria to assess speaking abilities with oral features of language such as—*interactive competence* (task management and turn taking), *fluency* (speed of delivery), *pronunciation* (stress and intonation) and *language* (accuracy) (CBSE, 2013, pp. 6–7). Considering all these limitations, this model does not seem to help learners produce a comprehensive and rich oral output that can have a variety of linguistic features and communicative functions. Moreover, the assessment criterion does not account for the totality of learners' oral performances, thereby raising doubts about the reliability and validity of the tool (Brown & Abeywickrama, 2010, p. 40). Lastly, as it is used for summative assessment, it lacks the potential to create conditions of assessment 'for' learning (Earl & Katz, 2006). These concerns have been voiced by practising teachers who feel that CBSE guidelines (2013) provide only a set of general instructions for task design and assessment criteria; therefore, lack of authenticity in these tasks fail to motivate learners to use English for communicative purposes (Ahmed & Agarwal, 2018, p. 17).

Using Assessment Practice as a Facilitative Tool

There is a need to place assessment practices in the broader context of teaching and learning. Towards this direction, Durairajan (2016, p. 55) proposes that there is a need to move from:

1. large-scale assessment to a more flexible classroom-based teacher-friendly assessment;
2. paper–pencil tests to alternate modes of assessment for summative (certification) and formative (classroom learning) purposes; and
3. assessment 'of' learning to assessment 'for' and 'as' learning to facilitate interface between teaching and assessment.

Such linking of assessment to teaching is expected to improve learning outcomes of ESL learners and help teachers and policy makers realize and achieve the real aim of assessment whereby one can keep reviewing the learning processes (Kapur, 2008, p. 10).

Form Focus in Language Assessment

There is a general understanding that learning a second language is a process which involves developmental sequences from 'limited knowledge' to 'more advanced knowledge' (Byram, 2004, p. 598). In an assessment scenario, this creates a need to expect learners show a gradual approximation to target norms over a period of time; so their language knowledge may be treated as 'dynamic' and not as a 'static' phenomenon independent of developmental sequences.

Traditional discrete point testing with a 'focus on form' (e.g. grammar-based multiple choice items or grammar cloze tests) violates the language acquisition principle, namely, L2 learning is a gradual approximation and foregrounds mastery over a series of isolated linguistic units or one grammar point at one time as valid linguistic knowledge (Oller, 1979, p. 37). As a critique of this philosophy of testing, *communicative tests* with 'focus on meaning' allow learners to use extended language production in real-life context (Morrow, 1979). Although this benefits L2 production and contributes towards making oral or written production fluent, it does not promise to measure advancements in accuracy of form or the role of form to express meaning for real-life communication. The limitations of both approaches paved the way for a third approach, i.e. 'focus on form' while attempting to solve communicative activities (Long, 1991; Long & Robinson, 1998). This approach advocates the use of a variety of pedagogic procedures to direct learners' attention to linguistic codes during the meaning making process. Such meaningful involvement of form subsequently motivates learners to participate in the learning process, as close to their real-life needs (e.g. use of different tenses to refer to different events or progression of time) (Prabhu, 2019, p. 199). It also captures a variety of aspects in language production (e.g. accuracy, complexity and fluency) in a comprehensive manner.

Task-Based Language Assessment: A Solution to the Problem

The pedagogic procedure to direct learners' attention to linguistic codes while using the target language is used is Task-Based Language Teaching approach and its application as Task-Based Language Assessment (TBLA). This model of assessment adds to the earlier model of communicative language testing by using a series of sequentially arranged real-life assessment tasks that can push learners to use a variety of linguistic forms to express meaning (Robinson, 1996, p. 111). Learners are required to act on a set of tasks to achieve a series of language learning outcomes. Through task involvement, learners display their language competence and use by producing extended pieces of discourse in the oral and/or written mode. Their language performance is assessed and based on it, feedback is provided and further teaching-related decisions are taken. Thus, TBLA has three crucial components: (i) *task as a stimulating input of language output*, (ii) *task-based construct of assessment* and (iii) *task-based criterion for assessment* (Bachman, 2002; Brown, 2004). By using TBLA, a teacher can evaluate how well learners are able to use the target language and show changes in inter language development over a period of time.

Task-Based Speaking Assessment (TBSA) in Classroom Context

An early attempt to develop a structure for a task-based approach to assess oral language ability was taken up by a project titled Guidelines for Task-Based University Language Testing (GULT) (Fischer et al., 2011). It reported that Task-Based Speaking Assessment (TBSA) can potentially include content (e.g. familiar ideas, themes and topics), align tasks with assessment objectives and consider learners' needs to help assessors estimate (assessment 'of' learning), prepare (assessment 'for' learning) and get learners involved (assessment 'as' learning) in assessment of language use in real-life situations (p. 15). Such tasks are normally designed as a series of tasks related to each other and used to assess various language skills in a context either individually (one skill at a time) or in an integrated manner (combination of skills) at different levels of abilities (p. 17). Moreover, TBSA also informs learners about their ability to communicate and can even make them a part of assessment procedure itself (Ellis, 2003, p. 285). All of these make TBSA a way to connect oral assessment to curricular objectives as well as to fulfil real-life communicative needs of ESL learners.

In extending these benefits of TBSA to the ESL context, it is important to attend to each of the three components of the TBLA model—*use of task as an input*, *defining the construct of assessment* and *assessing to provide feedback*. The model involves a number of issues to consider related to communicative and authentic use of the target language. These issues range from selection and design of tasks to deciding on

the content and language use abilities to be incorporated in the assessment criteria, training raters, assessing the performances and finally interpreting the scores/grades in sync with learners' future performances in real-life context (Ellis et. al., 2020) or target language use (TLU) domain (Bachman & Palmer, 1996, p. 23). These are essential components to fulfil predictive validity. Therefore, we propose that TBSA can be adopted in the Indian school context, especially to complement the ASL model (see Section "Assessing Speaking and Listening (ASL)"). The rest of the paper will demonstrate an improved version of the ASL model by adopting TBSA and argue that task features and task-specific assessment criteria can push learners to move beyond their current level of interlanguage use.

Role of Task Features in TBSA

The literature on task-based assessment offers varied dimensions in describing a 'task' with its rich and comprehensive components to cater to various assessment requirements. These dimensions are as follows: (a) *components of tasks* (e.g. meaning, real-world activity, outcome; Skehan, 1998, p. 95), (b) *task development characteristics* (e.g. background knowledge, context, real world language use reflections; Chalhoub-Deville, 2001, pp. 214–217), (c) *task selection criteria* (e.g. content areas, field specific tasks; Norris, 2001, p. 171) and (d) *task as a problem solving activity* (e.g. information gap, reasoning gap and opinion gap), which poses appropriate degree of cognitive challenge to the learners in using language functions (Prabhu, 1987, p. 48). An effort to incorporate features of these four dimensions make TBSA markedly different from other traditional pencil-and-paper based test (e.g. writing free-responses as essays or multiple choice tests) (Norris et al., 1998).

The literature around the construct of task has helped assessors elicit production through 'open-ended' tasks to assess language abilities in a comprehensive manner (Norris et al., 1998, pp. 188–207). Furthermore, use of open-ended tasks enables assessors to focus on each of sub-skills of speaking through connected discourse, and thereby get a detailed understanding of learners' strengths and weaknesses (Fischer et al., 2011, p. 26). Thus, TBSA helps a teacher align curricular elements to assessment and motivates them to use a set of authentic and real-life tasks for assessment.

A prominent component of TBSA is *task design factors*. Robinson (1996), for better L2 output among the ESL learners, has suggested a combination of task design factors—(i) task and task characteristics, (ii) learners' task performance abilities both as task completion and language use abilities and (iii) task complexity and task processing conditions. He substantiates the need to understand task complexity for pedagogical purposes and states that 'tasks should be sequenced for learners from simple to complex in order to promote success in performing complex tasks in the L2, as well as opportunities for further L2 learning and interlanguage development' (2011, p. 4). He mentions the need of considering complexity factors in language assessment, especially in designing parallel tasks. Such tasks can have

various levels of cognitive demands to assess learners at different levels of language proficiency (p. 5). Brown also supports the idea that the factors involved in designing and sequencing of tasks for pedagogic purposes are also inevitably related to the designing and sequencing of tasks for large-scale language assessments purposes (2004, p. 110).

Cognitive Processes Underlying Tasks

The idea of manipulating tasks to trigger cognitive processes to have a positive impact on ESL learning is supported by many researchers working in the cognitive domain (Robinson 1995, 2011; Skehan, 1998) and assessment researchers like Bachman (2002) and Brown (2004). In particular, we refer to Robinson's 'Cognition Hypothesis' which theorizes that the human brain has multiple attentional resources and complex tasks are likely to increase accuracy and language complexity but not necessarily fluency. He claims that '*task complexity* is the result of the attentional, memory, reasoning and other information-processing demands, resulting from design characteristics, are relatively fixed and invariant' (2001, p. 28). To guide these claims and realize their benefits in pedagogy and research, Robinson has provided a comprehensive taxonomy—Triadic Componential Framework (TCF)—of factors as a determinant of task performance into the three main categories*: task complexity, task condition* and *task difficulty* which form the base of selecting/designing and sequencing tasks for oral performance (2011, p. 6). A combinatorial effect of these three factors are likely to: (i) improve learner performances, (ii) create language learning opportunities and (iii) consider learner's level of language proficiency in designing assessment.

While manipulating task complexity (i.e. *cognitive factors)* in task design, two sub-factors play a role: (i) resource directing and (ii) resource dispersing (Robinson, 2011, p. 14). *Resource directing* refers to task demands made on the use of concepts in expressing and understanding task performance (e.g. ± intentional reasoning) by which learners' performances are likely to be more grammatically accurate, lexically varied and syntactically complex than its simple counterpart. *Resource dispersing* refers to task characteristics making performative or procedural demands (e.g. ± prior knowledge), which would require learners to attend to all three aspects—fluency, accuracy and complexity—in a relatively unfocused manner.

The studies that have applied the Cognition Hypothesis have systematic sequencing of tasks from simple to complex and examine task-specific aspects of speaking abilities—*linguistic* as well as *communicative benefits* (Robinson, 2001b, 2011; Kuiken & Vedder, 2011). Linguistic benefits like the ability to use a *variety of lexical items* are better achieved when the task involves conceptual demands. For example, [+ intentional reasoning] in a speaking task requires learners to state their likes and dislikes and present others' point of views about a place of visit. This requirement is likely to draw upon learner's ability of assigning internal state terms (*happy, sad, want*) to one's own or others' mind (Ishikawa, 2011; Robinson, 2011). Communicative benefits such as *an opportunity to negotiate, clarification requests*

and *confirmation checks* are evident through performance conditions (Robinson, 2001b, p. 36). For example, in an interactive task the need to discuss and come to an agreement about a place of visit offers communicative benefits to know each other's preferences and negotiate politely. Moreover, this involves asking/responding to questions, which may direct learners' attention to use of a variety *of syntactic structures* like subordination and forming different question types (e.g. wh-questions, yes/no questions).

In speaking assessment, as there is a need to tap into a variety of *linguistic categories* used for communicative purposes, these can be achieved through task features which can account for learning benefits. Drawing on Robinson's (2011) Triadic Componential Framework, in this paper, a study on Indian ESL learners is reported (Tamboli, 2017). This is an application of TBSA using the TCF model and attempts to provide a better alternative to the currently available ASL model in India. The study examines the following research question:

RQ 1: Do task complexity features incorporated in task design to elicit oral performance have an impact on language use across a series of tasks?

The Study

Participants

Twelve Indian ESL learners (6 male; 6 female), aged 13–14 years, enrolled in Grade IX in a state government regional medium school located in Maharashtra (India) participated in the study. The learners were exposed to the target language for five years and only in the classroom contexts. They had A2 to B1 level of oral language proficiency as per the global scale (CEFR, 2001). Their teachers reported that they could produce simple connected texts on topics of personal interest (e.g. planning a school trip). This was further ascertained by the researcher based on personal communication and analysis of writing and spoken samples prior to the study.

Task Design

The study applied the Triadic Componential Framework to design ESL speaking assessment. It was intended to understand learner growth across tasks and not to treat language capabilities as a static construct, which is an otherwise commonplace practice in summative language assessments used to certify learners. Thus, task design was an independent variable used to design the four tasks appropriate for A2-B1 level learners. They were differentiated through varying levels of text length and cognitive complexity as presented diagrammatically in Fig. 15.1.

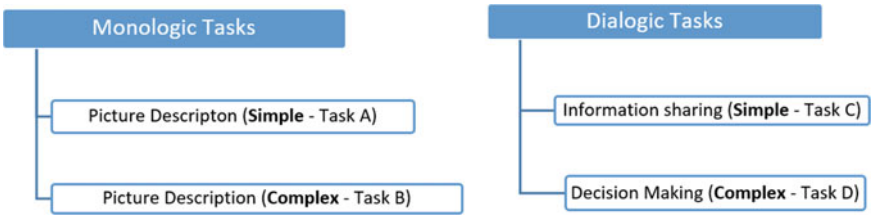


Fig. 15.1 Speaking assessment tasks

Given below are the actual tasks with an analysis of their task design features to understand how Robinson's Triadic Componential Framework as a model was applied.

Monologic tasks (*Task A* and *Task B)*

Task A is a picture description task (see Fig. 15.2). It is a simple picture because it has *a few objects* to be described in about 10 utterances.

Task B is also a picture description task but with two pictures (see Fig. 15.3). In this task, learners have to look at both pictures and compare and contrast two scenes in about 10 utterances. Finally, they have to indicate which event (picture) they would like to be in and give reasons. Therefore, Task B is at a higher cognitive complexity than Task A and the expectation is that Task B will yield higher fluency and linguistic variety.

Describe the picture given below. Speak for about 2 minutes.

Fig. 15.2 Task A: Single picture description

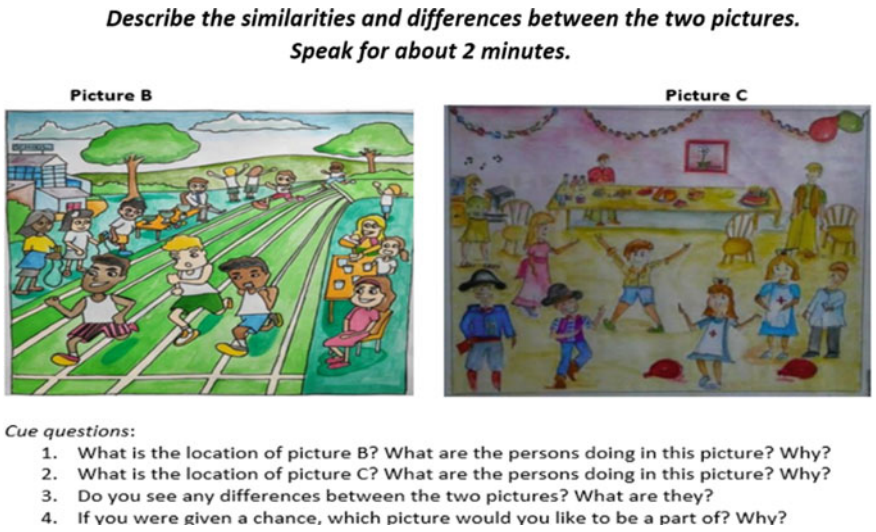


Fig. 15.3 Task B: Comparison of two pictures

Table 15.1 Task characteristics of monologic tasks

Cognitive complexity	Monologic simple (Task A)	Monologic complex (Task B)
Resource-directing variables	[+ few elements] *A single picture with a few elements*	[− few elements] *Two pictures with more elements*
	[− intentional reasoning] *Transferring information*	[+ intentional reasoning] *Transferring information, stating the intention, Selecting a favourite event and justify it*
Resource-dispersing variables	[+ few steps] *Only describe one picture*	[− few steps] *Compare and contrast between the two pictures*

Table 15.1 presents a list of features of the two monologic tasks[1] to explain the varying levels of task complexity within resource-directing and resource-dispersing dimensions:

Dialogic tasks (Task C and Task D)

These two tasks are conversation-based (dialogic) tasks conducted on the basis of role cards and are at two levels of complexity. In Task C (see Fig. 15.4), the simpler version of the dialogic task, learners work in pairs and express opinion on a single concept at a time (adapted from Friderick, 1984, p. 29). Learners are given two role

[1] We adapted and customised a set of three pictures from *Cambridge English: Young Learners English Test Flyer* (2013, 2014).

One day, you meet your friend at a market place and you get talking to each other. You talk about the things given to you in role card. Read it carefully and participate in a short conversation with your friend.

You need to: (i) *give and ask for information*, and (ii) *express likes and dislikes*

Role A: Imagine that you have moved to a new locality/ neighbourhood. So talk to your friend what you like and dislike about the town. Your friend will talk about the new school she/he has started attending. Your friend will say her/his likes and dislikes about the new school. *Ask your friend relevant questions to carry the conversation further.*

Town

- Name of the locality you have moved in
- How is life there (crowded/calm; neat/dirty; organized/chaotic)?
- What are the basic facilities there?
- How are the people in the locality?
- Your opinion about the locality

Role B: Imagine that you have started attending a new school. So talk to your friend about what your like and dislike about the school. Your friend will talk about the new town she/he has moved in. Your friend will say what she/he likes and dislikes about the town. *Ask her/him relevant questions to carry the conversation further.*

School

- Name of the school
- School timings
- Activities at the school
- Teachers and students at the school
- What you like about the school
- What you dislike about the school

Fig. 15.4 Task C: Information sharing

cards. When the first learner speaks, the second learner asks questions on the basis of the cues presented in the role cards and vice versa.

In Task D (see Fig. 15.5), the more complex dialogic task, the learners need to

Task: Imagine that your class is going for a trip next week. The place is not decided yet. You will have to discuss your choice with another friend and convince him/her why your choice is a good one. You also need to listen to what your friend has to say and express your opinion to his/her choice and if you like the place accept his decision and if you do not approve his choice then politely decline.

Goal of the task: *Remember that you and your friend have your own preferences and need to discuss with each other and come to a conclusion as to which place to go for the school trip.*

For the trip, you can think of any place of your choice. You give information about your preferences and talk about its advantages to make your friend agree to your choice. Your friend will talk about another place of her/his own preference. You can exchange information with respect to the following points: *(e.g., distance, transportation, length of stay, food, expense, weather, your opinion).*

Fig. 15.5 Task D: Decision-making

Table 15.2 Task characteristics of dialogic tasks

Task characteristics	Dialogic simple (Task C)	Dialogic complex (Task D)
Cognitive complexity		
Resource-directing variables	[+ here and now] Reference to events happening now	[− here and now] Reference to the event that will happen in the near future
	[+ few elements] A few details to talk about school and town	[− few elements] More number of elements to talk about with a lot of details
Resource-dispersing variables	[+ few steps] A few steps (give and check for information, express opinion)	[− few steps] More number of steps (ask and answer questions, compare between different options, and come to a decision which place to visit)
Task condition		
Participation variables	[+ convergent solution] Agreement is not required but just to discuss and express information	[− convergent solution] Agreement is required to decide a place of visit as a part of task outcome
	[+ negotiation not needed] Negotiation is not required	[− negotiation not needed] Need to negotiate with a partner

imagine that their class is going for a trip and they have to decide a place for this (adapted from Friderick, 1984, p. 110). Then as a part of planning for the trip, the learners are required to come to a decision about which place to visit and justify it.

Table 15.2 presents a list of features which contribute to two different levels of complexity across the two dialogic tasks. Note that for the dialogic tasks there is an extra factor 'interactive demands of the tasks' within the rubric of 'participation variables'.

All these four tasks were administered over a period of 15 days with a gap of 3 days after every task to ensure that there was no task familiarity effect and that the learners were not exhausted by doing all the tasks in a rapid succession (Weir & Wu, 2006, p. 170). The performances were audio recorded and then transcribed and analysed using a list of measures. In this paper, we report learner performances on the basis of two micro-measures of language that underlie both fluency and complexity. These are *lexical variety* and *syntactic knowledge*.

Findings and Interpretation

In accordance with a claim made by Robinson et al. (2009, p. 540) that specific conceptual–linguistic units of analysis can be used as one of the most appropriate ways to capture the effects of task complexity on performance, the performance of the

participants of the study were assessed on two specific linguistic features—*lexical variety* and *syntactic knowledge*—to observe actual instances of progress. Learner performance as a group is first reported based on frequency of occurrences of these two features. Specifically, these two features are reported here as our units of analysis to show teachers which finer aspects of form giving rise to meaning can be observed in spoken performances. This is likely to increase teachers' awareness and help them apply similar measures in their assessment practices.

Lexical Variety

Although there are standardized numerical measures to assess lexical diversity and density like TTR and Guiraid index, we do not report on such measures. Instead we analyse learner performance on the basis of lexical knowledge that teachers can use as an easy to detect measure and give learner feedback on language use based on this micro-feature. To understand the presence of lexical variety in learners' performances, we considered the measure of verbs, which form the pivot of sentences and content is expressed based on the use of a variety of verbs along with their arguments (NPs and PPs) (Levin, 1993; Sinclair, 2011). Verb usage has provided information about learners' ability to describe a variety of events, actions and descriptions (stative) arising from task requirements of Tasks A to D. The three types of verbs used with regard to their arguments are given in Table 15.3.

To know whether task complexity has a role to play in learners' use of particular verb types in their speech, the frequency count across the four tasks is presented in Figs. 15.6, 15.7, 15.8 and 15.9.

Table 15.3 Types of verbs

Verb complementation types	Use of verb complementation in oral production (Sinclair, 2011)	Examples	Task reference
Intransitive verbs NP (self-agent) + VP	To talk about: action and events (e.g. participate), physical behaviour (e.g. dance), and position (e.g. stand)	All children participated Two boys are dancing Sun is shining in the sky	Task B Task B Task A
Transitive verbs NP1 (agent) + VP + NP2 (object/theme)	To elaborate information: describe events (e.g. climb, ride), reporting verbs (ask, say)	We can ride in a boat at Mumbai Many people say that Mahabaleshwar is beautiful	Task D Task D
Causative verbs NP1 (agent) + VP + NP2 (patient)	It indicates an action which is caused to happen (e.g. make someone do something, help)	Teachers make us study more Planning before trip helps us enjoy well	Task C Task D

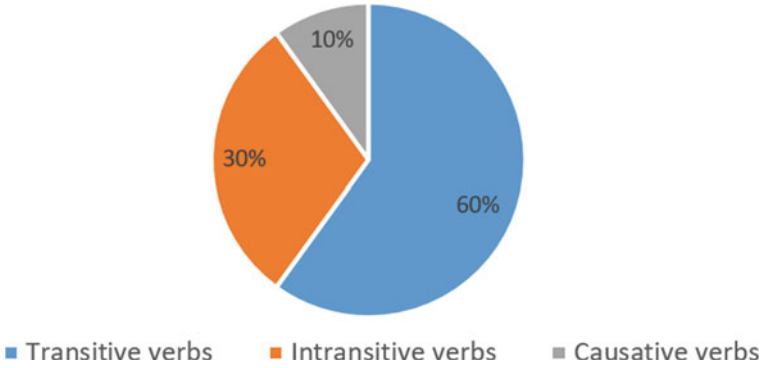


Fig. 15.6 Proportion of use of verb types in four tasks

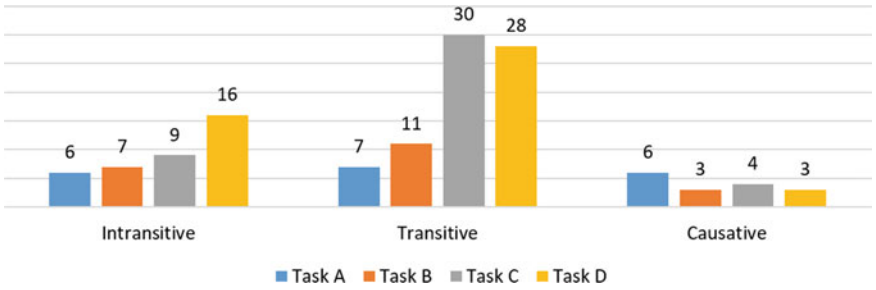


Fig. 15.7 Overall frequency count of verb types in four tasks

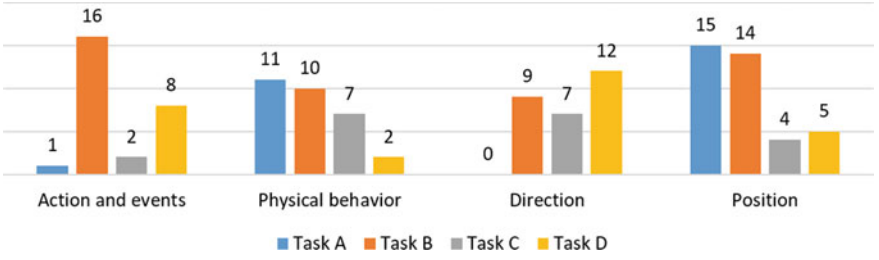


Fig. 15.8 Spread of uses of intransitive verbs across tasks

A comparison between two task conditions shows that in the monologic tasks a total of 40 verbs are used, whereas in dialogic tasks, a total of 90 verbs are used. Therefore, verb usage goes up from 31% in monologic tasks to a double of 69% in dialogic tasks.

Across the three verb categories, transitive verbs are highest (n=71) followed by intransitive verbs (n=38) and causative verbs (n=16) (see Fig. 15.6).

Figure 15.7 shows task-wise occurrence of the three sub-types of verb:

Fig. 15.9 Spread of uses of transitive verbs across tasks

From Fig. 15.7, it is observed that in the complex monologic task (Task B) and the dialogic tasks (Task C and Task D) there is an increase in number of intransitive and transitive verbs. But causatives are less in number across all tasks. Thus, task complexity features seem to have made learners experience some gain in use of verbs, especially in transitive constructions.

Let us look at the possible communicative reasons for use of the three verb subtypes.

Intransitive verbs: These are usually used to express states or describe events. Figure 15.8 shows that the varied instances of use of intransitive verbs across the four different tasks are due to task design factors.

In Table 15.4, actual examples from learners' performances are presented from the four tasks.

Transitive verbs: These are generally used to elaborate information as the objects could be events, humans or reporting verbs. Objects appear after linking verbs to describe size, colour, shape, qualities and types of things and people. In Fig. 15.9, the occurrences of transitive verbs where learners attend to various linguistic features are presented.

Table 15.4 Examples of intransitive verbs across tasks

Intransitive verbs	Examples	Uses
Task A	Sun is flashing	Action and events
	One boy is talking	Physical behaviour
	Sun is shining in the sky	Position
Task B	All children participated	Action and events
	Two boys are dancing	Physical behaviour
	Some girls are standing near the table	Position
Task C	Our school opens at 10 a.m.	Action and events
	In my village, a brook flows towards the river	Direction
Task D	We can sail in the boat	Action and events
	I went there	Direction

Table 15.5 Examples of transitive verbs across tasks

Transitive verbs	Examples	Uses
Task A	The people are waiting for the train	Event in picture
	There is a big café	Linking verb
Task B	Some girls are filling bottles with juices	Event in pictures
	They are cheering up the boys	Human participants
	There is a blue sky in the picture	Linking verb
Task C	We organize sports completion in new school	Event
	They save girl child	Human participants
	I forgot to say that there are basic facilities	Reporting verb
	My school's playground is very big	Linking verb
Task D	There are so many activities like riding a horse	Event
	Many say that Mahabaleshwar is very beautiful	Reporting verb
	There hotels are good	Linking verb

In Table 15.5, actual examples of transitive verbs from learners are presented from the four tasks.

In describing events there was a need to talk about more number of themes (playground and birthday party), people and their costumes, different places of a visit [− few elements] in the complex monologic and complex dialogic tasks. This led to higher number of transitive verbs in Task D (n=49) and Task B (n=40) followed by Task C (n=33) and Task A (n=27).

Reporting verbs are concerned with asking for and giving information—which is possible in task where more than one participant is involved; each participant expresses his/her opinion. So they possess information which is [− one way flow]. In learners' performances the use of reporting verbs for dialogic tasks (n=10) is detected as this task required [+ interactional reasoning].

Learners use human participants in the object position when they need to speak about something that affects a person rather than a thing. Across the four tasks, highest number of occurrences of this pattern was observed in Task C (n=12). In this task, the topics were about school and town where learners were required to talk about people. This might have made them use human as participants in the object position in this task (e.g. *They save girl child*).

Causatives: Causative verbs are concerned with an action which is a caused by an agent. The learners have used fewer causative verbs in comparison with intransitive and transitive verbs across the tasks. This reveals that the learners did not talk about how one person could cause another person to do something. However, there are a few instances of causative verbs as the learners talked about other people and presented different point of views (e.g. *My parents will not allow me to visit Goa*). Table 15.6 illustrates a few actual examples of use of causative verbs from learner oral data:

In all, the learners have used a variety of verb types across the four tasks for different communicative purposes. The verbs are used to describe a variety of events

Table 15.6 Examples of causative verbs across the tasks

Transitive verbs	Examples
Task B	Two girls are making juice for the persons who are coming to see the race
Task C	In new school, teacher makes us study more
Task D	My parents will not allow me to visit Goa

that requires different arguments. The gains in intransitive and transitive verbs supports that the learners do have a wide repertoire of verb complementation in mind. An increase in the use of verb complementation types depends upon task design features and resource-directing and resource-dispersing dimensions. Thus, inbuilt task features have helped learners actively use a variety of complementation to express rich content and in the process it has added to their expression of lexical knowledge used to express forms in the target language.

Syntactic Knowledge

To use knowledge of verb types in forming utterances/sentences, we chose to analyse syntactic knowledge. Although in measuring syntactic knowledge there are standardized numerical measures as T-units, Mean T-Unit Length (MTUL), clausal variety (Kuiken & Vedder, 2011, p. 95) and so on, we use a measure that teachers can relate to and can give feedback on learners' developmental use of SL (English). To tap into learners' syntactic knowledge, a micro-aspect of form-focused language use and the variety of verb phrases (VPs) occurring in learner performances were considered. This was to use the lexical knowledge of verbs to create different linguistic structures. So, lexical knowledge to form focus knowledge was considered which in turn pushed learners in expressing meaning as required through four tasks. Based on the verb patterns given in the *Oxford Advanced Learner's dictionary* (Hornby, 2015, p. R4), we considered two categories of verb phrases:

Simple VPs (mostly with the use of simple present tense, progressive present form): These are mono-clausal utterances indicating a basic linguistic knowledge to describe objects and/or present information. The examples of types of simple verb phrases are presented in Table 15.7.

Table 15.7 Examples of simple verb phrases across tasks

Sub-types of simple verb phrases	Examples of utterances	Task reference
V-ing	The train is coming	Task A
VN, simple present tense	I do not like competition	Task B
Wh-question	What do you like about school?	Task C

Table 15.8 Examples of complex verb phrases across tasks

Sub-types of Complex VPs	Examples of utterances	Task reference
V-Wh relative clause	There is a stationmaster who is flying a flag There is one boy whose legs are hurt	Task A Task B
V-that as demonstrative	I dislike that village I know that fort is birthplace of Sambhaji Maharaj	Task C Task D
VN-that relative clause	We should know Khandoba fort that is why we have to go there	Task D
V-to infinitive	We have to carry our luggage	Task D
VN-to in Wh-question	Which place do you want to go?	Task D

Complex VPs: These are multi-clausal utterances with infinitival clauses, relative clauses and other subordinating clauses. These involve a higher level of linguistic knowledge (e.g. 'that' is used to construct multiple clauses to elaborate upon information and to give reasons). The verb phrases included in complex VPs are presented in Table 15.8.

To find out whether task features contribute to learner performances, a distribution of all the nine types of VPs across the four tasks is presented (see Fig. 15.10).

As Fig. 15.10 indicates, there are instances of simple and complex clauses across the four tasks. The monologic tasks have facilitated use of relativization with 'Wh-marker (a total of 22 utterances), whereas the dialogue based tasks have resulted in a greater use of relativized clauses with 'that' marker (a total of 32). Hence, there is a growth in use of complex VP structure in the dialogic tasks that is more than in monologic tasks as they have [+ participant variable] as an additional feature.

The linguistic gains observed through participation across the four tasks will now be presented with a qualitative analysis of sample data from the performance of Learner 1 in the monologic and dialogic tasks. We would comment on the use of the two micro-features—*lexical variety* and *syntactic knowledge*—to show how the learner as a representative of the participants of the study could use a variety of complex structures in his/her oral production.

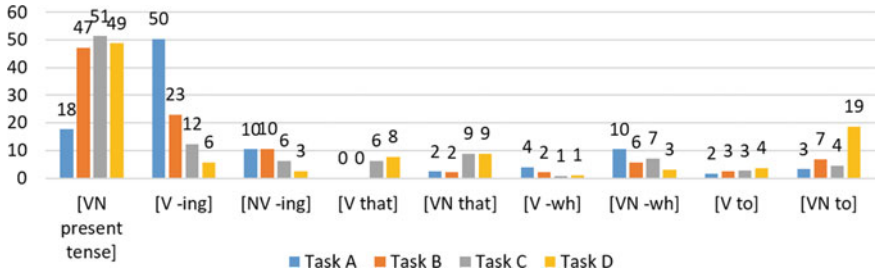


Fig. 15.10 Types of verb phrases across the tasks (in %)

Task A

> This is a picture of railway station. There are number of peoples present on platform who are waiting for their train. There is a big café. Just near the café, two persons are sitting with their coffee. There is a big bench. Two women and men are sitting on a bench. A lady with pink dress is playing with her baby, and another lady is reading a newspaper and also listening songs…

Task B

> I can see there are two pictures. First picture is outdoor game and second is indoor. We can see there is running competition between some children. Some children only few children are participating in competition, and in picture B, all the children are enjoying in the party and dancing on the music. Two girls are cheering up the children for participating in this competition. We can see there are two tables in picture A and a table in Picture B. On the, in the first picture we can see some glasses filled with cold drinks and also in picture B there some glasses with cold drinks and also cake.

As the monologic tasks are picture description tasks, the learner has used simple present and present continuous tense in declaratives with a VN structure. But there is a move from referring to less number of elements [+ few elements] to a comparison of more number of elements in picture set B [− few elements]; linkers are used to express comparison (*and, also*) which result in the use of coordinating clauses.

Task C

Role B	Hey, Hi Vaishnavi! What are you doing here?
Role A	I just came with my uncle **but** *what are you doing here*?
Role B	I am glad to meet you. I took admission in Maharshi Walmiki school and you?
Role A	No. *Tell me something* about your new school
Role B	My school timing is 10 a.m. to 5 p.m. and yours?
Role A	*Listen*, I have moved to a new city. It's Pune. It is a very nice city **but** life in a city is very crowded and also somehow dirty. *Tell me your prayer time or lunchtime*
Role B	In my school the prayer time is 11 a.m. and lunch time is 2 p.m. and your?
Role A	Okay. *Do not ask me about my school life.* I shall tell you something new about my new town
Role B	But you have to manage with it

In Task C, learner 1 with Role A now moves from transferring simple information and description to holding a conversation with learner 2 in Role B. The participant

also asks questions to continue the conversation and bring in new relevant sub-topics. This shows the use of [+ intentional reasoning], a higher task complexity feature. As a result, the learner displays a variety of syntactic knowledge like subordination (marked in bold in the text) and imperatives (marked in italics).

Task D

Role A	Hi Poornima. **What have you decided about our trip? Our principal has given a very good chance for deciding our place. It is very shocking for us to decide but what is your opinion, where we have to go?**
Role B	Yes. It is difficult but I have decided
Role A	Yes. What is it?
Role B	It's Kolhapur
Role A	Kolhapur?
Role B	Yeah. It is a very beautiful place in Maharashtra. Its distance is about 360 km from Pune
Role A	**But can we, but what transportation is available for Kolhapur?**
Role B	The train, buses, planes are available from Pune to Kolhapur or Mumbai. We can go anyone which affords our pocket
Role A	**But I do not think there are so many attractions at Kolhapur**
Role B	There is a temple of Mahalaxmi which is built from black stones

In performance on Task D, there is a move from only discussing and expressing information to making [+ convergent solution] with respect to the decision-making task; the learner attempts to express disagreement in a polite form (e.g. framing embedded questions) to make [− convergent solution]. This has made the learner use coordination (*but*), subordination (*for*) to give and ask for reasons and support a choice (marked in bold in the text).

In sum, the performance of learner 1 across the four tasks demonstrates that linguistic variations took place primarily due to the task requirements. It is observed that the increasing order of difficulty of the tasks along with the use of resource-directing and resource-dispersing dimensions, and participation variable have helped the learner use language in a variety of structures and to express different kinds of language functions ranging from informing, describing, comparing, evaluating, expressing opinion (agreement and disagreement) and steering the conversation, to name a few.

The study presented above has demonstrated that task complexity has resulted in a variety in language use in the oral modality, albeit small but significant for the learners. If teachers can be made aware of the beneficial role of TBSA on interlanguage development, it may help them define test construct, design real-life tasks and develop assessment criteria in sync with cognitive and interactive factors of real-life speech. This is also likely to help teacher–assessors tap learner performances and linguistic gains over a period of time and make assessing speaking a more valid and reliable activity within the framework of formative assessment. It is also likely to increase teachers' proficiency and raise their metalinguistic awareness, whereby

they can identify improvement (or lack of it) in learner language to express meaning communicatively and help them pay attention to form to express meaning. It can prepare them to provide meaningful and level appropriate feedback. Thus, in the next section, we show how teachers can gauge such linguistic gains in a step-wise manner. A way to do this would be to learn how to design a series of tasks with an increase in cognitive load and incorporate features in assessment criteria within the framework of TBSA that can give rise to assessment for learning.

Steps in Designing Tasks and Assessment Criteria for Speaking

On the basis of the study design and findings reported above, the paper would now attempt a step-wise demonstration of issues in task design and assessment criteria for teachers to apply TBSA for classroom-based assessment and to provide feedback. It is hoped that this step-wise detailing would help ESL teachers develop their assessment instruments to tap the underlying linguistic abilities of the learners according to contextual needs.

Step One: Task Selection (or Design)

In designing tasks for assessment of speaking, teacher–assessors can use the following general guidelines:

1. use a variety of speaking tasks under both the conditions—monologic and dialogic—as each task type helps learners in unique ways;
2. include a variety of task complexity variables which direct learner attention to different linguistic aspects and perform better;
3. include linguistic as well as communicative abilities (e.g. language functions) in the speaking assessment construct;
4. account for task-specific linguistic components in the assessment construct;
5. include task-specific linguistic and communicative components in designing of assessment criteria.

All these factors (1) to (5) contribute to an understanding of the ways and benefits of designing tasks by considering both ‘cognitive’ and ‘performance’ conditions. Such tasks can help teachers provide learners with the practise of certain linguistic aspects like lexis, structures and discourse organizing features. For example, in a sequence of tasks presented in this paper, there is an increase in syntactic knowledge found in responses to Task D as a result of including cognitive features such as *reasoning demands* and *more steps* to make the task more complex. This has pushed learners to attend to form to express meaning. More number of task requirements such as to describe, select/reject the places of visit, justify and plan their decisions carefully have made learners use *relativization* and *transitive and causative verbs.*

Table 15.9 Scoring checklist for content (for Task D)

1	Asks for and provides general information about places to be visited
2	Asks for and expresses opinions about the places
3	Compares between different opinions
	Agrees, disagrees and supports decision
5	Negotiates for a choice
6	Comes to a decision of a place of visit

Step Two: Creation of a Content Checklist

In designing task-specific assessment criteria as a primary step, teacher–assessors can create a *scoring checklist for content*. This makes scoring learner performances easier since each task is expected to produce different content. In Task D learners are required to use content in terms of language functions to finally achieve the task goal (come to a decision). As a distinct feature, this criterion is simple to implement in terms of estimating learner performance and scoring them as in presented in Table 15.9.

Step Three: Creation of Task-Specific Comprehensive Criteria

To draw the pedagogical benefits of using the two linguistic units—lexical variety and syntactic knowledge—let us look at how to incorporate a valid estimation of these features in task-specific speaking assessment criteria. This is because language teachers would not be able to measure the presence of these features as it is time consuming and cumbersome. But having an awareness that these form focused features can help them draw learner estimates and give feedback to create conditions of assessment 'for' learning, we show how these features can be incorporated in assessment criteria. We believe that an inclusion of such features in task-specific criteria can help teachers give feedback to ESL learners who can improve their language based on what their teachers make them 'attend to' through the feedback.

The creation of a comprehensive model of task-specific assessment criteria to represent the important sub-skills of speaking is a crucial step towards systematic and fair assessment and feedback. CEFR guidelines can be used to design a holistic assessment criterion for a dialogic task where learners may need to take a few decisions (2001, p. 26, 81). The criteria can be divided into three parts: content, language complexity and accuracy as presented in Table 15.10. In content, two categories of communicative abilities can be considered: (i) fluency with interactional strategies and (ii) the use of language functions. In language complexity the use of syntactic knowledge features (as found from the study reported in Section "The Study" above), can be incorporated. By doing so the teacher–assessor can look for the variety in

Table 15.10 Task-specific holistic assessment criteria for a decision-making (dialogic) task

<table>
<tr><th rowspan="2">Accuracy</th><th rowspan="2">Language complexity</th><th colspan="2">Overall content</th></tr>
<tr><th>Use of language functions</th><th>Fluency and interactional strategies</th></tr>
<tr><td rowspan="2">Use of simple structures and complex structures with a few errors; but meaning is not affected</td><td>A good mix of simple and complex VP based utterances</td><td rowspan="3">Provides relevant information with specific details and decides on the focus of the task (the key idea based on which a decision has to be taken)
(e.g. see Table 15.9 for a task-appropriate content checklist)</td><td>Speaker volunteers information with some effort (pauses and repetitions)</td></tr>
<tr><td>In complex utterances presence of relativized clauses with ‘that’ and wh-markers</td><td rowspan="3">Can perform on the following functions:
ask and respond to questions
express opinions
evaluate an opinion; take a decision and express it; and support/reject a decision</td></tr>
<tr><td rowspan="2">Attempts to use tenses and sub-verb agreement; but they need not be uniform (e.g. In my locality there are parks and gardens…, There will be many facilities…)</td><td>Use of verb complementation with a good mix of intransitive, transitive and causative verbs</td></tr>
<tr><td>Use of linkers to express opinion (but, because), give reasons (because), and come to a decision (therefore, so)</td></tr>
</table>

Note that this criterion is provided as sample that can be adapted for classroom purposes. This task type is an authentic complex dialogic task. So, it can be used to assess learners’ communicative use of a variety of language functions. A teacher can use the above comprehensive criteria as a checklist of content and linguistic features that make oral productions rich in meaningful and context appropriate

presentation of verb phrases and use of relativized clauses that are likely to be used because it is a reasoning task.

Step Four: Incorporating Micro-Aspects of ‘Form Focus’ in Communicative Tasks for Assessment Criteria

In designing task-specific speaking criteria (see Section “The Study”), one can incorporate the use of a variety of linguistic structures. This would help teachers in giving form-focused feedback to develop content. However, since this is a detailed and a far more nuanced way of designing assessment criteria, teachers can treat this as an optional step. They can use this information at a later stage and with learners who would can handle this kind of detailed feedback on their performance.

Table 15.11 presents task-specific assessment criteria for syntactic knowledge with respect to the types of clauses that can feature during language production (oral or written) across four tasks (mentioned in Section “Participants”).

Table 15.11 Assessment criteria for use of types of verb phrases

Task A	Task B	Task C	Task D
Use of [V-ing] to describe simple things happening in the picture (e.g. All people are smiling)	Use of [VN] present to talk about identify differences, compare between two pictures, and state the opinion (e.g. I do not like competition)	Use of [VN Present] in declarative and question sentences to state likes and dislikes and ask and answer questions (e.g. What do you like about your new locality?)	Use of [VN Present] in declaratives to make suggestions (e.g. We can tell headmaster our choice) and question structures to give and ask for opinions (e.g. Which place do you have in your mind?) Use of [VN –to] to talk about future plans (e.g. Which place do you want to go to?)
Use of 'wh -there' construction to refer to visual picture and to be more specific (e.g. There is a station master who is flying a flag)	Use of 'wh -there' construction to refer to visual pictures and to be more specific in giving information (e.g. There is a boy whose legs are hurt)	Use of 'wh-there' construction to give information (e.g. There are some boys who also have some groups)	Use of 'that as a determiner' to refer to action which will happen in near future (e.g. We cannot see more place in that Jejuri fort)
	Use of 'wh-relativazation' as relative pronoun in declarative statements (e.g. Picture A depicts who is first)	Use of 'that as determiner' to denote a place which is not visible at that point of time (e.g. There are gardens and in that garden some children play)	Use of 'that marker' as relative pronoun in making assumptions about future of what they are not very sure about (e.g. There is a fort in the sea that we may see)
		Use of 'that clause' after a reporting verb (say, tell) to refer to what one had said in order to keep the conversation going one (e.g. I forgot to say that there are so many basic facilities available)	Use of 'that clause' after a reporting verb (say, tell) to refer to what one had said and link conversation to the task (e.g. Our teacher told us that all students together need to decide the place)
		Use of a question inside a statement to ask or make a request in a polite manner (e.g. We can but what transportation is available for Kolhapur?)	Use of a question inside statement to ask or make a request in a polite manner (e.g. We can but what transportation is available for Kolhapur?)

(continued)

Table 15.11 (continued)

Task A	Task B	Task C	Task D
			Use of 'that' as conjunction to give reason/support the choice (e.g. There is a so pretty story so that we have to go here)

Secondly, a variety of verb types and their respective uses can be incorporated for the four tasks as suggested in Table 15.12.

Table 15.12 Assessment criteria for use of types of verb complementation

Task A	Task B	Task C	Task D
Use of intransitive verbs to talk about: Action and events (e.g. Sun flashed)	Use of intransitive verbs to talk about: Action and events (e.g. Two boys are dancing) Physical behaviour (e.g. All the children participated)	Use of intransitive verbs to talk about: Action and events (e.g. We all play)	Use of intransitive verbs to talk about: Actions and events (e.g. We will sail in the boat)
Position (e.g. some girls are standing near the table)	Position (e.g. Some boys are sitting out of the ground)	Direction (e.g. I am coming)	Direction (e.g. I went there)
Physical behaviour (e.g. One boy is talking)	Direction (e.g. The boys are running towards finish line)		
Use of transitive verbs to talk about: Events in the picture (e.g. The people are waiting for the train)	Use of transitive verbs to talk about: Events in the pictures (e.g. Some girls are filling bottles with juices for the sports competition)	Use of transitive verbs to talk about: Events (e.g. We organize sports competition in new school)	Use of transitive verbs to talk about: Events (e.g. There are so many activities like riding a horse)
	Human as participants/emotions (e.g. They are cheering up the boys)	Human as participants/emotions (e.g. They save girl child)	Human as participants/emotions (e.g. All our classmates will be happy)
		Reporting verbs to ask for and give information (e.g. You said that there is no cleanliness)	Reporting verbs to ask for and give information (e.g. Many people say that Mahabaleshwar is very beautiful)

Using both of the linguistic (micro) categories—*syntactic knowledge* and *lexical variety*—has given us a detailed and rich explanation regarding the positive effects of using task factors in speaking assessment. The patterns of linguistic growth can be explained by the fact that carefully designed tasks help learners pay attention to task-specific linguistic aspects of form and meaning in a compositional manner. This can form an important part of task-specific assessment criteria. Similarly, while giving feedback, to direct learner attention to such linguistic aspects and guide them attend to their existing linguistic repertoire, it is essential to consider 'task complexity' and 'task condition' factors in TBSA.

However, teachers need to note that a limitation of this estimate is that these do not capture the interactive nature of learner performance; nevertheless, these are basic building blocks of language proficiency and can be used to evaluate learners' ability at an early stage and give feedback to help pay attention to form and its variety to improve content through communicative language use.

In sum, the major contribution of using task-specific performance into the design of a more nuanced and systematic evaluation criteria is that it can serve to make assessments effective by (a) rewarding learners if they attempt more complex responses by using improved language (with a rise in complexity and fluency) and (b) using the criteria to give feedback such that learners can get an estimate what they have done well and which features they need to improve. That would create the atmosphere for setting up a truly formative assessment model.

Conclusion

The example of TBSA presented in this paper is to highlight that task complexity can have a positive impact on assessment practices of both the stakeholders—learners and teacher–assessors (Popham, 2008). To make it happen, there is a need to make teacher–assessors aware of the benefits of aligning task design with complexity factors. The TBSA model is a part of formative assessment. It is structured based on Robinson's Triadic Componential Framework and is used to select/identify 'task complexity' and 'task condition' factors in speaking assessment. It is to help teachers assess learners' linguistic abilities at different levels of learning and to provide feedback.

The paper has presented an example to show how ESL teachers can use 'task design factors' to tap on task-specific linguistic gains and through the design of task-specific assessment criteria. In addition to this it is also possible to design task perception checklist to make both teachers and learners reflect on each task and their demands. Furthermore, to make them realize their abilities, the teachers may use task-specific self-assessment criteria. Self-generated feedback using assessment criteria provides information about the strengths and weakness and scope to improve. Such criteria can reflect a gradual increase in task requirements to tap the developmental second language knowledge of Indian learners. Finally, the paper suggests the incorporation of Task-Based Speaking Assessment to bring the education policy

recommendations of the recent *National Education Policy* in practice, i.e. a shift from rote memorization to practising formative assessment in promoting learning and testing higher-order skills (2020, p. 18).

Acknowledgements I am grateful to my Ph.D. supervisor Professor Lina Mukhopadhyay for helping me to develop the paper and to Professor Jacob Tharu for giving me insightful feedback on the paper. I also thank my colleague Sreelekshmi for the help with drawing the pictures and the learners who participated in the study.

References

Ahmad, K., & Agarwal, P. (2018). The role of CBSE ASL in enhancing the speaking and listening skills of students. Explorations: Teaching and Learning English in India. *British Council India, 8*, 13–28.

Bachman, L. F. (2002). Some reflections on task-based language performance assessment. *Language Testing, 19*(4), 453–476.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1). Oxford University Press.

Brown, H. D., & Abeywickrama, P. (2010). Principles of language assessment. In *Language assessment: principles and classroom practices* (2nd ed., pp. 25–51). Pearson Education.

Brown, J. (2004). Performance assessment: Existing literature and directions for research. *Second Language Studies, 22*, 91–139.

Byram, M. (Ed.). (2004). *Routledge encyclopedia of language teaching and learning*. Psychology Press.

Cambridge English: Young Learners English Test, Flyers Part 1, Volume 1, Sample papers, 2013. Cambridge University Press.

Cambridge English: Young Learners English Test, Flyers Part 2, Volume 2, 2014. Cambridge University Press.

Central Board of Secondary Education. (2013). Introduction: Assessment of speaking and listening skills in English–ASL. In academic circular (27/2013*) Guidelines for assessment of speaking and listening skills for classes IX to XI and training of teachers* (pp. 1–28). Retrieved from http://cbseacademic.nic.in/web_material/Circulars/2013/27_ASL_for_Teacher_s_%20Book.pdf

Chalhoub-Deville, M. (2001). Task-based assessments: Characteristics and validity evidence. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 210–228). Longman.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press. http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf. Retrieved 23 Jan 2011

Durairajan, G. (2016). Assessing and testing learners' language proficiency. *Language and Language Teaching, 5*(1), 52–56.

Earl, L., & Katz, S. (2006). *Rethinking classroom assessment with purpose in mind: Assessment for learning, assessment as learning, assessment of learning*. Governments of Alberta, British Columbia, Manitoba, Northwest Territories, Nunavut, Saskatchewan, and Yukon Territory: Western and Northern Canadian Protocol for Collaboration in Education.

Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2020). *Task-based language teaching: Theory and practice*. Cambridge University Press.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Erling, E., Ashmore, K., & Kapur, K. (2016). *Supporting language learning through formative assessment.* TESS-India. https://www.open.edu/openlearncreate/pluginfile.php/134914/mod_resource/content/4/SE13_AIE_Final.pdf

Fulcher, G. (2003). *Testing second language speaking.* Pearson Education Limited.

Fischer, J., Chouissa, C., Dugovicova, S., & Virkkunen-Fullenwider, A. (2011). *Guidelines for task-based university language testing*. Council of Europe.

Hornby, A. S. (2015). *Oxford advanced learner's dictionary*. Oxford University Press.

Ishikawa, T. (2011). Examining the influence of intentional reasoning demands on learner perceptions of task difficulty and L2 monologic speech. In P. Robinson (Ed.), *Researching the cognition hypothesis of language learning and performance* (pp. 308–329). John Benjamins Publishing Company.

Kapur, K. (2008, July). *Assessment for improving learning in schools in India: A perspective* [Paper Presentation]. IAEA, Cambridge, UK. https://iaea.info/documents/assessment-for-improving-learning-in-schools-in-india-a-perspective/

Friderick, K. (1984). *Keep talking: communicative fluency activities for language teaching.*

Kuiken, F., & Vedder, I. (2011). *Task complexity and linguistic performance in L2 writing and speaking.*

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation.* University of Chicago press.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). John Benjamins.

Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). Cambridge University Press.

Maharashtra State Board of Secondary and higher Secondary Education, Pune: 411004, W.E.F./ Research/ Revised syllabus / 1973 Date: 16/03/2006

Morrow, K. (1979). Communicative language testing: Revolution or evolution. *The Communicative Approach to Language Teaching*, 143–157.

Ministry of Human Resource Development: Government of India. (2020). *The National Education Policy-2020.* New Delhi.

Norris, J. (2001). Identifying rating criteria for task-based EAP assessment. *A focus on language test development: Expanding the language proficiency construct across a variety of tests* (pp. 163–204).

Norris, J., Brown, J., Hudson, T., & Yoshioka, J. (1998). *Designing second language performance assessments.* University of Hawai press.

Oller, J. W. (1979). *Language tests at school*. Longman.

Popham, W.J. (2008). *Transformative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Prabhu, N. (1987). *Second language pedagogy.* Oxford University Press.

Prabhu, N. (2019). *Perceptions of language pedagogy* (Durairajan, G., ed.). Orient BlackSwan.

Ramanathan, H. (2008). Testing of English in India: A developing concept. *Language Testing, 25*(1), 111–126.

Ratnam T., & Tharu J. (2018) Integrating assessment into classroom instruction to create zone of development for teachers and learners: Some perspectives from India. In H. Jiang & M. Hill (eds.), *Teacher learning with classroom assessment*. Singapore.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning, 45*, 99–140.

Robinson, P. (1996). Task-based testing, performance-referencing and programme development. *The University of Queensland Working Papers in Language and Linguistics, 1*, 95–117.

Robinson, P. (2001a). Task complexity, cognitive resources and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287–318). Cambridge University Press.

Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*, 27–57.

Robinson, P. (2011). Second language task complexity, the cognition hypothesis, language learning, and performance. In P. Robinson (Ed.), *Researching the cognition hypothesis of language learning and performance* (pp. 3–37). John Benjamins Publishing Company.

Robinson, P., Cadierno, T., & Shirai, Y. (2009). Time and motion: Measuring the effects of the conceptual demands of tasks on second language speech production. *Applied Linguistics, 30*(4), 533–554.

Sinclair, J. (2011). *Collins co-build English grammar* (2nd ed.). Harper Collins Publisher.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Tamboli, V. (2017). *The role of task complexity and task conditions in ESL speaking assessment*. [Unpublished doctoral dissertation]. The English and Foreign Languages University.

Weir, C. J., & Wu, J. R. (2006). Establishing test form and individual task comparability: A case study of a semi-direct speaking test. *Language Testing, 23*(2), 167–197.

Chapter 16
Whole Text Reading Comprehension: An Application of Task-Based Language Assessment

Vrishali K. Patil

Abstract In ESL context, we often assess reading on the basis of a set of sub-skills, while assessment of whole text comprehension is not used very commonly. This chapter demonstrates how whole text comprehension is a useful construct to assess reading in a holistic manner whereby teachers can get a sense of the process of comprehension as well as the product of learning. The assessment of whole text comprehension is done by evaluating oral summary recalls of ESL learners. The construct serves as an application of task-based language assessment model. In this model, Robinson's Triadic Componential Framework can be added to select texts for summarizing. The paper illustrates in detail the cognitive mapping of a reading text, by using propositions which are later developed to create summary propositions that form a basis to assess oral summaries.

Keywords Task-based assessment of reading · Whole text comprehension · Summary task · Cognitive task complexity · Classroom assessment

Introduction

Reading comprehension is a complex and multidimensional skill involving both linguistic and cognitive processes. A set of sub-skills such as *word decoding, phonological awareness, word meaning identification, syntactic awareness*, *inference generation, comprehension monitoring* and *whole text representation* (Snowling & Hulme, 2005) contribute to the development of comprehension skills, while cognitive processes as *working memory span* and *attention* play a vital role in comprehension success (Just & Carpenter, 1992). In classroom assessment, the development of comprehension skills is evaluated based on prescribed texts and using a multitude of tasks like fill in the blanks, short answer type questions, essay type questions, reference to context and so on. For summative (end of term) assessments, similar assessment techniques are used, and they result in rote learning of content for

V. K. Patil (✉)
EFL University, Hyderabad, India

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_16

achieving high scores. Assessment of unseen passages is at times done when learners are instructed to identify or supply key words to solve fill-in-the-blanks type items. Note that through these tasks only a piecemeal assessment of one's reading abilities is attempted. Higher order cognitive skills involved in comprehension process do not find a place in such assessments. This creates a gap between learning objectives for a holistic development of reading skills and a mapping of learners' knowledge through assessments. One way to attend to this gap would be to use a process-based approach to the assessment of reading comprehension within the formative (during instruction) assessment framework.

The aim of this paper is to consider steps to assess reading in a holistic manner by adopting the task-based language assessment (TBLA) model where the processes of reading comprehension can be assessed as well as the product. Thus, this paper is organized in the following manner: it begins with validating the use of whole text comprehension and establishing a rationale for using summary as a TBLA task to assess knowledge of whole text comprehension. Thereafter, it illustrates the use of systematically developed task and assessment criteria for holistic reading assessment. It concludes with a discussion of how the task can be used to assess '*individual differences*[1]', which, otherwise take a back seat in summative and large-scale reading assessments.

Whole Text Comprehension

A holistic approach to assess reading development can be attempted by examining learners' knowledge of 'whole text comprehension'. This is a higher order reading skill and can be assessed through free summary recalls either in the oral or written modalities. This model has been proposed by Lynda Taylor in her seminal work titled 'Testing Reading through Summary: Investigating summary completion tasks for assessing reading comprehension ability' (2013) and later applied by Mukhopadhyay in a study on classroom-based assessment (2017). To show evidence of whole text comprehension, learners need to build a mental representation as a mental map. For this, they need to understand

(i) the literal meaning of a text by attending to the surface linguistic features;
(ii) the inferential meaning from the links present between the ideas across paragraphs of a text by attending to latent meaning of the sentences;

and express their understanding

[1] 'individuals vary in the amount of activation of working memory they have available for meeting the computational and storage demands of language processing. This conceptualization predicts quantitative differences among individuals in the speed and accuracy with which they comprehend language. In addition, it is capable of accounting for some qualitative differences among readers' (Just and Carpenter, 1992: p. 124). Here, we address individual differences in recalls due to text complexity and text structure difficulty and the amount of learner effort required.

(i) as mental representations in terms of key ideas that can be linguistically expressed as summary propositions (Taylor, 2013: pp. 95–97).

In whole text comprehension, lower level sub-skills are not considered piecemeal as all of them are required to perform on this higher order skill. This becomes a holistic method of assessing the mental representations of learners as they attempt to comprehend and reconstruct the whole text in their mind and express it linguistically based on a series of summary propositions or clausal units drawn from a set of propositions (or clausal units) that hold a key idea together and is usually represented as a paragraph. This method becomes a reliable tool of holistic reading assessment as it encompasses the *entirety* of comprehension.

Building on Taylor's work, this paper presents a cognitive model to assess whole text reading comprehension by situating it within the framework of TBLA. Thereafter, Robinson's framework of task complexity (2011) is applied to demonstrate how to select texts and what are the steps in designing assessment task and criteria for doing whole text reading assessment. While the ideas presented in this paper primarily address assessment of reading, they are equally applicable for pedagogical purposes and can be used to teach whole text reading comprehension and give meaningful feedback to ESL learners. In fact, Taylor has proposed that teachers can be trained to create summary propositions to assess summary recalls in a transparent and systematic manner (2013: p. 218). Furthermore, building on Taylor's work that uses Van Djik's cognitive model to identify summary propositions (SPs), this paper in its last section proposes a model to explain how a teacher can deal with alterations to SPs as individual differences are likely to occur in the recalls.

Summary as an Application of TBLA to Assess Whole Text Comprehension

Before discussing the features of summary as a valid reading comprehension assessment tool, let us look at the frequently used traditional ways to assess reading. This will help in understanding how summary as a contrast to the above ways can capture the mental processes of reading a text for meaning. In Table 16.1, applications of various types of reading assessment are presented.

We can observe that different task types of reading assessment make different demands on working memory, attention, content knowledge and linguistic competence. For example, tasks such as multiple choice questions (MCQs), cloze test and short answer questions assess reading as a product. In contrast, a free recall summary task can assess and provide an evidence of learners' understanding of texts as a whole: it looks at reading comprehension as a process and then a product. It offers an appropriate format for assessing reading comprehension because it captures the cognitive demands in holding 'mental representation of a whole text' together in the short-term memory.

Table 16.1 Techniques in measuring reading comprehension

Reading assessments	Task type	What does it (may) measure	Approach	Scope for assessment
Cloze test	Fixed response	Semantic and syntactic knowledge	Product	It measures only surface linguistic forms
Multiple choice questions	Fixed response	Possibility of measuring partial cognitive understanding	Product	The accuracy of responses may not be interpreted as purely cognitive; at times exact interpretation is not possible
Short answer questions	Limited response	Possibility of measuring and interpretation of cognitive understanding	Product	The interpretation could be subjective if a detailed key is not available. Preparing key for numerous possible and acceptable answers is challenging
Free recall summary task	Open-ended responses	Whole text construction	Process	A key made for content assessment by creating summary propositions is essential, though challenging

As whole text comprehension can be reliably assessed through free summary recalls, let us look at the benefits of this model of assessment. This assessment model can be considered as an offshoot of TBLA where real-world behaviour is elicited through a carefully chosen set of texts that differ from each other in levels of complexity and in variations in communicative use of language functions (Norris, 2016, p. 233) and/or to capture different levels of ESL learners.

The rationale for using free recall summarizing task is that the task and its related text comprehension share two critical operations: (i) identifying main theme (key ideas) of a text, and (ii) judging the relative importance of text ideas based on their relevance to their main theme (Koda, 2005: p. 173). Careful reading requires understanding and identification of the micro- (e.g. *Vinay is a dentist*.) and macropropositions (*More than a dentist Vinay is famous as a wild life photographer.*), their interconnectedness in the text and an integration of old to new information to hold information within the mental model and finally create a well-knitted discourse (or the summary) (Khalifa & Weir, 2009; Taylor, 2013). All these aspects of reading are well captured in a summary task. Furthermore, this is a less intrusive technique to assess the reading process (Taylor, 2013: p. 219). It also makes visible the strengths and weaknesses of a reader's understanding (Patil, 2020). Hence, summaries could be an apt measure for understanding and assessing text representation knowledge

of ESL learners. Here, a teacher can give periodic feedback and take necessary steps in her teaching to ensure that learners' reading development are based on her observations from such formative assessments.

Assessing Reading Using Summaries

For academic assessment purposes, a variety of summary tasks can be used. They can range from fixed response to limited response and free response formats as depicted in Table 16.2:

At the base of all these tasks, an assessor has to identify and use the list of key ideas and links between them. Thereafter, the key ideas can be presented within a cloze task (multiple choice or supply type) as discrete point assessments or for a free summary recall as a performance-based assessment.

Researchers have tried to answer the questions like—how can a text summary be generated? What is the minimal unit of analysis that will lead to the creation of a logically built text summary? Research shows that (i) propositions, a linguistic representation as mono clausal units, and (ii) summary propositions (SPs), based on key ideas (Field, 2004, p. 225; Van Dijk and Kintsch, 1983pp. 109–144), are validated units of analysis in summary studies. Furthermore, researchers have used summary tasks to examine a variety of capabilities as listed below:

1. the awareness of *causal relationship* and its effect in oral recall in L1 context (Zinar, 1990);
2. the relationship between the *serial position* of a key idea in terms of its *primacy* and *recency* effects in a text, the related importance and probability of the presence of key ideas in the free recall in L1 context (Freebody & Anderson, 1986);

Table 16.2 Types of summary tasks in academic assessment

Type	Choice type	Modality	Assessment	Purpose of reading (Revised Bloom's Taxonomy, 2002)
Fixed response	Summary cloze task (fill in the blanks: MCQ)	Reading	Score	Understand and analyse
Limited response	Summary cloze task (supply type)	Reading + writing	Key of phrases	Understand, analyse and integrate
Free response	Free summary	Reading + speaking/reading + writing	Assessment rubric with key of SPs	Understand, analyse, integrate, evaluate and create

3. children's strategies for processing informational text to understand and remember new information through oral summaries (Cote et al., 1998);
4. the nature of retold meanings in narrative and expository texts using oral summaries (Kucer, 2011a);
5. the usefulness of immediate oral recall protocol for testing reading comprehension in L2 context (Bernhardt, 1983);
6. the effect of test method using global response modes of summary and the recall protocol (Riley and Lee,1996);
7. the basic condensation rules employed by children and adults when they summarize expository texts (Brown & Day, 1983);
8. strategic differences in awareness of task demands in summarization task, ability to identify important elements and ability to transform text into its gist between poor and good learners (Winogard, 1984);
9. the characteristics of EFL readers' summary writing with respect to ability to distinguish important ideas from less important ones in L2 texts and application of rules of summarization during summary writing (Kim, 2001);
10. the extent of transfer of rules of summarization from first language to a second language (Corbeil, 2010);
11. summary completion task for assessing reading comprehension ability using a key consisting of summary propositions (Taylor, 2013).

The findings from all these studies can be summarized with common threads of knowledge. The studies show that in terms of knowledge representation learners display application of various higher-level cognitive factors, starting from

- *identifying* a text base or a generic macrostructure (e.g. story grammar, expository text structure);
- for every text *applying* the knowledge of *identifying key ideas* or SPs*, links* between the ideas at sentential and paragraph levels;
- *selecting* the significant ideas for understanding and *building mental representation,* and while doing so;
- *super-ordinating* these propositions with their knowledge of text type, text structure, global experiences;
- *reorganizing* them to *create* plausible summary propositions and
- *deleting* certain propositions in the process.

These findings depict that during summary recalls a reader applies the above-listed *rules of summarizing* to hold the mental representation in the working memory and short-term memory. Thereafter, the reader expresses this knowledge by reconstructing the text as 'summary recall'. These findings are significant for a classroom teacher to understand the reading development for her learners in terms of their ability to use these higher-level cognitive factors in their academic performances.

It is important that the tasks used in testing conditions must generate interactional authenticity, meaning the language use should resemble real-life language use. In other words, the abilities used in test performance should be involved in a way similar to real-life contexts. This is also the basis of TBLA model. In this regard,

creating summary is a real-life task especially in the academic context and goes a long way to fulfil the criteria of purposeful assessment within the TBLA framework. A summary task (either oral or written) thus qualifies on these aspects of TBLA and also validates itself as a cognitive assessment tool for reading comprehension.

Aligning Whole Text Reading Comprehension to Cognition Hypothesis

The design of summary tasks to assess whole text reading comprehension would need a selection of a set of texts. This is a requirement of the TBLA model to reading assessment. In such a context, the set of texts would have to be sequenced along a continuum of complexity to yield linguistic and content development benefits. In the selection of texts, along the complexity gradation, we now propose to apply Robinson's Triadic Componential Framework (Robinson, 2001a, 2001b, 2003, 2005, 2011) because it gives a clear set of guidelines to choose tasks (here texts) on graded scale of difficulty that is likely to push ESL learners for target language development.

Robinson's framework distinguishes between two dimensions of *task complexity* as two types of demands: cognitive–conceptual (i.e. resource-directing) and performative–procedural (i.e. resource-dispersing). Robinson predicts that *increasing complexity along resource-directing dimension has the potential to direct learners' attention to both complexity and accuracy* (Robinson 2001a, p. 294). Resource-directing tasks are those in which the demands on language use made by an increase in task complexity can be achieved by manipulating the manner in which the information is presented. As resource-directing features of a cognitively complex task direct learner's attention to content, which means conceptual demands direct L2 learner's attention towards different linguistic forms. Consequently, resource-directing cognitively complex tasks may result in an L2 output with a higher structural and lexical variety and complexity because of the increased cognitive demands. The six resource-directing factors in the framework are [± here and now], [± few elements], [± spatial reasoning], [± causal reasoning], [± intentional reasoning] and [± perspective taking] (Robinson & Gilabert, 2007, p. 163–7). With use of these complexity features, the selection of reading texts could be manipulated to increase complexity in a graded manner that could result in improved performance towards structural and lexical variety and complexity.

Similarly, tasks can be increased on complexity along 'resource dispersion' dimensions to make extra resource demands, which cannot be met by using any specific linguistic system. Resource dispersion variables include manipulating task complexity along pre-task planning time [± planning time]; tasks with background knowledge [± prior knowledge]; [± single task]; [± task structure]; [± few steps] and [± independency of steps] (Robinson & Gilabert, 2007, pp. 163–7). Increasing complexity along these resource dispersion dimensions replicates processing conditions under which real-time language is often used. So, incorporating some of the

resource-directing and -dispersing conditions in text selection will help to design better summary tasks because teachers would get an estimate of the complexity features present in each text and predict learner task completion success accordingly.

Applying Cognition Hypothesis in Text Selection

To better understand cognitive processing underlying whole text comprehension, we propose applying Robinson's Cognition Hypothesis. This is primarily done in (a) selection of reading texts and in (b) design of criteria to assess summary recall performance. We need to understand the cognitive aspects of complexity that each of these elements offer to the construct of reading assessment that can be used for formative classroom-based assessment. For that, it is essential for a teacher to first understand *how* and *which* inbuilt aspects of texts contribute to the cognitive load while creating mental representations of texts.

The prediction of the cognition hypothesis gives us a possibility to look for the complexity feature of 'text' as a task for learning as well as assessment purposes. A teacher would need to know how task features could be manipulated around the dimension of task complexity both in selection of texts and assigning procedural conditions to accomplish the summary task. In class, where usually the teacher works with prescribed texts, manipulating task complexity of each text could be a cumbersome task for him/her. However, new texts could be selected as authentic texts at certain complexity levels for desired holistic performance. Alternatively, prescribed texts could be analysed for complexity features. In all, a teacher can use this knowledge to diagnose the difficulty areas or bottle necks in learner performances.

Text Structure as a Resource-Directing Feature

The way a text is structured depends on the purpose of text formation and author's argumentation. Text structure refers to how idea units or propositions are hierarchically organized to convey the desired message to the readers. Depending on the purpose of reading, text structures are classified into two broader types, viz. narrative and expository. Text type and/or a structure of a text can significantly impact reading comprehension performance of test-takers and at different proficiency levels (Kobayashi, 2002). To achieve construct validity of reading comprehension tests, the most frequently naturally occurring text types, narrative and expository are frequently selected (Weigle, 2002; Khalifa & Weir, 2009; Kucer, 2011a; Taylor, 2013)[2]. If a teacher trains her learners to identify text structures features (e.g. story grammar for

[2] Even in the Indian examination system, these two text types are frequently used for assessing reading comprehension. An analysis of reading comprehension section of three question papers of grade VIII, IX and X shows that the percentage of use of expository texts (55.55%) in unseen

narrative texts and exposition-based macrostructure features) with respect to their semantic and syntactic features, then the learners would get an insight into why a particular text structure is used and comprehension may become better.

Text Internal Rhetoric as a Resource-Directing Feature

In addition to the structure- and purpose-related characteristics of text types, cognitive complexity of any text may rest within the structural complexity of specific text. Within the broader text types, there are various internal rhetoric structures, which arise due to the causal mechanisms, lexical links and rhetorical predicates. This can become significant in terms of creating mental representations in the minds of readers and impact their summary construction. With respect to the assessment of structural and linguistic features, research has proved that narrative texts are easier to comprehend than the expository texts (Kintsch, 1982; Kim, 2001; Kent, 1984; Kucer, 2010, 2011a, 2011b). However, Lynda Taylor's study has proven that when mental mapping is assessed, difficulty is irrespective of the task type because it is the internal rhetoric which gives rise to several levels of comprehension difficulty. If we analyse prescribed texts for task difficulty, we would find that it cuts across genres. It is not the text type as such that makes a text difficult, but it is the additional presence of the complexity features of internal rhetoric that leads to the difficulty in building a text representation. For a teacher, an understanding of text features is significant for cognitive complexity and is crucial and with respect to two aspects: one its generated structure (text type) and two its internal rhetorical structure. This would be useful in teaching as well as in classroom assessments as it would give her a chance to understand which complexity features pose difficulties for her learners and how they give rise to individual differences. She can also gauge and prepare feedback and teaching strategies to help her learners.

A Sample Analysis of Text for Cognitive Complexity

In TBLA, learners' abilities to attend to linguistic requirements of a task in a comprehensive manner in terms of accuracy, complexity and fluency are assessed (Skehan, 1998; Norris, 2016, p. 241). But when a text is complex in terms of its propositional representation, it can create certain cognitive gains for a learner's reading comprehension ability. This gain is in the form of a leap from using general language aspects to using language to express whole text comprehension either as an oral or written summary. Though expository texts are difficult to process than narrative texts, one can question its applicability to all the texts falling under each of these two sub-types.

---

passages is more than that of the narrative texts (44.44%) for all the three grade exams (out of 27 passages from nine question papers—i.e. three question papers for each grade).

Texts differ according to individual macrostructure of type or genre (Meyer & Rice, 1984) and internal text-specific rhetoric (Meyer, 1982; Winogard, 1984; Freebody & Anderson, 1986)? If 'an effort' load of a text to perform is created through an increment of cognitive load, then what role does cognitive complexity play in naturally occurring reading comprehension texts? Do authentic texts need to be manipulated according to task complexity features? Or is there a way to understand such features in authentic texts that can be inferred based on learner performances of summaries of whole text reading comprehension? To find answers to these questions, the first step is to analyse the complexity of reading texts along complexity features.

The presence of task complexity features across texts can be identified based on the text, typically its macrostructure and text-specific internal rhetoric structure. Texts can be analysed for task complexity as resource-directing dimension. The resource-directing features of task (here text selection) complexity are now analysed based on the features of a sample text (refer to Appendix 1 for the text) and are given in Table 16.3.

In Table 16.3, in addition to analysing to the linguistic–cognitive features that add to resource-directing features, we have also added resource-dispersing elements that a teacher can plan and add to the summary task to make it further graded with additional levels of difficulty. So, if the task is done [+planning], the results yielded would be different as would be the case if the text is [+prior knowledge] of (i) a familiar theme and/or (ii) the type of summary recall (one text-based summary vs. review summary arising out of a few thematically linked texts).

Assessing Text-Based Summary Recalls

We have discussed the task complexity features that can help teachers in selecting texts with respect to their complexity features. This is a primary requirement of the TBLA model. We now need to consider how to assess summary recalls in a systematic and transparent manner. This is because summary recalls are open-ended responses and are likely to give rise to individual differences. So, a teacher needs to know how to assess such responses and give meaningful feedback. In this regard, a first step would be to assess the content of the summaries. This can be done by listing 'summary propositions' which would be the content key to assess performances. Additionally, teachers can also look for structural complexity and accuracy; but we do not discuss the implementation of these features in the assessment criteria in this chapter.

Creation of Summary Propositions (SPs)

Summary propositions (SPs) are a set of higher-level propositions where several text propositions occur in a sequence to form 'an integrated high-level macrostructure

Table 16.3 Task complexity features of a text: 'Self-Reflection'

S. No.	Cognitive features for task complexity	Presence	Frequency of occurrence of idea units	Examples from text
A	Resource-directing features			
1	± here and now	+	12	Watch for the times when you complete…, sometimes you are reminded…, find a quiet place
2	± few elements	+	37	Mad rush of one subject after another, what you are doing, what you are learning, self-reflection
3	Spatial reasoning	+	2	Find a quiet place, at your desk
4	Causal reasoning	+	6	So that new information can be quickly retrieved, because there is some kind of test
5	Intentional reasoning	–	–	–
6	Perspective taking	+	5	If you are going to take notes…, if you have never taken the time to reflect, try it now
B	Resource-dispersing features			
8	Planning time	–		
9	Prior knowledge	+	1	Self-reflection ability in study
10	Single task	–		
11	Task structure	+	1	Reflective writing, explanatory

which gives global meaning for the more specific or local sequences in the text' (Taylor, 2013: p. 95). In other words, a paragraph might contain several propositions around a key idea. These can be represented by one or two SPs. The rationale for using SPs is that in determining accurate and coherent summaries, SPs can be matched with most participants' common mental representations of texts. A set of SPs from each text then can become 'a key' against which learner responses for that text can be systematically assessed. This also ensures that interrater reliability is maintained by controlling an assessor's subjective impressions.

Selecting SPs

SPs include main ideas of a text and the links between those ideas. Following Taylor's study (2013, p. 99) in the creation of SPs a teacher, can be guided by (a) the orthographic paragraph (b) structuring of the text and (c) main identifiable points of topic shift. This can help in the creation of an optimum set of SPs that can

i. accurately reflect the content of the original text;
ii. include elements/propositions of the original text which most readers would find salient (and also the elements which might cause difficulty for poor readers);
iii. form a coherent and standalone text in itself and
iv. be long enough to generate a sufficient amount of structurally interdependent sentences (Taylor, 2013: p. 212).

Note that (i) and (ii) constitute the semantic rules for drawing a summary while (iii) and (iv) constitute the structural rules for creating the summary. Thus, SPs of a text as a content-based key for assessing whole text comprehension become a valid tool for evaluating learner responses.

Individual Variations in Summary Recall

When learners read and reconstruct whole text summaries by *applying rules of summarizing*, due to *internal rhetoric of a text structure,* they can exhibit certain patterns as *individual differences* in the form of variations in constructing SPs to express whole text representation knowledge. This happens because every learner builds unique text representations with some common features of structural representation from a text. In summary recalls, a teacher may find a set of common trends across learner recalls. However, since this whole process is a matter of text reconstruction, there would be variations in recall representing various stages of learning and use of working memory constraints. Teachers need to address and provide explanations for such variations if the feedback on reading performance is to create further learning opportunities within the formative assessment model. These should not be treated as 'errors' as these are not random; but the variations are instances of systematic reconstructions of the texts based on identifying key ideas and their structural–cognitive links. Based on the finding of a study on whole text representation as oral summary recalls of Indian ESL learner in grades 7 and 8 (Patil, 2020), we have identified four types of alterations to assess individual differences. These are listed in Table 16.4.

These alterations can be assessed vis-a-vis the key of SPs created for each text. Based on this discussion, we attempt to work with three important conclusions:

1. Young ESL learners' demonstration of knowledge of whole text representation of academic texts can be assessed in terms of cognitive and linguistic difficulties

Table 16.4 Definitions for alterations in summary recalls

Type of alteration	Definition	Linguistic/Cognitive complexity feature attributed [resource directing]
Addition	Propositions that are coherent additions to the existing SP knowledge. These are the examples of inferences generated to form new SPs or propositions—which are neither the SPs from the key list nor the substitutions of SPs in the key list, but are new summaries or new ways of presenting text-based propositions	Inference making—lexical, pragmatic and propositional because of [± intentional reasoning] [± causal reasoning] [± perspective taking]
Substitution	Propositions that are coherent but distinct replacement of key ideas/content. (with inference generation ability)	Inference making—lexical, pragmatic or propositional and/or lexical divergences because of [± here and now] [± spatial reasoning] [± intentional reasoning] [± causal reasoning] [± perspective taking]
Deletion	SPs from the recall which are not mentioned in the key are called deletions. (i.e. the SPs not present in the knowledge representation by the learners); at times there can be partial deletion of a SP	Redundancy, elaboration of content, clausal complexity because of [− here and now] [− spatial reasoning] [− intentional reasoning] [− causal reasoning] [− perspective taking]
Distortions	Propositions that are coherent and meaningful alterations but thematically incorrect (Grammatical inaccuracies are not considered here.)	Idea units are not related to the content

related to the target language use (TLU) domain (Bachman & Palmer, 1996: p. 59) of the group (Patil, 2020).

2. Individual differences of learners in whole text comprehension can be addressed by considering a set of SPs that are built on generic macrostructure (text structure) and internal rhetoric of the specific text.
3. The key of SPs can be used to identify alterations in summary recalls (Patil, 2020; Taylor, 2013).

In the next section, we elaborate upon these three statements by providing a systematically developed holistic task which a teacher can use for assessing knowledge of whole text comprehension in the form of summary (either in the oral or written modality.

Designing a Classroom-Based Assessment of Whole Text Comprehension

According to the process approach to assess reading, whole text representation knowledge can be assessed as common knowledge of four factors: (1) text type, (2) internal rhetoric structure of texts, (3) chunking of idea units to create summary propositions and (4) applying generalizable summary rules to produce the summary text.

Higher order cognitive processes like *inference generation, chunking of ideas* and *establishing connective links* between the ideas to prepare plausible summaries are more **reliable** attempts than just 'product oriented' use of rules of summarizing for reading assessment (Taylor, 2013: pp. 60–63). Assessment of reading should be encouraged in such a way that both the cognitive processes of learning to arrive at the product and the product are assessed. In this way, a context would be created to support assessment *for* learning.

Since every learner is an individual with different learning abilities, it is of utmost importance that they are guided to develop their individual levels of learning. This features as individual inference generation abilities on the basis of which a learner can create text base according to what he/she can comprehend. *Learner effort* is one of the significant aspects of task-based learning (Prabhu, 1991) and *variance in actual learner efforts* may also cause the individual *differences/alterations* in production, and can be a window to understand why micro-variations occur in the process of learning to comprehend. Hence, it is important that the learners be assessed for their 'individual learning differences' along with the 'common knowledge' of summarizing, which will make the development of reading and its assessment truly 'holistic'.

A central claim made in this paper is that whole text reading comprehension can be assessed through process approach by using a summary task. The task is an example of TBLA and creates conditions for assessment *for* learning in young ESL learners. Based on this position that summaries are a way of assessing *for* learning to read holistically and in the form of whole text representation, we now present a complete summary assessment task to demonstrate how effectively ESL learners are able to respond to the sub-aspects of the whole text comprehension as interim phases of understanding before the whole text representation takes place.

A teacher can use this assessment following the guidelines listed below to create a holistic assessment tool. This can be used at several levels: (i) to test knowledge of text organization through knowledge of macrostructure, text-specific internal rhetoric structure and links between propositions; (ii) to test the ability to identify key ideas and arrange them on a hierarchical or rhetorical structure-based relationships through the use of linkers. The framework presented here can also be used by teachers to help learners identify connective links (either linguistically explicit or implicit)

between the key ideas; find out the ability of creating sequential[3] or synthesizing[4] summaries (Ratteray, 1985) and lastly to investigate their *common knowledge* and *individual differences* to tap growth. The list of following features should be taken into consideration while designing the task.

A summary task can be developed systematically by fulfilling the following characteristics:

a. make a stepwise selection of texts based on its type or genre, internal rhetoric structure, links between ideas and arguments; this can be done by analysing/adding resource-directing and resource-dispersing features;
b. give suitable instruction so that learners recall the content;
c. construct appropriate summary propositions as the assessment key to be used for assessment and later to for feedback; and
d. make a note of the *alterations* of SPs to make space for individual differences in performances.

The Triadic Componential Framework can be applied to this assessment task for text selection

1. first to understand the cognitive complexity features in the text and summary task;

2a. later to make related inferences from the learners' performance on the task and

2b. creation of task-based criteria of assessment, i.e. task features, whereby implied cognitive processes can be assessed.

Steps to Develop a Sample Summary Task

A teacher may follow the steps given below to design a reading assessment applying task complexity features from the Triadic Componential Framework.

STEP 1 Selection of texts

1a. Identify TLU level (a threshold level of proficiency at B1-B2 level of reading as specified in CEFR is required) and text domain.

An expository text titled *Self-Reflection—What does it Mean* (K12 Reader Reading Instruction Resources, 2012) (refer to Appendix 1 for the full text) is chosen. It is a

[3] Sequential summaries are those which retain the original order in which information is presented. These are very often used in the academic context as one kind of writing task at intermediate level (A2/B1).

[4] Synthesizing summaries are the ones which alter the original sequence of the text to achieve specific objectives and are mostly an academic reading writing task of a higher order usually for advanced level (B2/C1) EFL/ESL learners.

Table 16.5 Linguistic analysis of self-reflection

Text	Text length (no of words)	Frequency of clause type				Mean clausal length
		Simple clauses	Coordinating clauses	Subordinating clauses	Total	7
Self-reflection	420	4	14	25	60	

problem–solution text with *response rhetorical* relationships as it organizes the text around a set of questions and answers.[5]

1b. Analyse the text for its linguistic complexity.

To achieve construct validity, it is necessary to match clausal complexity of the selected text with the clausal complexity of prescribed texts in the textbook at this level. This kind of linguistic analysis can be done by a researcher and/or by a teacher. In Table 16.5, we give the linguistic analysis of the text *self-reflection.*

It is not possible for a teacher to do a detailed linguistic analysis of each and every text in the textbook without a training of how to do such analysis. But since texts are designed for a particular grade the teachers have the benefit of assuming that those belong to a level of linguistic complexity. However, when a teacher gives new reading texts to learners, then knowledge of linguistic and cognitive complexity of that text would require an analysis of text complexity. So, teacher awareness on linguistic analysis for task complexity features would be a crucial step to help them with text selection.

It is prerequisite that the teacher trains her learners on various text types, macrostructure elements and internal rhetoric structure every time a new lesson is taught so that learners are acquainted with diverse text rhetorical structures. This will help learners identify topic shift markers spread across the macrostructure of a text and establish links between the micro- and key ideas. This will result in training them to create relevant mental representations of summaries.

1c. Check for features of cognitive complexity.

For this, refer to Table 16.3 where the text on *self-reflection* is analysed for the presence of text complexity features. The text has presence of resource directing [+ here and now], [+ spatial reasoning], [+ causal reasoning], [+ intentional reasoning] and [+ perspective taking] features.

If the text is a part of the prescribed syllabus or a new text with a familiar topic, then resource dispersions will be low as [+ prior knowledge to theme] will assist in summary creation. In sum, the addition of a few resource-dispersing features adds to task complexity.

[5] Note that this is not a prescribed text but it has been matched with prescribed texts for matching TLU level, linguistic complexity and features of cognitive complexity of learners in Grades 7 and 8, studying in an English Medium Instruction School in India.

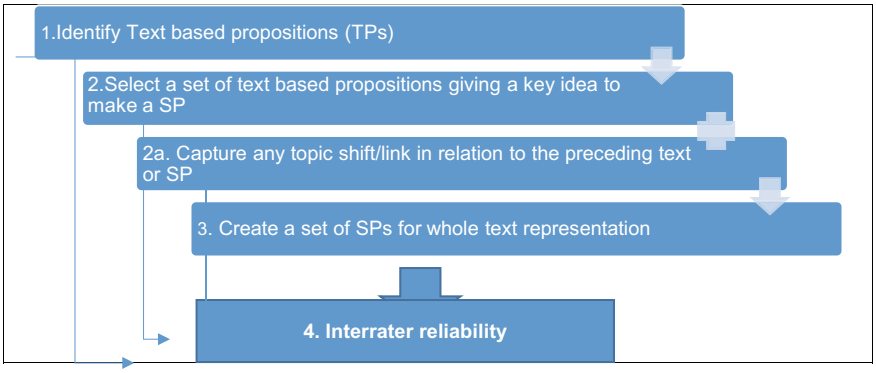


Fig. 16.1 Steps in creation of summary propositions

STEP 2 Prepare a list of lexical items (phrases) which might create difficulty in understanding meaning.

STEP 3 Identify chunking of text that hold the main idea.

STEP 4 Make a key list of text-based propositions (TPs). These are mono clausal units (refer to Appendix 1 with TPs marked in parenthesis and backward slashes (/)).

STEP 5 Design a list of summary propositions (SPs) from the text by taking into consideration what can be common knowledge about key ideas and rules of summarizing.

STEP 6 Validate the evaluation tool for interrater reliability by comparing the key of fellow teachers.

We present steps (3–5) diagrammatically in Fig. 16.1:

In Table 16.6, we present a sample of chunking of ideas from text propositions (TPs) of the introductory paragraph of the text 'Self-Reflection'. This will help teachers understand how to arrive at SPs (steps 3–5). The teacher can use this as a key for assessing summary performance. It is equally applicable for the oral and written modalities.

STEP 7 Make a checklist of task complexity features.

This is also called the table of task specifications. It helps to maintain construct and content validity of the assessment (Bachman & Palmer, 1996, p. 21; Brown & Abeywickrama, 2010: pp. 30–33). In Table 16.7, we present the specification of the task:

STEP 8 Design the method of evaluation using SPs.

The summary recalls of the participants are to be evaluated with the help of list of SPs for the texts set up as a key to assess performance (for a full list of SPs refer to Appendix 2). Grading and scoring indicate to students the importance of what is being assessed. A teacher may create a grade sheet (or a score sheet) as per his/her

Table 16.6 Derivation of TPs and SPs to prepare content assessment criteria

Excerpt from self-reflection text and chunking markers	Measuring Unit	TPs (single clause with single idea, addition of a noun phrase as a subject)	SPs (single idea arriving from many clauses or sentences, orthographic breaks)
(Do you ever feel like you are running from one activity to another/, or studying a mad rush of one subject after another?)/(When things happen at such a fast pace/, it is easy to lose sight of/what you are doing, /or what you are learning.)/(It is important to stop yourself every now and then/to think about/what you have been doing/or learning.)	punctuation marks linguistic markers answers to the questions given in the introduction macrostructure of the text topic shift markers	TP01: do you ever feel like TP02: you are running from one activity to another TP03: do you ever feel like studying a mad rush of one subject after	SP01: We feel the frenzy of doing things continuously
		TP04: when things happen at such a fast pace TP05: it is easy to lose sight of TP06: what you are doing TP07: or what you are learning	SP02: As a result, we are likely to lose focus in work
		TP08: It is important to stop yourself every now and then TP09: you think about what you have been doing TP10:you think about what you have been learning	SP03: At times we must stop to take stock of what we are actually learning from our work

Table 16.7 Specifications for whole text representation sample task

Task type	Output modality	Cognitive steps involved	Objectives
Free response- text removed	Oral or written recall	*Identifying* key ideas by means of applying *selection, generalization, reconstruction* (to accommodate SPs to generate alterations in terms of inferences)	• The ability of reconstruction • The ability to generate inferences (lexical, pragmatic and propositional) • Demonstration of *alterations* • Understanding text internal rhetoric

classroom assessment needs. Also, the teacher has to be mindful of the fact that the language of the SPs presented by the learners will be only near equivalent to the ones designed as plausible. There is a chance that learners may recall some of the text-based propositions depicting single ideas only. These propositions also should be recorded and assessed. So, a teacher needs to

(a) check recalled knowledge in terms of the matched ideas with key ideas of the text and the link between them; and
(b) identify 'individual differences' in terms of 'alterations'.

A classroom teacher is the best judge of her learners' cognitive–linguistic abilities, and this kind of assessment gives her a chance to understand and align her expectations as a teacher with individual differences of learners and create her own assessment estimate to best judge the learning process. She can make observations for diagnostic purposes as well as for feedback. This assessment is not a one shot assessment, but could be carried out as a series of assessments to create learning opportunities and compare differences arising due to the graded levels of text complexity.

It is also important to note that summary writing and recall task in academic assessment require a threshold level of language proficiency for learners to identify and organize information and integrate that into a well-knitted discourse structure, either written or oral. And since we are capturing the cognitive mapping of learners at threshold level proficiency, the modality of recall is expected to not have impact on their expressions of mental representation. Modality of a task can hamper the learning output in the tasks where language assessed is at a higher TLU domain than learners' current level. The task presented in the paper assesses idea appropriacy and not linguistic accuracy. Hence, the factors related to written and oral summaries, i.e. writing or speech proficiency, knowledge of spellings and grammar, etc., are not likely to affect the mapping of mental reconstruction of knowledge. A teacher can negate the possible effect of modality by designing a meaningful instruction for the task where learners would understand the purpose of the task as recreation of ideas. They would also be motivated to generate better content output by being internally guided with text complexity features and those would push them to produce the target language with a variety of linguistic structures. Considering all these points in the discussion, let us see the demonstration of use of instructions for such a task.

A Sample Summary Task for Assessing Whole Text Reading Comprehension

Read the following passage silently. Take as much time as you need to understand it.

(Once the learner is ready –>)

Oral Modality—Instruction:*Recall the details of this text in your own words. Telling everything that you remember. Do not worry about correct pronunciation or grammar.*

Table 16.8 Task-specific assessment criteria

Task features	Implied knowledge	Description
Content	Content knowledge	Key ideas: understands the concept of self-reflection, ways of self-reflection, reason to self-reflect steps involved in self-reflection, explains the example of school text book organization
Accuracy	Linguistic abilities/grammar	Some grammatical errors are present; but they do not make recall incomprehensible (look for subject verb agreement and tense features; referential pronouns; coordination and subordination of ideas; prepositional phrases
Identification of ideas, links, application of rules of summary construction, alterations	Cognitive processes	• Identification of TPs; construction of SPs • Applying selection and deletion rules to distinguish more important ideas from less important one • Ability to generate inferences and their sub-types in the forms of *addition, substitut*ion and *distortions* • Type of summary (sequential- or synthesis-based)

Written Modality—Instruction:*In the space given below, try to write down as much of the text as you can remember. You can write short sentences, or in notes. Try to write in your own words. Writing down everything you remember is more important than correct grammar and spelling. Take a new sheet if you run out of space.* In Table 16.8 we give the task-specific assessment criteria for teachers.

STEP 9 Assess individual differences.

Summary recall is not an imitation task; it is an interplay of knowledge at various levels constituting of generating inferences as well as application of general summarizing rules. These can be associated with what Prabhu (1991) has called as learners 'effort of deployment' which involves learners handling their world knowledge and abilities of inferring, reasoning, etc., using linguistic sources they command. Such variations in mental representations as alterations are the instances of 'cognitive

attempts' by learners to show 'growth' and complete the text base in a logical manner. Hence, a teacher–assessor should not aim only for uniformity in comprehension representation. Rather he/she should also be able to address to these '*individual differences'/alterations* as learners attempt to use higher-level cognitive processes to show reading development. Since alterations depict growth, it is essential to address those in assessment of whole text comprehension. Also, alterations are learner-specific; hence it is more work for a teacher at the beginning. However, within the formative assessment model, this would help a teacher to maintain cognitive graph of each learner in terms of his/her cognitive development in comprehension. Individual variations which are inevitable can be used to understand which complexity factors pose processing difficulty for learners, to give them feedback, to help them acquire whole text representation as higher cognitive process and to apply learning to new texts with a variety of internal rhetoric structures.

Deleting Ideas

The process of meaning making is not a uniform and linear process. A reader consistently makes decisions about what he/she considers significant and adds to the making of 'new' substructures and deletes what is insignificant in the processes of reconstruction and integration of ideas. What a reader may find insignificant to delete mostly consists of trivial information in a text and the redundancies (Kintsch & Van Dijk, 1978). This forms the rule of deletion while constituting summaries. Thus, understanding what is deleted is a significant step for a teacher as it helps in understanding the mental processes underlying the application of rule of deletion.

Adding Ideas

Creating additional propositions constitutes an alternative way of chunking ideas while reading a text. A reader might chunk using the links he/she thinks essential for making inferences, which could be regarded as a higher level of chunking as opposed to chunking of ideas while reading through a text for the first time. So, apart from the chunking done for arriving at summary propositions, a reader may apply his own strategies of deciding what is important in the text for chunking. Thus, in the process of knowledge recall, additions are evidences for an alternate pattern of chunking while reconstruction of whole text knowledge in summary recalls. Those are also inferences which learners make by applying world knowledge to the situation in the text. Hence, like deletion, a teacher also needs to address additions of individual learners in a similar way. This can be done by first distinguishing coherent additions from incoherent ones, secondly, analysing learner-specific inferences and lastly identifying the causes of difficulty with respect to the task complexity.

Substituting Ideas

Research has explained that *substitution* or *superordination* is a cognitive skill which occurs at specific levels or substructures of text comprehension in the process of recall and summarizing. Readers construct general categories to include specific details to maintain links between the propositions; but creation of macrolevel links also happens to maintain the conceptual framework of a text (Kirkland & Saunders, 1991). Thus, substitution plays a key role in attaining a conceptual framework for comprehension and is aided by working memory to arrive at a meaningful summary (Kintsch & Van Dijk, 1978). Substitutions are developmental strategies in creating a mental reconstruction of the text. Hence, it is crucial to study substitutions in the learners' recalls along with the deletions and additions. Substitutions are occurrences where lexical items as well as idea units are substituted to generate. Identifying substitutions and their inference types will help a teacher to analyse which features of task complexity lead to creation of substitution.

Idea Distortions

Distortions may be *coherent* and *meaningful idea units* but the ones that are *thematically irrelevant or non-coherent.* Some learners substitute the meaning with inferences in a coherent way but that does not match with the underlying theme/topic of a text. However, substitutions are different from distortions in a way that, though both are coherent and inference generating, substitutions are *thematically connected* to the text while distortions are *deviated* from the theme. So, distortions need not necessarily be linguistic errors in communications; but they are additional material that is not thematically coherent. A teacher has to identify distortions and analyse which complexity features might have led to learners' attention to the global knowledge rather than text-based ideas and the links.

Note that in each of these alteration types the inferences are guided by one common factor—the resource-directing complexity features which give rise to processing and storage difficulty and finally give rise to individual differences. While summarizing the content, learners make inferences of structures of *narration* or *information sharing* or *description of the text.* In contrast, inferences and critical comments are based on *evaluation* [+clausal reasoning] and *decision-making about the argument* [+perspective taking]. When as teachers we know which features may have added to task difficulty (e.g. by looking closely at the alterations), we can understand learner-specific needs and attempt to provide diagnostic feedback.

Pedagogical value of summary tool to assess reading

Though this tool of assessing whole text representation is high on cognitive validity and reliability, it may not be practical for use during summative assessment due to time constraints and lack of resources to ensure interrater reliability. However, this as

an assessment tool can surely be used for formative purposes and as an 'alternative assessment' to tap growth in whole text reading comprehension periodically. From this, a teacher can infer individual differences and plan her feedback to improve learners' comprehension abilities. This would also help her plan her the next set of reading and writing classes.

Whole text representation is versatile in the sense that it could be administered either in the oral or written or combined mode. Hence, it is readily available for a teacher once his/her criteria as a set of key SPs are designed. The same criteria could be used by all the teachers teaching for same grade making it practically a feasible and generalizable format of assessment. This is achievable by giving teachers proper training in assessing whole text reading comprehension.

Using whole text representation-oriented tasks, we can map the development of the process as well as the product of learning. Thus, this could be a valid learning tool of teaching, learning and assessing not only ESL/EFL as a subject but also content subjects where English is the medium of instruction. From content subjects, a variety of internal rhetorical structures can be worked with as they would provide a variety of structures to learners and would help them develop whole text reading in a holistic manner by developing knowledge of variety of text types and sub-types.

Further, this tool has a strong potential for positive washback as it taps the process of reading. It creates a possibility for an ESL teacher to move further from the common practice of only corrective feedback towards practicing constructive feedback. A teacher can understand cognitive challenges that learners are likely to face to do whole text comprehension. Based on this understanding, she can provide diagnostic feedback and improve classroom teaching and assessment practices.

Conclusion

This paper has highlighted the need for holistic assessment of reading in the domain of TBLA. It has illustrated that whole text representation is a comprehensive way of mapping cognitive growth in reading assessment. In doing so, the application of 'task/text complexity features' from Robinson's Triadic Componential Framework helps in a positive way to understand individual differences in learning. Thus, teachers need to have an awareness about (i) assessment *for* and *as* learning and (ii) the cognitive aspects of task design to add text complexity features. This would make classroom-based assessment of reading meaningful. They can make ESL learners aware of their stages of learning to comprehend texts by providing them constructive feedback based on text complexity features and summary propositions.

To conclude, the chapter recommends task-based reading assessment as an alternative and formative assessment for better learning and assessment of higher-order skills involved in reading comprehension, a crucial aspect of school and tertiary education. The task can be prepared with lessons for the English class as well as Science and Social Science classes at school level. It can give teacher an idea how to identify task-related cognitive difficulties and what effort individual learners need to

make on each task. This would lead them to provide feedback and transform teaching. It would also create an opportunity for learners to receive diagnostic feedback as they can compare their task performance vis-à-vis' the assessment criteria presented to them. It would be beneficial for teachers to examine multiple aspects of language ability at a certain TLU level with respect to growth in knowledge domains such as content development and skills of summarizing through inference generation and knowledge reconstruction. Alongside this, teachers can also look for development of variety, complexity and accuracy of grammatical knowledge and pragmatic knowledge of texture through a single performance. This assessment design is likely to have a high positive washback effect as it would make teachers reconsider teaching and assessing reading according to the process of comprehension that happens in the learner mind. Thus, the assessment task is equally suitable as a pedagogic task for developing whole text reading comprehension within the Indian school context. In fact, the very recent National Educational Policy visualizes "to close the gap in achievement of learning outcomes, classroom transactions to be shifted, towards competency-based learning and education; the assessment tools (including assessment '*as*', '*of*' and '*for*' learning) to be aligned with the learning outcomes, capabilities and dispositions as specified for each subject of a given class", (pp. 5, 13). It recommends the assessment of learners' knowledge of concepts and their cognitive abilities like critical thinking, analysis and application of knowledge with a focus on 'regular formative assessment for learning' (p. 5). Considering the descriptors of learning and assessment as given by NEP, whole text reading assessment could be a justifiable way to reach desired learning outcomes.

Acknowledgements I am thankful to my Ph.D. supervisor Professor Lina Mukhopadhyay for helping me to develop the paper and to Professor Geetha Durairajan, Dr. Sudharshana and Dr. Santosh Mahapatra for useful feedback on the earlier versions of the paper.

Appendix 1

Self-Reflection—What Does It Mean

(Do you ever feel like you are running from one activity to another/, or studying a mad rush of one subject after another?)/(When things happen at such a fast pace, /it is easy to lose sight of/what you are doing, /or what you are learning.)/(It is important to stop yourself every now and then/to think about what you've been doing/or learning.)

(Self-reflection means stopping the mad rush of activity/and calming yourself and your mind/so your brain can evaluate the input/it has already received.)/(Some people prefer to do their self-reflections mentally, /while

others keep a journal or written notes of some kind. /Either method can work, /depending on/what your own personal style is. /Whether written or purely mental, /the process is the same.)

(School textbooks are often divided into chapters and units of study. /This can make it easier for someone/who wants to begin the process of self-reflection to get started.)/(Watch for the times/when you complete a unit of study in any of the subjects/you are learning: math, social studies, science, art or any other topic/that might be specific to your school. /Sometimes you are reminded/that the unit is over/because there is some kind of test or quiz. /Use these natural breaks as opportunities to stop and reflect./) (Find a quiet place. /This can even be sitting at your desk at school/when you finish something early/and the other students are still working.)

(If you are going to take notes/, take out paper or your reflection journal. /Jot down some notes on things/that you learned in this unit/that you did not know/before you began. Let your mind ponder/on the notes you have written/and make some connections.)

(Sometimes just pausing to think deeply/allows your brain to make connections/so that new information can be quickly retrieved/when you need it again.)/(Next, think about things/that you still wonder/. For example, /maybe you learned about a certain body system, /but you're not sure how it works together with the other body systems. /Maybe you learned a new way to solve a math problem, /but you're not sure when to use it.)/(Writing down your questions will help you remember/to continue seeking answers/the next time you are exposed to the same topic.)/(Self-reflection is an essential skill for a successful student/. If you have never taken the time to reflect, /try it now.)

Appendix 2

Key of Summary Propositions (SPs)

SP01	We feel the frenzy of doing things continuously
SP02	As a result, we are likely to lose focus of our work
SP03	At times we must stop to take stock of what we are actually learning from our work
SP04	Self-reflection means evaluating what we are learning with a calm mind from time to time
SP05	We may have different methods of reflecting, but the mental process remains the same
SP06	It is easy to start the process of reflection by following division markers like units and chapters in books, which help, identify when to begin pondering
SP07	We can reflect on our work or learning anytime in any quiet place

(continued)

(continued)

SP08	We can write down all new things we learnt and make connections with what we already know
SP09	This exercise allows our brain to remember any new information for future reference
SP10	We can also think about aspects of knowledge that we cannot apply in everyday life
SP11	We need to question ourselves and write those down so that in future it will help us understand
SP12	Therefore, it is never too late to practice this essential skill of 'self-reflection' to be a successful student

References

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford University Press.

Bernhardt, E. B. (1983). Testing foreign language reading comprehension: The immediate recall protocol. *Die Unterrichtspraxis/teaching German, 16*(1), 27–33.

Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Centre for the Study of Reading Technical Report, No. 270*.

Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices* (2nd ed.). Pearson Education Incorporated.

Coté, N., Goldman, S. R., & Saul, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes, 25*(1), 1–53.

Corbeil, G. (2010). Exploring the effects of first-and second-language proficiency on summarizing in French as a second language. *Canadian Journal of Applied Linguistics/revue Canadienne De Linguistique Appliquée, 3*(1–2), 35–62.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. (CEFR). Cambridge University Press.

Field, J. (2004). *Psycholinguistics: The key concepts*. Psychology Press.

Freebody, P., & Anderson, R. C. (1986). Serial position and rated importance in the recall of text. *Discourse Processes, 9*(1), 31–36.

Just., M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122

K12 Reader Reading Instruction Resources (2012). www.k12reader.com. Retrieved Aug 2013 from http://www.k12reader.com/worksheet/self-reflection/

Kent, C. E. (1984). A linguist compares narrative and expository prose. *Journal of Reading*, *28*(3), 232–236. http://www.jstor.org/stable/40032707

Khalifa, H., & Weir, C. J. (2009). *Examining reading: Studies in Language Testing, 29*. UCLES/Cambridge University Press.

Kim, S. A. (2001). Characteristics of EFL readers' summary writing: A study with Korean university students. *Foreign Language Annals, 34*(6), 569–581.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363–394.

Kirkland, M. R., & Saunders, M. A. P. (1991). Maximizing student performance in summary writing: Managing cognitive load. *TESOL Quarterly, 25*(1), 105–121.

Kobayashi, M. (2002). Method effects on reading comprehension test performance: Text organization and response format. *Language Testing, 19*(2), 193–220.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.

Kucer, S. B. (2010). Readers' tellings: Narrators, settings, flashbacks and comprehension. *Journal of Research in Reading, 33*(3), 320–331.
Kucer, S. B. (2011a). Going beyond the author: What retellings tell us about comprehending narrative and expository texts. *Literacy, 45*(2), 62–69.
Kucer, S. B. (2011b). Processing expository discourse: What factors predict comprehension? *Reading Psychology, 32*(6), 567–583.
Ministry of Human Resource Development. (2020). *National education policy*. Government of India
Meyer, B. J. (1982). Reading research and the composition teacher: The importance of plans. *College Composition and Communication, 33*(1), 37–49.
Meyer, B. J., & Rice, G. E. (1984). The structure of text. In P. David Pearson, R. Barr, & M. L. Kamil (Eds.) *Handbook of reading research 1*. Lawrence Erlbaum Associates, Inc., pp. 319–351.
Mukhopadhyay, L. (2017). Assessing coherence through cloze-based and free summaries: A study of adult EFL learners. *Language and Language Teaching, 6*(1), 14–18.
Norris, J. M. (2016). Current uses for task-based language assessment. *Annual Review of Applied Linguistics, 36*, 230.
Patil, V. (2020). *Using summaries for assessing reading comprehension: a study of young ESL learners*. Unpublished doctoral dissertation. EFL University
Prabhu, N. S. (1991). *The learner's effort in the language classroom.*
Ratteray, O. M. (1985). Expanding roles for summarized information. *Written Communication, 2*(4), 457–472.
Robinson, P. (2001a). Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition. *Second Language Research, 17*(4), 368–392.
Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27–57.
Robinson, P. (2003). The cognitive hypothesis, task design, and adult task-based language learning. *University of Hawai'I Second Langauge Studies Paper 21* (2).
Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching, 43*(1), 1–32.
Robinson, P. (Ed.). (2011). *Second language task complexity: Researching the cognition hypothesis of language learning and performance.* (Vol. 2). John Benjamins Publishing Company.
Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *International Review of Applied Linguistics in Language Teaching, 45*, 161–176.
Skehan, P. (1998). Task-based instruction. *Annual Review of Applied Linguistics, 18*, 268–286.
Snowling, M. J., & Hulme, C. (Eds.). (2005). *The science of reading: A handbook*. Blackwell.
Taylor, L. (2013). Testing reading through summary: Investigating summary completion tasks for assessing reading comprehension ability. *Studies in Language Testing, 39* (UCLES/Cambridge University Press).
Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.
Winograd, P. N. (1984). Strategic difficulties in summarizing texts. *Reading Research Quarterly, 19*(4), 404–425.
Van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension.
Zinar, S. (1990). Fifth-graders' recall of propositional content and causal relationships from expository prose. *Journal of Reading Behaviour, 22*(2), 181–199.

Chapter 17
Tasks in Language Acquisition Research: More Than What Meets the Blinking Eye

Shruti Sircar and Lina Mukhopadhyay

Abstract The paper reviews tasks used for studying children's language acquisition—judgement methods (grammaticality/acceptability judgements, yes–no/truth value judgements); repetition/elicited imitation and dictation; elicited production using picture description and narrative retelling. Supported by studies with children learning languages, like Bangla, Malayalam as first language, and English as second language, the review outlines the task design, task administration and scoring criteria employed for each task. The central aim of the paper is to highlight some of the key methodological considerations with respect to linguistic and cognitive task demands, age appropriacy, and general limitations like attention, memory, sociocultural variations that might interfere with learners' task performance. These considerations need to inform task design since young children are known to be sensitive to pragmatic and metalinguistic influences, and hence are easily biased by carelessly designed experiments.

Keywords Grammaticality judgement · Truth value judgement · Elicited imitation · Dictation · Picture description · Narrative retelling · Language acquisition · Assessment · Methodological constraints

Introduction

A number of paradigms are used to assess child language particularly syntax. They use the same principles of assessment as used in a typical language test in pedagogic contexts. Like any language test, they are designed to eliminate any context or situational cues so that only formal aspects of language, i.e. morphology, grammatical structure, and lexis can be directly tested. However, tests used in language acquisition

S. Sircar (✉)
Department of Linguistics & Contemporary English, EFL University, Hyderabad 500007, India
e-mail: shrutisircar@efluniversity.ac.in

L. Mukhopadhyay
Department of Training and Development, EFL University, Hyderabad 500007, India

N. P. Sudharshana and L. Mukhopadhyay (eds.),
Task-Based Language Teaching and Assessment,
https://doi.org/10.1007/978-981-16-4226-5_17

research depart from the usual language tests in that the former are often tests of comprehension, and these tests/tasks (story comprehension, grammaticality judgements, truth value judgement, picture-selection) have their own assumptions and operational constraints. A good test is one where the task or methodology does not compromise the construct that is being measured or assessed in any way. For instance, if the cognitive demands for a task are too high for a child, the test would give a 'false negative', i.e. the child would fail to produce or comprehend the linguistic item not because he/she lacks linguistic knowledge but because he/she is unable to retrieve or reflect the knowledge through his/her performance.

Language acquisition tasks assess linguistic performance (i.e. what children say and do) which is an indirect assessment of their linguistic competence (i.e. what children know). The data collected from tasks elicit some sort of an observable behaviour from which the internalized linguistic competence can be inferred. In this connection, our paper reviews some of the widely used tasks used for studying children's language acquisition, tasks which study comprehension, i.e. grammaticality judgement, truth value judgements, anomaly detection and online processing; and those which involve some element of production, i.e. elicited imitation, dictation tasks and tasks which involve essentially production, i.e. picture story description and narrative retelling.

The paper outlines some of these task types and illustrates them with studies of children learning languages, like Bangla, Malayalam as first language and English as second language. The intention is to highlight some of the methodological considerations of task demands and age appropriacy, to ensure that general limitations like attention and memory do not interfere with test results. It also throws light on some aspects of language that do not usually get formally assessed in a pedagogic contexts.

In describing each task, the design that we use is that of a general description of the task, followed by a sample task, which includes a description of the target language structure and/or meaning, the task stimuli and design, followed by the general findings of the study. The tasks study either first or second language acquisition in children. Sections "Grammaticality Judgment Task–Elicited Imitation Task" describe receptive/comprehension-based tasks which do not involve production except either a yes/no response or a written choice while Sections "Dictation Task–Narrative Retelling Task" describe production-based performance.

Language Acquisition Tasks: Issues in Design and Measurement

Grammaticality Judgement Task

The task: Grammar as a mental construct is not accessible to conscious awareness and therefore cannot be explicitly observed or tested. All measurements are therefore

by definition indirect. Grammaticality judgement tasks require speakers to indicate whether a sentence in the target language is well formed and possible; the response can be a simple yes/no or can be a scale, a Likert scale of 5 or 7 points on which judgement response needs to be indicated. The scale is designed to ascertain the level of surety of one's judgement.

The modality of task stimuli could either be written or oral; though the latter is more difficult to implement since it should be done individually and in an environment with no noise or any other distraction (see Murphy, 1997 for a detailed discussion on modality impact on judgement). The written form has a different constraint; it can be used only with older and literate learners.

The advantage grammaticality judgement has over spontaneous language production task is that it can be designed to study a specific language structure, for example, the dative alternation. Different forms of grammaticality judgement tasks have been used to study the structure of language: (a) forced choice, where two sentences are provided and speakers have to choose the one that is well formed; (b) bipolar judgement ('yes or no') or judgement on a 5 point or 7 point Likert scale; (c) judgement on a set of sentences with picture support (this is different from a picture-sentence matching task) and (d) a timed version of the judgement task where the time taken to respond is also recorded.

Grammaticality judgements need to be treated as distinct from acceptability judgements. Utterances that are grammatically well formed might be deemed unacceptable by a native speaker since it exceeds the limits imposed by the human parser. (1) is grammatically well formed but would be seen unacceptable by many language users.

(1) The rat the cat the dog followed chased died.
(To mean, the rat that the cat chased that the dog followed died)

The reason (1) is seen as unacceptable is not because of grammar, since language allows recursion of such kind. It is because of processing constraints. It is difficult to keep more than two unfinished nodes active in the short-term memory (see Chomsky & Miller, 1963 for reference).

On the contrary (2b) which is grammatically ill-formed in English would be rated as acceptable, and is in fact used in common speech, though the same speakers would vehemently shoot it down in a grammaticality judgement test.

(2) a. Which book did you return without reading____?
b. Which book did you return without reading it?

One of the criticisms levelled against this task type is that in the task the language becomes an object of scrutiny (Bresnan, 2007) and needs to be judged for its form rather than meaning. The task is a syntactic cousin of the lexical decision task in word recognition studies. The ability to comment on form is a metalinguistic ability which makes it unsuitable for use for children below 12 years because of the absence of operational thought.

The test sentences in the task need to be designed in a way that the sentences vary in only the target structure and not in any other way. It has been found that the choice of the determiner, the type of noun phrase (lexical vs. pronominal), the case

of the noun phrase, and the transitivity of the verb affect processing time, though not necessarily accuracy. In order that memory constraints not implicate the results, the sentences should have the same constituent length (and if possible similar word length). The instructions play a surreptitious role in the way the sentences are judged: An instruction like—'Are these sentences possible in the language?' would make speakers think up contexts where they could be marginally used. Another kind of response is evinced if the instructions are 'Would you use these sentences in your language?' Therefore, the instructions need to reflect the task objectives and should be trialled prior to the actual administration.

The experiment: The grammaticality judgement task we describe here required 10 to 13 year-old second language learners of English to make binary judgements on the well formedness of ditransitive verbs and their participation in dative alternation.

Target structure: Ditransitive verbs like 'give' in English alternate between a prepositional dative (3a) and a double object dative (3b) reading which is reflected in the alternating positions of the direct and indirect object in the sentences:

(3) a. John gave$_1$ [$_{\text{direct obj}}$ a book] to [$_{\text{indirect obj}}$ Mary].
b. John gave$_2$ [$_{\text{indirect obj}}$ Mary] [$_{\text{direct obj}}$ a book].

The theme 'a book' in (3a) is the direct object, and the goal is encoded in a prepositional phrase 'to Mary', therefore called a 'prepositional dative' (PD). In (3b) the goal 'Mary' occurs next to the verb where the preposition 'to' is incorporated. The two objects are both governed by the verb, therefore, called a 'double object dative' (DOD).

Verbs showing the dative alternation form a sub-class (4a). Those that do not participate in the double object dative construction and only occur in a prepositional dative construction (PD) form a separate sub-class (4b).

(4) a. alternating class: *give, lend, throw, pass, send, mail*
b. non-alternating class: *push, carry, present, supply*

Fine-tuned semantic features determine whether verbs alternate or do not alternate. Verbs of continuous causation of transfer (i.e. *push, carry*) and verbs of fulfilling (i.e. *present, supply*) do not occur in double object dative constructions (Gropen et al., 1989; Levin, 1993). The objective in the task was to examine whether second language learners are aware of these semantic distinctions, which can be discerned from their judgements of these verbs in grammatical and ungrammatical sentences, i.e. in PD and DOD frames.

Participants: Sixty-seven 10–13 year-old Bengali learners of English, at four grades (Grades 5, 6, 7, and 8) with at least 5 years of exposure in the language, participated in the study.

Design*:* Six alternating verbs and four non-alternating verbs were used in monoclausal sentences in both PD and DOD frames (16 grammatical, 4 ungrammatical). 5 filler sentences were used. Each sentence was eight to ten words long. All sentences were in the active voice, in simple past tense, and were monoclausal sentences with one finite verb.

Children were first shown two practice items and they were asked to identify sentences that were grammatical. Thereafter, they were instructed to read the sentences in the worksheet and tick (✓) or cross (X). There was no time limit set for the task. Practice items had similar frames but had benefactive verbs e.g., (Mary *baked*$_1$ a cake for her sister/Mary *baked*$_2$ her sister a cake).

Scoring: Since the double object frame is more restricted and marked, correct acceptance of a double object frame was scored '2' and acceptance of the prepositional dative frame as '1'. Similarly, correct rejection of the former was scored '2'. The scoring scheme was used to examine the learning of the marked construction.

Findings: The important trends were as follows: (a) judgement of non-alternating verbs was less accurate than alternating verbs; (b) the accuracy increased for all subclasses of verbs with age and grade, and (c) rejecting ungrammatical sentences were poorer than accepting grammatical sentences. Thus, fine-tuned semantic restrictions took longer to be figured out, though it was eventually learnt by grade 8.

Truth Value Judgement Task

The task: Unlike the grammaticality judgement task where the *form* is judged, in the truth value judgement task, the basis of judgement is the 'truth value' or the *meaning* of the sentence. The task requires the child to make a bipolar judgement about whether a sentence accurately describes a particular situation: the situation is presented either in the form of a picture, or a video, or a preamble. The success of the task clearly resides in the simplicity of the response and therefore can be used with children as young as 2;5.

The task takes usually two forms: either a bipolar yes/no or a reward-punishment format, where the child either rewards a puppet for making a true statement or punishes it for making a false statement about a situation, and the latter was popularized by Stephen Crain and his colleagues to evaluate children's knowledge of complex syntactic principles and quantification (Crain, 1991).

Tasks in language acquisition research attempts to study the availability of particular aspects of knowledge in child language. However, if performance interferes with our understanding of competence, our understanding of language acquisition would be affected. The truth value judgement task minimizes this interference. The assumption is that the only way a child can understand whether the sentence is true is by parsing the sentence and matching it with the context presented. The sentence parsing depends on the stored representation, and therefore, the child response would be a window to the linguistic representations in the child's brain.

The experiment: The experiment is a truth value judgement task where 4;0 and 6;0 children learning Malayalam as their first language judge cleft constructions (Jasmine, 2018). The design was devised by Aravind et al. for English cleft constructions and had been adapted for Malayalam. The sentences were placed in the context

of a picture. The task recorded the accuracy of response as well as the time taken to respond to the sentence.

Target construction: Clefting is the predominant way of asking and answering questions and mark focus in Malayalam (Jayaseelan, 2001; Madhavan, 1987).

(5) a. aar-aaNə puucca-(y)e ooDikkunn-atə
who-COPULA cat-ACCUSATIVE chase.present-NOMINALIZER
'Who is it that is chasing the cat?'

b. patti-(y)aaNə puucca-(y)e ooDikkunn-atə
dog-COPULA cat-ACCUSATIVE chase.present-NOMINALIZER
'It is the dog that is chasing the cat.'

(6) a. patti aar-aaNə ooDikkunn-atə
dog who-COPULA chase.present-NOMINALIZER
'Who is it that the dog is chasing?'

b. patti puucca-(y)aaNə ooDikkunn-atə
dog cat.ACCUSATIVE-COPULA chase.present-NOMINALIZER
'It is the cat that the dog is chasing.'

(Jasmine, 2018, pp. 79–80)

(5a, b) are subject cleft question and its response, and (6a, b) those of object cleft. One of the aims of the study was to check whether there was a subject–object asymmetry in the judgement of clefts. However, a felicity condition was added. The presupposition (the pragmatic constraint) was manipulated in the following way: here we use English (Aravind et al., 2016) for easy reference.

(7) a. Look, something is chasing the cat. I wonder what it is!
b. It is the dog that is chasing the cat. [subject felicitous]
c. ?It is the cat that the dog is chasing. [subject infelicitous]

(8) a. Look, the dog is chasing something. I wonder what it is!
b. It is the cat that the dog is chasing. [object felicitous]
c. ?It is the dog that is chasing the cat. [object infelicitous]

Participants: Two groups of Malayalam-speaking children from three age bands—age band I (4;4 to 5;2), age band II (5;4 to 6;0) and age band III (6;1 to 6;10)—participated in the study. All children spoke Malayalam as their first language and most of them had monolingual parents.

Design: The task was a timed variant of the truth value judgement task and was designed by Aravind et al. (2016). Two pictures with oral sentences were presented

to children one after the other on the computer screen with DMDX display which recorded both accuracy and reaction time of children's response.

The first sentence was a 'question' picture which established that the expected answer was a clefted construction: the Malayalam of the English sentence 'Look, something is chasing the cat. I wonder what it is'. The picture had two–three characters who were involved in a reversible action, and one character was concealed by a box with a question mark. The concealed character could be thematic agent 'the dog' or the thematic object 'the cat' in (7), thus prompting either a subject or an object cleft construction. This was followed by a second picture where the part concealed is revealed to the child. The picture is accompanied with an aural sentence (either 7b or 7c) (Figs. 17.1 and 17.2).

The children were instructed to indicate whether the second sentence was correct or not. The children responded to the 'truth value' of the sentence i.e. they matched the question with the answer based on the picture. All sentences presented to children were grammatically well formed; therefore, judgements were based on meaning rather than form. The intention was to see whether the felicity condition affected the accuracy and response time for subject and object clefts.

Children were trained to press a blue key to say 'yes' and a red key to say 'no': the keys were at the right and left of the keyboard respectively. In order for the children to be comfortable with the response format, and familiar with the keys, prolonged practice was provided, until the children were conversant with the keys. Since the time taken to respond was also calculated, it was of crucial importance that children had no confusion about the keys.

In the sentences, all other things were kept constant except the felicity condition and the subject–object conditions. All verbs were reversible, and all nouns were definite. The base order for the clefts was SOV, the default word order in Malayalam.

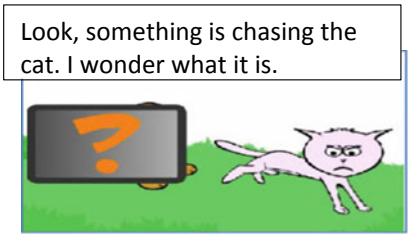


Fig. 17.1 Cue picture for subject cleft. @Picture from Aravind et al. (2016: p. 56)

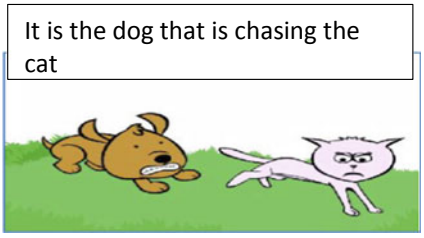


Fig. 17.2 Target picture for subject cleft. @Picture from Aravind et al. (2016: p. 56)

A 12,000 ms timeout was used, giving enough time to children to understand the sentences and to respond. Altogether, 16 target sentences (4 contexts × 4 sentence types [subject felicitous, object felicitous, subject infelicitous, object infelicitous]) and 12 filler sentences and 8 practice sentences were used.

Scoring: Since all sentences were grammatically and semantically correct with reference to the second picture, all 'YES' responses were coded as correct and 'NO' responses as incorrect.

Findings: Accuracy results showed that felicitous objects were judged more accurately than felicitous subjects, and infelicitous objects were lowest in accuracy. The interpretation of clefts, both subject and object improved with age, and the effect of felicity was seen more in older children than younger children. The response time analyses also showed that felicitous objects were understood faster than subjects, and infelicitous objects took the longest time to be interpreted. This asymmetry in subject–object was attributed to the default word order and the fact there was preference for the clefted constituent to be preverbal; therefore, S O $_{\text{cleft}}$ V was preferred over S_{cleft} O V).

Elicited Imitation Task

The task: Elicited imitation requires learners to hear an aural sentence and then attempt to repeat it as accurately as possible. This task has been used for nearly half a century to measure first language knowledge (Bloom et al., 1974; Slobin & Welsh, 1973), abnormal language development (Fujiki & Brinton, 1987) and also in some cases linguistic performance of second language learners (Ellis, 2005; Naiman, 1974; Vinther, 2002).

To repeat verbatim an aural sentence that one has heard only once is possible only if one understands its meaning. To understand its meaning, the sentence needs to be parsed correctly, and for parsing one needs to use linguistic representation that is stored in the mind. Therefore, a child cannot repeat a sentence if the structural information in the sentence far exceeds his/her knowledge of the structure in question. The child's ability to repeat a sentence, therefore, could be used to unveil the underlying syntactic representation that the child is using.

Imitation involves a series of cognitive processes: (a) hearing the input, parsing it and forming a representation (in memory), that includes information at various levels (phonological, syntactic and semantic), (b) holding this representation in the short-term memory for a while and (c) starting the articulation process and producing the sentence from the accessed representation (Bley-Vroman & Chaudron, 1994).

Short-term memory where information is held has limited capacity and therefore can retain very little information at any point of time. Success of retention depends on the number of units being processed, and the manner in which information has been previously stored. When a long and complex representation is held, chunking becomes necessary so that the representation can be held until it is repeated. Also,

the ability to chunk is related to language competence, i.e. more proficient speakers are better at chunking, therefore, more efficient in forming representations and more accurate in reconstructing the sentence.

The experiment: Thirty 10–12 year-old second language learners of English were asked to repeat a question with long-distance dependencies, orally presented to them (Sircar, 2015). The repetition/imitation revealed the underlying representations of *wh* extraction. Learners focus on conveying the meaning illustrated by the sentence, and do so in language that they know, thus often departing from the structure represented in the sentence presented to them.

Target structure: Ross (1967) showed that in forming questions in English a wh constituent cannot be extracted from embedded clauses which are relative clauses (9b), complex NPs (10b) and adjunct clauses (11c), but can be extracted from complement clauses (12b) and that clauses (13b). '__' indicates the position from where the wh constituent is extracted.

(9) a. Prema knew the girl who won the award.
b. *What did Prema know the girl who won __?

(10) a. Prema knew the story that Nick married a princess.
b. *Who did Prema know the story that Nick married ___?

(11) a. Prema laughed when the clown spilled water on the little girl.
b. *What did Prema laugh when the clown spilled ___ on the little girl?

(12) a. Prema knew the girl will win a prize.
b. What did Prema know the girl will win ___?

(13) a. Prema convinced me that I should meet a doctor.
b. Who did Prema convince me that I should meet ___?

The relative clauses, complex NP clauses and adjunct clauses are syntactic islands and do not allow subject or object wh extraction. The imitation task required children to imitate such type of sentences.

Participants: Twenty Indian language learners of English, 10 each from Grade 6 and Grade 8 participated in the study. Learners of Grade 6 had a mean age of 11;3 (SD = 0.73) and Grade 8 a mean age of 14;6 (SD = 1.1). They were speakers of Dakkani Hindi or Telugu and had studied English at least from Grade 1.

Design: Twenty sentences, 12 grammatical and 8 ungrammatical were used. All sentences had an embedded clause and were of comparable length, with 13 to 16 syllables. All embedded clauses had transitive verbs with animate subjects and inanimate objects. For parity, each question had a prepositional phrase (e.g. *from the house*) primarily to add length to the sentence and to avoid any form of rote repetition. Tense was kept constant in the sentences, with the main clause in the present tense and the embedded clause in simple past. Sentences in the tests were randomized and presented to learners. Two declarative sentences were used as training items to ensure that learners understood the instructions.

Scoring: Accurate repetitions were scored as '1' while inaccurate repetitions were scored 0 for grammatical targets. Lexical errors (wrong names, verb substitutions, synonyms, etc.) were ignored while scoring. For the ungrammatical targets, accurate repetitions were scored 0. Recast where the target ungrammatical structure was corrected was scored 1. Recast where the obligatory structure was completely deleted was scored 0. Incomplete repetitions and other ungrammatical repetitions were also scored 0. Given below are examples of errors and recasts in the task (Sircar, 2015).

(14) a. *What does Yuni hear the news the plumber stole ____from the house?
b. What does Yuni hear the plumber stole ____from the house? (Score 1)
c. What did the plumber steal from the house? (Score 0)
d. What does Yuni hear the news …. (Score 0)

Findings: Sentences with islands were difficult to repeat, and learners tended to recast them, most of the time without much success. Older learners tended to recast them by deleting the offensive constituent (as in 14b) more often than younger learners. Grammatical targets with object extraction was more difficult to imitate than targets with subject extraction, and children used a medial wh in many of the grammatical targets, indicative of the stepwise fashion in which wh moves, often leaving a copy behind. Elements of resumption was also seen in their imitation.

Dictation Task

The task: A task parallel to the elicited imitation task is the *cross-modal dictation task*, where the learners hear a sentence and write it down. The advantage of this task is that it can be done in groups or as a whole class activity and participants need not be tested individually like the elicited imitation task which is time and resource consuming. The flip side of this task is that it places extra demands of writing speed, spelling and punctuation accuracy. However, these effects can be minimized by ignoring the spelling and punctuation errors while scoring and providing enough time between two sentences for children to write them down.

The study we report looked at the acquisition of English prepositions and their lexical/functional divide, where the functional words were often found to be delayed in acquisition (Jayasinge, 2015). The study was able to throw light on the ranking of lexical/functional features in acquisition.

Target structure: English prepositions have been found to indicate a lexical-semantic meaning or serve a purely functional-syntactic purpose.

(15) a. He put the book on the table.
b. His translation of the book was awful.

'On' in (15a) has a lexical meaning which indicates a location, whereas 'of' in (15b) has no specific meaning but is needed functionally to relate two nouns,

'translation' and 'the book'. Littlefield (2006) further divided prepositions into four categories using [± lexical] and [±functional] features, arguing that the combination of these two properties determines the order of acquisition of a number of prepositional elements in English. [+lexical] features contribute semantic content, and [+functional] are needed to check the case of their complements. These fundamental distinctions result in the four categories of prepositional elements as in (16a-d), and the acquisition proceeds from the most to the least lexical, according to her claims.

(16) a. Adverbs: *put down the cup* [+lexical, -functional]
b. Particles: *he ate it up* [-lexical, -functional]
c. Semi-lexical prepositions: *run to the store* [+lexical, +functional]
d. Functional prepositions: *translation of the book* [-lexical, +functional]

Littlefield (2006) found evidence that children acquire the prepositions in a particular order: adverbial, particle, semi-lexical and finally functional. Littlefield's study on naturalistic speech of first language children found [− functional] features were shown to rank higher than [+ lexical] features in acquisition.

One of the aims of the present study was to find out whether this ranking was good for Sinhalese learners of English and to examine whether, there was an advantage of either [+ lexical] features or [− functional] features at initial stages and disappeared at later stages of learning.

Participants: 260 Sinhalese students from four grades: 65 of them from Grade 4 (mean age 8.07), 65 from Grade 6 (mean age 10.10), 65 from Grade 8 (mean age 12.09) and 65 from Grade 10 (mean age 14.10) participated in the study.

Design: The task had 40 sentences, with ten sentences each for adverbs, semi-lexicals, particles and functionals. In constructing the sentences, the length of the sentence (six words per sentence, except in sentences with adverbs with four words), the syntactic construction and the tense of the verb were controlled. All the sentences were in active voice and in simple past. Each construction denoted only one event and had only one finite verb. Words used in the task were frequently occurring and in everyday use.

The sentences were recorded, and a 35-second pause was provided after each sentence to enable learners to write the sentences down. Each learner was provided a sheet with 40 numbered spaces to write the sentences. A five-minute break was provided after 20 sentences, and then the rest of the sentences were played. Instructions were provided orally in English and Sinhala.

Scoring: Each correctly written sentence was scored '2' if the target preposition and the other parts of the sentence were grammatically correct. Minor spelling errors were ignored for the analysis. If the latter had errors but the target preposition was intact, then it was scored '1', and if the preposition was omitted or substituted or misspelt, it was scored '0'. Accuracy of repetition and the type and nature of errors for each category was also computed.

Findings: If the fine-grained approach is shown to be effective with prepositions, then ESL learners would use the different categories in the sentences differently: this was ascertained, therefore justifying the four-way classification. Adverbs and semi-lexical prepositions were better imitated than particles and functional prepositions (i.e. [+ lexical] seemed to rank higher than the [− lexical] features). This 'lexical' advantage disappeared in Grade 10 with all categories being used with comparable accuracy. Younger learners omitted the functional prepositions, and older learners' substitutions were category specific.

In the sections above tasks (1–4) were designed to test comprehension rather than production, while in the subsequent sections tasks (5–6) reported were designed to test comprehension and production. Note that the production in the elicited imitation and dictation demands a verbatim recall and does not involve composition which is an essential aspect of tasks that study production. Therefore, tasks (1–4) are not considered as typical 'production' tasks though they may involve some form of verbal output.

Picture Description Task

The task: One of the time-tested methods of eliciting language samples for first/second/foreign language learners is a picture description task, where a single picture or a picture strip is provided. A variation of this task is the story-generation task where a picture strip which tells a story is provided, and the learner has to create a story and tell it either orally or in writing. It has the benefits of both spontaneous and elicited production tasks.

Spontaneous speech or writing is a connected language sample that is a useful context for observing the integrated function of all components of speech. Unlike reading or imitation, spontaneous speech requires formulation which draws upon greater cognitive and linguistic resources, therefore at times, is fraught with fillers, pauses and is less intelligible than more constrained elicitation tasks (Kempler & Van Lancker, 2002). Therefore, a method which elicits information-rich language yet is constrained would be critical for language assessment. Picture story or picture description has this dual benefit. The language sample thus elicited is connected and can be used for assessment of a range of skills: grammatical and lexical diversity and complexity, story grammar, formulation, discourse skills, etc.

Design: The factor that is of crucial importance is the selection of the scene/picture. If the task is used with children, the scene needs to be engaging to a wide range of children with respect to age and cultural background, and needs to be illustrated in a child-friendly manner. The picture description task should evoke a familiar everyday theme that would be accessible to a diverse group of children. If the task is meant to tap grammatical accuracy and lexical complexity and diversity, the scene should lend itself to the use of diverse syntactic structures and lexical categories (nouns verbs, adjectives, prepositions, etc.). It should illustrate events of different kinds, which

require simple clauses, complement clauses and embedded structures. The choice of the scene depends on the target analysis.

The experiment: The study reported (Mukhopadhyay, 2015) looks at a picture description with a twist. The task was used to establish that when cognitive demands are reduced, whether the task output becomes linguistically more complex, subject to the proficiency of the learners (see Nair & Sircar, Chap. 8 of this volume for a similar study). The cognitive and linguistic demands were lowered by asking bilingual learners to describe a picture first in their L1 (Bangla) and then in their L2 (English). This was compared with a picture description task in L2 without any L1 mediation. The study design was adapted from Lanauze and Snow (1989) in which Spanish and English children's writing ability were compared.

Target analysis: The focus of the study was to establish that the use of L1 as a pre-task helps in lowering the cognitive and linguistic demands of the task, and therefore attentional resources can be allocated to linguistic features of complexity, diversity, and organizational-discoursal features.

Participants: Thirty-four 11–12-year-old Bengali learners of English participated in the test. At the time of the study, they had 6 years of classroom exposure to the target language (English). The learners were categorized on the basis of teacher reports into four groups: good in L1 and L2 (GG); good in L1 but not L2 (GP); poor in L1 and good in L2 (PG) and poor in both (PP). The aim was to see whether L1-mediation benefits all groups equally in all domains of language.

Task input: As detailed earlier, the language samples of the children depend on the selection of the picture. Two pictures were used: Picture A, to examine writing output in English only and Picture B, to examine it in a L1 mediated condition. In the study, the focus was on syntactic and lexical complexity, and therefore, the pictures selected had the following features (Tables 17.1 and 17.2):

Fig. 17.3 Picture A for description

Fig. 17.4 Picture B for description

Table 17.1 Specifications Picture A

Passage attribute	Rationale	Occurrence
Broad appeal	To create a scene that appeals to children from diverse cultures, ages, language abilities	Familiar scene: two children getting ready (here and now)
Sub-scenario/events	To focus attention and provide topics to elicit language sample	Events: during bed time two boys are getting ready to go to bed, changing dress and putting clothes on the rack, on a chair Inference: brothers/friends in a hostel room, chatting before falling asleep
Representation of words	Nouns	Bedroom, bed, quilt, pillows, chair, pictures/portraits, door, doormats, wall, boys, clothes, pyjamas, socks/mittens, towels, balls, teddy bear/toy
	Verb phrases	Wear clothes, put on the socks, sit on the floor/bed, look at each other
	Adjective	Colour terms: orange, red, blue, brown; emotions: happy, excited
	Prepositional phrases	On the chair/doormat/bed/floor/wall, in the bed, at the door
Structures	Simple, coordinating and subordinating clauses	A variety of sentence structure; tense unspecified
Discourse connections	Adverbials (to denote time); causality	One day, then, after that…as, because, since, so

Table 17.2 Specifications Picture B

Passage attribute	Rationale	Occurrence
Broad appeal	To create a scene that appeals to children from diverse cultures, ages, language abilities	Familiar scene: of a man selling balloons and children running towards him
Sub-scenario/events	To focus attention and provide topics to elicit language sample	Events: a man selling balloons; the man leaning against a tree; balloons tied to a walking stick; the man and children looking at the balloons; children running towards him; a girl skipping, one boy's arm on the other's shoulder Inference: children want to buy balloons; they are happy; the man is tired and therefore leaning against the tree
Representation of words	Nouns	Vendor, old man, stick, pants/trousers, coat, walking stick, balloons, thread, children, boys, girls, shirts, skirts, frocks, tree trunk, branches, leaves, grass, shrubs, field, pool of water,
	Verb phrases	Walk on the field, run towards the man/vendor, lean on the trunk, skip towards the tree, tied to the strings, hanging from the walking stick
	Adjective	Describe objects/humans: barefooted, old, young colour terms: orange, red, blue, brown, green, polka dots emotions: happy, excited, joyful, tired, amused
	Prepositional phrases	On the tree, against the trunk, across the field, at the man,
Structures	Simple, coordinating and subordinating clauses	A variety of sentence structure; tense unspecified
Discourse connections	Adverbials (to denote time); causality	One day, then, after that…as because, since, so, but

Instruction and Task Administration: Learners were instructed to write at least ten sentences describing the scene and include:

- names of objects, their colour and location
- description of human figures (colour of dresses, physical description, etc.)
- actions and emotions.

Picture A was described in L2, and learners were given feedback on the task. After a week, Picture B was to be described in first their L1 and on the subsequent day in L2. However, in describing Picture B in L2, they were not allowed access to their L1 texts, ensuring that it was a translation of ideas rather than word by word translation.

Scoring: Learner picture descriptions for the two conditions (L2-only and L1-mediated conditions) were coded for: *syntactic complexity*—mean T-unit length in words, T-unit sub-types: subordinating and coordinating and *text cohesiveness*—frequency of referential markers (addition, causal, temporal and adversative markers). A frequency count of the measures of syntactic complexity and cohesiveness were used for final analysis.

Findings: The main findings of the study were as follows: (a) all groups benefited in syntactic complexity and text cohesiveness from L1 mediation, (b) the quantum of benefit differed across proficiency levels: GG, GP and PP groups showed significant increases in syntactic complexity; only GG group showed an increase in content elaboration and text cohesiveness; GP and PP groups showed an increase in the use of referential markers but no increase in content elaboration. This suggested a support for CALP transfer in groups GG and GP as they had threshold level proficiency in L2.

Narrative Retelling Task

The task: Narrative retelling requires a child to retell the story that she/he heard. Usually to aid memory, the story is told with picture support. A picture strip-based narrative telling or retelling task is used to assess a child's knowledge of lexis, morpho-syntax and story grammar and integrates comprehension and production.

Narrative retelling is widely used to elicit spontaneous connected speech (and/or written responses) from children and adults for research and clinical purposes. The study reported in the later half of this section shows how narrative retelling is used to assess children's lexical skills as a predictor of their success in reading comprehension. We show that narrative retelling tasks generate language samples that are similar to oral language samples used in norm-referenced and standardized tests. Speakers generate their own words, thus narrative retelling can be used equitably for linguistically diverse groups and for assessing in particular multicultural–multilingual populations. Unlike elicitation tasks of standardized tests, the format can provide a more holistic view of spoken language skills (Stockman, 1996).

We know the language used to (re)tell stories is very different from regular conversational speech. But the advantage of using this task type is that the topic, content, and length of verbal discourse can be controlled, and unlike conversational skills, it can be used more easily across different speakers without requiring them to produce the same words or grammatical constructions. Unlike elicited imitation tasks, where the stimuli is a sentence (though a complex one!) the stories are far too long to be

memorized; therefore, learners are forced to draw on their own lexical and linguistic resources to retell the stories. Moreover, the experimenter in telling the story can control to a certain extent the amount, topic and complexity of language elicited. The task allows for a range of skills to be tested apart from grammatical and lexical skills, i.e. speech intelligibility, lexical diversity, discourse skills, formulation skills, verbal memory (Culatta et al., 1983; Liles, 1993), as well as cultural specificity of recall (Mandler et al., 1980).

In fact, story generation and story retelling pose different speaking demands on the test-taker. A story-generation task, which is popularly used as a picture story telling or writing task, requires the learners to use pictures to tell or create a story. This is different from telling a story that one already knows. The story retelling requires one to comprehend a story and retell it in their own language, so that the ideas and content are represented adequately. Therefore, the role recall plays in retelling are more pronounced, and therefore, can also be a test of comprehension, unlike story generation. When a learner creates a story, the content needs a form, and therefore, the sentences are less complex, and descriptions are less elaborate. Another advantage that narrative retelling has over story generation is that since the content is known, establishing the contextual information and comparing the performance across learners becomes easier for the experimenter.

The modality in which the input is provided plays a crucial role: it has been found that an auditory input with picture support leads to more linguistically complex retelling than an auditory only input (Gibbons et al., 1986; Hayes et al., 1986). However, this advantage is modulated with age, exposure to input and proficiency in the language of retelling.

The study: The task we report was used in the India-based 'MultiLiLa' project led by the University of Cambridge in collaboration with Indian and UK academic institutions. The oral narrative retelling task performance was used to measure the development of lexical diversity of school-going multilingual Indian children aged 7 to 12 years and coming from low socio-economic status (SES) families. The study aimed to find out if children's knowledge of oral vocabulary correlated with their success in reading comprehension in English, which was the language of medium of instruction at school. Reading comprehension was measured through the English ASER test (2017) on components of sentence and text reading followed by two novel comprehension questions specially designed for the study (for details refer to Tsimpli et al., 2020).

Participants: From a large cohort of 1800 children in three different sites—Delhi, Hyderabad and Patna—studying in grades four and five from low SES families, a smaller group of 90 children studying in English medium schools in Hyderabad were considered for the study.

Fig. 17.5 Sample story from MAIN @Gagarina et al. (2012: p. 25)

Task Input: The task used in the study was sourced from *Multilingual Assessment Instrument for Narratives* (MAIN) prepared and validated by Gagarina et al., (2012). The task had a six-picture strip with a story script that was aurally played on the computer. The task was controlled for:

(a) story characters (the number, time when they are introduced and their representation in pictures);
(b) background and foreground information (cognitive complexity and visual density) and
(c) content (how the story develops and concludes).

MAIN includes four-picture strips corresponding to four different but comparable stories controlled for the features (a–c) listed above. Figure 17.5 is a sample story 'The Cat Story' which had six pictures and the corresponding audio transcript (Fig. 17.5).

Scoring: The oral narrative retelling samples were transcribed and measured for lexical diversity in the language of medium of instruction (English). Lexical Diversity was measured using the CLAN transcript software at three levels—*word* (all inflected forms as different types), *morpheme* (counts homographs or different parts of speech as different tokens) and *lemma* (all inflected forms as same type).

A sample analysis of two sentences on the CLAN at the word level (TXT tier) morpheme (MOR tier) level and grammar (GRA or POS tier) level is presented in Figs. 17.6 and 17.7.

*TXT: *butterfly is on the tree*
%mor: n|+n|butter+n|fly cop|be&3S prep|on det:art|the n|tree .
%gra: 1|2|SUBJ 2|0|ROOT 3|2|JCT 4|5|DET 5|3|POBJ 6|2|PUNCT
*TXT: *cat is seeing the butterfly* .
%mor: n|cat aux|be&3S part|see-PRESP det:art|the n|+n|butter+n|fly .
%gra: 1|3|SUBJ 2|3|AUX 3|0|ROOT 4|5|DET 5|3|OBJ 6|3|PUNCT

Fig. 17.6 CLAN analysis for lexical diversity

<pic 1/2> One day there was a playful cat who saw a yellow butterfly sitting on a bush. He leaped forward because he wanted to catch it. Meanwhile, a cheerful boy was coming back from fishing with a bucket and a ball in his hands. He looked at the cat chasing the butterfly. <pic 3/4> The butterfly flew away quickly and the cat fell into the bush. He hurt himself and was very angry. The boy was so startled that the ball fell out of his hand. When we saw his ball rolling into the water, he cried, "Oh no! There goes my ball!" He was sad and wanted to get hos ball back. Meanwhile the cat noticed the boy's bucket and thought, "I want to grab a fish." <pic 5/6> At the same time the boy began pulling his ball out of the water with his fishing rod. He did not notice that the cat had grabbed a fish. In the end, the cat was very pleased to eat such a tasty fish and the boy was happy to have his ball back.

Fig. 17.7 Audi text for 'the Cat story' @Gagarina et al. (2012, p.54)

The measures of lexical density were as follows: Type Token Ratio (TTR = V/N), Guiraud Index (GI = V/$\sqrt{}$N) (Guiraud, 1954) and Moving Average Type Token Ratio (MATTR) with a window size of 18 words (Covington & McFall, 2010).

Findings: The main trends were as follows: (a) lexical diversity counts increased with age and grade as measured through Guiraud Index on count of lemmas; and (b) higher lexical diversity showed a corresponding increase in reading comprehension. The findings add to the existing body of research that vocabulary knowledge as lexical diversity is an important indicator of development of reading for meaning.

Methodological Considerations

Language competence is hard to capture. Neither is the underlying knowledge nor the mechanisms that enable language acquisition amenable to observation and usually lie hidden in the mind of the child. To uncover and reveal these, researchers depend on an array of tools. This paper was about these tools describing the techniques that language acquisition researchers use in studying language development. The focus was on the description of the techniques of data collection, transcription and coding of speech/ written samples. The paper presented six such tools and described the research aims these methods served, the details of task implementation and the type of data the methods yielded. Each method was a sort of journey from the research

aims to the research methods in studies that authors had personally engaged in or had been associated with. Each of the six methods provided the attendant advantages and disadvantages of data collection.

Experimental paradigms for assessing children's linguistic knowledge (of grammar and vocabulary) can be broadly divided into three types: production, comprehension and judgement. *Production methods* use various techniques to make children produce particular sentence types with the hope that the errors they make will reveal something of theoretical interest. In *comprehension methods*, children demonstrate their comprehension of a sentence that is verbally presented to them by choosing a matching picture from a selection (either explicitly by pointing or implicitly by looking). *Judgement methods* require children to make judgements about the well-formedness of a sentence.

Of the six tasks that were reviewed, two were judgement tasks (grammaticality and truth value judgements), one was free production task (picture description) and the other a guided production task (narrative retelling), and two were tasks where stimuli had to be comprehended and repeated (elicited imitation and dictation). So the six tasks form an assessment continuum from comprehension to guided to free production.

Judgement: The main advantage of the grammaticality judgement paradigm is that it allows a teacher or researcher to answer questions that cannot be directly addressed using production or comprehension measures. It investigates children's knowledge of grammar (both syntax and morphology) in a relatively explicit manner. Moreover, it provides unambiguous, numerical data that does not require extensive scoring, or checking for interrater reliability, and are in fact suitable for most commonly used statistical analyses (e.g. ANOVA, regression). The more crucial part of the paradigm is that it allows a researcher to create novel (ungrammatical) sentences which are in the target language but probably allowed in other languages in the world, or to create sentences which violate universal linguistic constraints, rejection of which would establish their knowledge of language constraints. The disadvantage of the paradigm is that creating grammatical and ungrammatical targets, counterbalancing the targets and fillers, controlling for sentence length, syntactic and morphological segments can be relatively demanding for the researcher. Also, the paradigm is most suitable for use with relatively older children. However, the difficulty in design is compensated by the ease of scoring and analysis.

In pedagogic contexts, a teacher could use a grammaticality judgement task either as role plays or as a worksheet to assess children's understanding—grammaticality and semantic value—of target structures. Though these tasks are a bit intrusive, a teacher can use them through the play-way method to access children's understanding of abstract concepts of language. The tasks have a broad scope as they can be used equally well with all age groups and sociocultural backgrounds and for first and second/foreign language learning as well as for clinical purposes.

Production: The main advantage of free (and elicited) production studies is that the researcher can exert a reasonable degree of control over what children are likely to say and hence manipulate the variable(s) of interest. The main disadvantage is that

elicited production tasks are probably the most difficult for children to complete. Hence, children may fail not because they lack the required knowledge but because they do not understand the nature of the task, or one or more of the various task components (e.g. interpreting the scenario to be described, choosing the right words, planning the utterance) interfere with their ability to produce the correct form. Thus, production does not necessarily reflect competence. Moreover, since children vary in their utterances, they are less likely to be bounded, and therefore, comparisons across children become difficult. The speech/writing samples unless objectively analysed for linguistic features (like complexity and diversity) either manually or with designated software, require verification and a good interrater reliability scheme. Furthermore, responses, spontaneous or elicited, are difficult to code and analyse. The prompts that are used for elicitation need to be trialled to ensure that they yield the intended language constructions. Thus, what looks simple can be demanding and time consuming.

The picture description and narrative retelling tasks do not assess a specific set of (or target) syntactic structures unlike the judgement and imitation tasks; they lend themselves to a holistic language sample collection. The cognitive and linguistic load of the tasks need to be carefully manipulated according to task complexity features grouped in Robinson's triadic componential model (refer to Robinson, 1997 for details) without which performance can be negatively affected. Note that the MAIN manual (2012) with the four story tasks has been carefully controlled for such features, and can be used with bi/multilingual populations with transcripts available in European and Indian languages.

Data analysis and assessment of language proficiency can be done in a far more flexible manner either on micro- (lexical and/or morphosyntactic levels) or macro-features (discourse level or story grammar) of language; a teacher/researcher can choose specific aspects to measure or analyse, based on contextual needs. Although the studies reported in this paper have used quantitative measures like T-units to measure syntactic complexity and CLAN software to measure lexical diversity, teachers can use task-specific assessment criteria to create learner estimates and to give feedback.

Imitation: The primary advantage of elicited imitation tasks and dictation tasks is the high degree of control that they afford over the precise form and wording of the target utterance. This control is possible only when the targets are designed with utmost care and attention so that the targets differ only in the critical segment and all other things that might interfere with repetition are controlled. This design feature is important if the target is to link and interpret any errors in repetition with the critical segment. What becomes a point of theoretical interest are the errors children make in repeating sentences and the changes they make to ungrammatical sentences. One of the disadvantages of this method is that it cannot be reliably used with adults and older children, who at some stage will be able to repeat a sentence verbatim using a pure 'parroting' strategy.

The elicited imitation and its counterpart cross-modal dictation task involve comprehension of target structure(s) and their immediate recall. So these are higher

level tasks than either comprehension or production tasks as they require working memory (central executive) functions of storage, processing (recall grammatical sentences with the target structure) and updating (recast ungrammatical sentences) information. Furthermore, recasts of ungrammatical sentences add to the cognitive load.

Conclusion

The six language acquisition tasks presented above are very close to language tests and can be used for classroom language assessment. However, one needs to understand the advantages and limitations of these tasks so that teachers can make informed choices on which tasks to use for what purpose and with which age group of learners. To use the tasks for classroom purposes, teachers have to consider a host of contextual and learner internal factors and task factors that could possibly impact the success of the design, administration and validity of results. The variables likely to impact task design and performance are as follows: *learner internal variables* like age, context of learning the target language (first/second/foreign; naturalistic vs. classroom), working memory span, level of proficiency, *task factor variables* such as aspect of language to be assessed (comprehension vs. production) data collection type (spontaneous vs. elicited; guided vs. free), modality (oral vs. written) and in it *linguistic constraints* (word and syllable length) and *cognitive constraints* (time allotted, resource-directing vs. resource-dispersing elements, steps in the tasks). A combination of these factors would play a significant role in performance. The tasks can be used by researchers and teachers alike; but they would need to understand that there would always be 'a trade-off' between task design and analysis, i.e. tasks which are easy to design are usually difficult to score and analyse, while tasks which are easy to score and analyse are difficult to design. In conclusion, though the tasks presented in this chapter look like they can be used only for language acquisition research have more use beyond what meets the blinking eye.

References

Aravind, A., Hackl, M., & Wexler, K. (2016). Syntactic and pragmatic factors in children's comprehension of cleft constructions. *Language Acquisition, 25*(3), 284–314.

Bley-Vroman, R., & Chaudron, C. (1994). Elicited imitation as a measure of second-language competence. In E. E. Tarone, S. Gass, & A. D. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 245–261). Erlbaum, L. Associates.

Bloom, L., Hood, L., & Lightbown, P. (1974). Imitation in language development: If, when and why. *Cognitive Psychology, 6*, 380–420.

Bresnan, J. (2007). Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In S. Featherston, & W. Sternefeld (Eds.), *Roots: Linguistics in search of its evidential base* (Studies in Generative Grammar) (pp. 77–96). Mouton de Gruyter.

Chomsky, N., & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 269–321). Wiley.

Covington, M. A., & McFall, J. D. (2010). Cutting the Gordian knot: The moving-average type–token ratio (MATTR). *Journal of Quantitative Linguistics, 17*(2), 94–100.

Crain, S. (1991). Language acquisition in the absence of experience. *Behavioural and Brain Sciences, 14*, 597–650.

Culatta, B., Page, J. L., & Ellis, J. (1983). Story retelling as a communicative performance screening tool. *Language, Speech, and Hearing Services in the Schools, 14*, 66–74.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*(2), 141–172.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Valimaa, T., Bohnacker, U., & Walter, J. (2012). *MAIN: Multilingual Assessment Instrument for Narratives. Report*. German Ministry for Education and Research (BMBF).

Gibbons, J., Anderson, D. R., Smith, R., Field, D. E., & Fischer, C. (1986). Young children's recall and reconstruction of auditory and audiovisual narratives. *Child Development, 57*(4), 1014–1023.

Gropen, J., Pinker, S., Hollander, M., Goldberg, R., & Wilson, R. (1989). The learnability and acquisition of the dative alternation. *Language, 65*(2), 203–257.

Hayes, D. S., Kelly, S. B., & Mandel, M. (1986). Media differences in children's story synopses: Radio and television contrasted. *Journal of Educational Psychology, 78*, 341–346.

Jasmine, M. G. (2018). *Acquisition of subject and object clefts in Malayalam* (Unpublished PhD dissertation). EFL University.

Jayaseelan, K. A. (2001). Questions and question-word incorporating quantifiers in Malayalam. *Syntax, 4*, 63–93.

Jayasinge, R. R. (2015). *Acquisition of fine-grained categories of prepositions by ESL learners* (Unpublished PhD dissertation). EFL University.

Kempler, D., & Van Lancker, D. (2002). The effect of speech task on intelligibility in dysarthria: A case study of Parkinson's disease. *Brain and Language, 80*, 449–464.

Levin, B. (1993). *English verb classes and alternations. A preliminary investigation.* Massachusetts Institute of Technology Press.

Liles, B. (1993). Narrative discourse in children with language disorders and children with normal language: A critical review of the literature. *Journal of Speech and Hearing Research, 36*, 868–882.

Fujiki, & Brinton, B. (1987). Elicited imitation revisited: A comparison with spontaneous language production. *Language, Speech, and Hearing Services in the Schools, 18*(4), 301–311.

Madhavan, P. 1987. *Clefts and pseudo-clefts in Malayalam* (PhD dissertation). CIEFL.

Mandler, J. M., Scribner, S., Cole, M., & DeForest, M. (1980). Cross-cultural invariance in story recall. *Child Development, 51*(1), 19–26.

Mukhopadhyay, L. (2015). Using L1 knowledge to enhance L2 writing performance. *The EFL Journal, 6*(1), 15–32.

Murphy, V. (1997). The effect of modality on grammatical judgment task. *Second Language Research, 13*(1), 34–65.

Naiman, N. (1974). The use of elicited imitation in second language acquisition research. *Working Papers on Bilingualism, 2*, 1–37.

Pratham (2017). Annual Status of Education Report (ASER Rural) 2016. New Delhi: Pratham. Available at http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202016/aser_2016.pdf

Robinson, P. (1997). Individual Differences and the Fundamental Similarity of Implicit and Explicit Adult Second. *Language Learning, 47*(1), 45–99. https://doi.org/10.1111/0023-8333.21997002

Ross, J. R. (1967). *Constraints on variables in syntax* (Doctoral dissertation). Massachusetts Institute of Technology.

Sircar, S. (2015). Elicited imitation as a measure of ESL learners' grammatical knowledge. *The EFL Journal, 6*(2), 43–62.

Slobin, D. I., & Welsh, C. A. (1973). Elicited imitation as a research tool in developmental psycholinguistics. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development* (pp. 485–497). Holt, Rinehart & Winston.

Stockman, I. (1996). The promises and pitfalls of language sampling analysis as an assessment tool for linguistic minority children. *Language, Speech and Hearing Services in the Schools, 27*, 355–366.

Tsimpli, I. M., Balasubramanian, A., Marinis, T., Panda, M., Mukhopadhyay, L., Alladi, S., & Treffers-Daller, J. (2020). Research Report of a Four-Year Study of Multilingualism, Literacy, Numeracy and Cognition in Delhi, Hyderabad and Patna.

Vinther, T. (2002). Elicited imitation: A brief overview. *International Journal of Applied Linguistics, 12*(1), 54–73.