

Word-Level and Sentence-Level Automaticity in English as a Foreign Language (EFL) Learners: A Comparative Study

Dongmei Ma¹ · Xiaoru Yu² · Haomin Zhang³

Published online: 17 June 2017
© Springer Science+Business Media, LLC 2017

Abstract The present study aimed to investigate second language (L2) word-level and sentence-level automatic processing among English as a foreign language students through a comparative analysis of students with different proficiency levels. As a multidimensional and dynamic construct, automaticity is conceptualized as processing speed, stability, and accuracy which are indexed by reaction time, coefficient variation and accuracy rate. Sixty students (39 undergraduate students and 21 graduate students) who majored in English participated in this study. They completed the lexical semantic classification task, the sentence construction task, the sentence verification task under two different modalities (visually- and aurally-presented situations). Multivariate analyses were conducted to examine the differences between the students with different proficiency levels pertaining to their L2 performance. The results indicated that the processing speed was not found to be a good indicator of automatic language processing. Moreover, both levels of students appeared to reach a plateau in word-level processing but there were some variations in students' sentential processing. Finally, the findings showed that automatic language processing seemed to be module-specific and non-sharable across different modalities and skills.

Keywords Automaticity · Proficiency · English as a foreign language

✉ Dongmei Ma
dongmei@seu.edu.cn

Xiaoru Yu
x.yu@let.ru.nl

Haomin Zhang
haominzh@cmu.edu

¹ School of Foreign Languages, Southeast University, Nanjing, China

² Center for Language Studies, Radboud University Nijmegen, Nijmegen, The Netherlands

³ Department of Modern Languages, Carnegie Mellon University, Pittsburgh, PA, USA

Introduction

Conceptualization of Automaticity

Automaticity is a complex construct with various conceptual and operational definitions. [Kahneman \(1973\)](#) argues that automaticity requires little attentional control or awareness in cognitive processing and it also refers to automatic processing as opposed to controlled processing. Interface and non-interface theoretical approaches have been proposed to understand the automatization of language acquisition ([Anderson 1983](#); [Logan 1988](#); [Paradis 2009](#)). The interface approach states that controlled processing can be transferred to automatic processing through repeated use and practices ([Schmidt 1994](#)) while the non-interface approach argues that there is no speeded-up controlled processing to automatic processing, but there is only a continuum from predominant reliance on controlled processing to predominant reliance on automatic processing ([Paradis 2009](#)). Regardless of the connection or disconnection between controlled and automatic processing, automaticity, in general, is a subconscious condition and often tied with quick speed and high efficiency when learners perform multiple complex tasks ([DeKeyser 2001](#)). In the realm of language learning, automaticity was characterized as a more efficient, more accurate and more stable performance of language processing ([Segalowitz 2003](#)). More recently, [Segalowitz \(2010\)](#) has refined this conceptualization of automaticity as processing speed, processing stability, and processing flexibility.

Word-Level and Sentence-Level Automaticity in Second Language

With regard to automatization of specific linguistic features, [Grabe and Stoller \(2011\)](#) proposed that lower-level linguistic processes, such as lexical access, syntactic parsing and semantic proposition formation, may necessitate automatic or subconscious language processing whereas higher-level processes (e.g., inference and comprehension) may require conscious use of background knowledge. The processes of language comprehension involve attentional resources from lower-level, data driven processes to higher-level, conceptual processes ([Breznitz 2006](#); [Koda 2005](#); [Stanovich 2000](#)). [Lim and Godfroid \(2014\)](#) mention that, unlike lower level processes, higher-level processes may not be automatized because they require conscious processing and are highly context-dependent. Therefore, automaticity of language processing is often associated with word-level automaticity (e.g., automatic word recognition).

[Favreau and Segalowitz \(1983\)](#) explored the relation between L2 reading speed and automatic processing during lexical access and they drew upon the lexical decision task to manipulate the occurrence of priming stimuli and target words. Sixty bilingual speakers (thirty Francophones and thirty Anglophones) participated in the study. The findings indicated that half the participants read first and second language at the same speed and the rest read second language slowly. More critically, bilinguals with equal 53 L1 and L2 reading rates showed evidence in automatic processing based on reaction times. Nonetheless, the results of bilingual speakers with slower L2 reading speed showed indications of automatic word recognition in L1 rather than L2. Recent training studies tried to test whether automatic word recognition can be facilitated through various modes of intervention. [Akamatsu \(2008\)](#) investigated the effect of word recognition training on word-recognition performance among Japanese-speaking EFL learners. Students were measured by a lexical decision task before and after a 7-week training. The findings testified to the effectiveness of training in improving the speed and accuracy of word recognition. In addition, the study also included working memory capacity and word frequency as variables to test their

relative contributions to automatization of word recognition. The results indicated that working memory span did not necessarily affect the improvement in the speed and efficiency of word recognition. Meanwhile, improvements in recognizing low-frequency words were more associated with automatization. [Sato et al. \(2013\)](#) investigated the impact of a technology-enhanced learning tool (a multimedia learning application) on improving the automatization of word decoding among EFL learners. The learning tool with a time-control function was found to enhance the automatization of word decoding skills.

In addition to word-level automatization, a number of studies have explored the automatization of morphosyntactic or grammatical features. [Dekeyser \(2007\)](#) investigated the automatization of morphosyntactic knowledge by probing how 61 adults acquired the grammar rules of an artificial language over one semester. After the participants completed an 8-week comprehension and production practice, a test was administered to them to assess how well they could comprehend and produce grammatical structures in that artificial language. The results revealed that morphosyntactic knowledge improved gradually across time and automatization of L2 knowledge followed the similar learning curve as other skills. It is also worth mentioning that the practice effect could lead to acquisition of skill-specific morphosyntactic knowledge. Additionally, [Jiang \(2007\)](#) explored the differences between native English speakers and non-native Chinese-speaking English learners in processing grammatical idiosyncrasies, namely plural noun and verb subcategorization, through a self-paced reading task. The participants were asked to read both grammatical and ungrammatical sentences as quickly as they could in a word-by-word manner. The findings demonstrated that native speakers had a delay in reading ungrammatical sentences of both structures whereas non-native speakers were only sensitive to grammatical errors in verb subcategorization. Likewise, [Rodgers \(2011\)](#) investigated the automatization of morphosyntactic knowledge (verbal morphology) in comprehension and production across foreign language learners with different proficiency levels. 85 undergraduate students majoring in Italian were divided into three subgroups based on their proficiency level and they completed a picture identification and a picture description task to measure their receptive and productive skills. The results indicated that learners with higher language proficiency responded faster and more accurately in both tasks compared with learners with lower proficiency level. The study suggests that automatization of verbal morphology develops with language proficiency level. More important, the findings comparing receptive and productive skills lent support to the differences in automatization of modalities of linguistic capacities.

Based on the review of literature, several research objectives were identified for this study. Firstly, previous studies predominantly focused on automatic word recognition and word-level processing skills. However, the question remains as to whether automaticity of word-level knowledge could determine overall L2 automaticity. [Paradis \(2004\)](#) argues that single words are not ideal for the understanding of language representation because linguistic systems inclusive of phonology, morphology and syntax involve cognitive functioning in language processing. Decontextualized word recognition or decoding may not provide a clear picture of language automatization. It is crucial to scrutinize the automaticity of language skills through lens of word-level processing and sentential processing among L2 learners. Secondly, automaticity of language skills seems to be moderated by general language proficiency ([Rodgers 2011](#)). Nonetheless, few studies have addressed to what extent language proficiency would affect automaticity in second language processing. This study aims to investigate whether language proficiency would trigger differences in L2 automaticity. Thirdly, the current literature has shown the possibilities of automaticity in different macro-level linguistic modalities (reading, writing, listening and speaking), however, most studies were restrained to automaticity in the reading modality ([Rawson 2010](#)). It is crucial

Table 1 Participant pool

Subjects	Mean age (years)	Proficiency	Standardized tests passed	Years of English learning (years)
Undergraduate (<i>N</i> = 39)	21	Low	TEM-4	8
Graduate (<i>N</i> = 21)	25	High	TEM-8	11

to expand the scope to automaticity in other linguistic modalities and this study emphasizes both aural and visual representations of language processing (listening and reading). Finally, little research has explored the transferability of different modalities of second language processing (DeKeyser 1997; Rodgers 2011; Segalowitz and Fishman 2005). The current study also aims to explore the possible relation between reading and listening automaticity in second language in order to shed light on whether processing skills underlying the two receptive modalities are sharable.

To address the research objectives presented above, three research questions were posed accordingly.

1. Does language proficiency affect automaticity in word-level and sentence-level processing skills through aurally-presented stimuli?
2. Does language proficiency affect automaticity in word-level and sentence-level processing skills through visually-presented stimuli?
3. Are there any relationships in word-level and sentence-level automaticity between the two task modalities?

Method

Participants

The study was conducted at a national key university in China, and 60 participants who were English majors participated in this study. Table 1 shows some basic information about the students' background. Among them, 39 undergraduate students passed a standardized Test for English Majors (TEM) Band 4 and 21 graduates passed TEM¹ Band 8. The undergraduate students had learned English for at least 8 years while the graduate students had learned English for at least 11 years. Each student's actual score of the standardized proficiency test was checked to ensure that they were accurately classified to the low or the high proficiency groups.

Measurements

As shown in Table 2, three tasks were adopted from Hulstijn et al. (2009), and Lim and Godfroid (2014), namely lexical semantic classification, sentence construction and sentence verification. Linguistic processes underlying these tasks were lexical access, syntactic parsing and semantic proposition formation. As we discussed above, processing speed, stability and accuracy seem to be three important properties of automaticity (DeKeyser 2001; Segalowitz

¹ TEM: Test for English Majors, a standardized national test in China for undergraduate English majors. TEM4 is taken by sophomore undergraduates and TEM8 is taken by senior undergraduates.

Table 2 Measurements, linguistic processes involved, and indices

Task	Linguistic process	Data index
Lexical semantic classification	Lexical access	RT, CV, ACC
Sentence construction	Syntactic parsing	RT, CV, ACC
Sentence verification task	Semantic proposition formation	RT, CV, ACC

2010). Therefore, three indices including reaction time (RT), coefficient of variance (CV) and accuracy rate (ACC) were used to quantify the three respective features in automatic language processing.

All visual and aural stimuli were encoded by the E-prime software program. The listening and reading tests consisted of four stages, instruction, training, transition and administration. Test procedures and instructions were provided at the beginning. After reading the instruction, the participants were asked to complete a training session. On the first screen, they saw a red cross to remind them that a stimulus would be presented. Four stimuli were presented in the training session. There was an interval of 500 milliseconds (ms) between every two stimuli. The participants were supposed to give their response after the display of the stimulus. A note popped up to indicate whether it was correct or not when the participants submitted their response. The transitional page showed that the participants could repeat the training or start the experiment immediately. In the actual experiment, an interval of 800 ms was used as a buffer to reduce the participants' cognitive load. In addition, the participants were not told if their response was correct or not in the actual experiment.

Semantic Classification Task The semantic classification task was used to measure how learners comprehend language at the word level. The participants were required to make a quick judgement about whether a word they heard and read was a living being. In this task, there were 48 items, among which were “animal”, “baby”, “bed”, “book” etc. When subjects heard or read the word “animal” or “baby”, they were supposed to press “J” on the keyboard for these two items refer to living creatures. On the contrary, when the item referred to a non-living being, such as “bed” and “book”, the participants were supposed to press “F” on the keyboard.

Sentence Construction Task The sentence construction task was designed to test the participants' ability to parse sentences. In this task, the first part of a sentence was visually or aurally presented to the participants and then two options were provided visually or aurally. They were asked to select the following part of the sentence by pressing correspondent letters (F and J) on the keyboard. The participants were informed in advance that the sentence might not be complete even with the correct option. Test items were kept short and simple in order to rule out the confounding effect of semantic analysis, thus focusing on the sentence parsing ability.

Sentence Verification Task The sentence verification task was the one in which the participants were asked to judge whether a visual or an aural statement made sense. Nonsensical sentences included those which violated factual knowledge or grammatical knowledge. Four sample stimuli are shown below:

1. A horse is an animal that can fly.
2. My uncle made me for a snowman.
3. He went inside, though it started to rain.
4. Summer is the hottest time of a year.

Table 3 Descriptive statistics of the listening test

Task	Index	Proficiency	Min	Max	Mean	SD
Semantic classification	RT	Low	745.63	1504.46	1102.50	152.94
		High	846.04	1505.66	1178.24	176.68
	CV	Low	.15	.47	.31	.08
		High	.14	.42	.28	.07
	ACC	Low	.83	1.00	.97	.03
		High	.61	1.00	.96	.07
Sentence construction	RT	Low	4953.87	8092.86	6816.60	687.47
		High	5602.50	8147.96	6926.22	578.72
	CV	Low	.07	.31	.15	.05
		High	.07	.25	.13	.04
	ACC	Low	.48	1.00	.84	.11
		High	.69	1.00	.92	.08
Sentence verification	RT	Low	2559.97	4326.17	3543.41	419.69
		High	2772.89	5393.05	3542.47	602.38
	CV	Low	.15	.41	.30	.06
		High	.17	.37	.26	.06
	ACC	Low	.61	.98	.89	.08
		High	.84	1.00	.93	.04

The first sentence violates the factual knowledge that horses cannot fly. The second sentence is grammatically inaccurate and the correct sentence should be “My uncle made me a snowman.” The third sentence is inaccurate because the conjunction should be causal ones such as “since” and “because”. The participants needed to respond to these stimuli through pressing the letters on the keyboard. Before the testing, the participants were explicitly told to focus on meaning instead of form.

Results

Descriptive statistics of the listening and reading tests are shown in Tables 3 and 4. Overall, high-proficiency students had relatively higher accuracy rates than low-proficiency students in both task modalities. Pertaining to processing stability, there was no salient discrepancy in CV between the two proficiency groups within each task modality. With regard to processing speed, high-proficiency students only responded more quickly than low-proficiency students to the aural presentation of the sentence verification task and the visual presentation of the sentence construction task. A Shapiro–Wilk’s test showed that indices of RT ($p > .05$) and CV ($p > .05$) in all the measures were approximately normally distributed for both groups, but ACCs ($p < .05$) violated the normality in the measures. Meanwhile, the former two indices did not reject the homogeneity of variance assumption while the latter did, as suggested by a Levene’s test ($p > .05$). Therefore, analyses of RT and CV were based on the parametric tests, MANOVA and Pearson’s correlation while the index of ACC was analyzed using non-parametric statistics including Mann–Whitney Test and Spearman’s rho.

Table 4 Descriptive statistics of the reading test

Task	Index	Proficiency	Min	Max	Mean	SD
Semantic classification	RT	Low	563.14	931.13	722.43	85.69
		High	627.57	972.24	775.69	103.55
	CV	Low	.18	.40	.26	.06
		High	.16	.36	.26	.05
	ACC	Low	.86	1.00	.96	.04
		High	.93	1.00	.97	.02
Sentence construction	RT	Low	969.42	2189.87	1445.05	283.19
		High	898.24	1787.15	1395.95	259.35
	CV	Low	.19	.55	.35	.08
		High	.21	.55	.36	.08
	ACC	Low	.82	1.00	.94	.05
		High	.82	1.00	.95	.04
Sentence verification	RT	Low	1956.70	4591.79	2993.14	530.61
		High	2268.85	4731.89	3301.54	648.60
	CV	Low	.21	.45	.34	.06
		High	.21	.47	.33	.06
	ACC	Low	.73	1.00	.91	.06
		High	.82	1.00	.94	.05

Group Differences on Word-Level and Sentence-Level Automaticity

Processing Through Aurally-Presented Stimuli (Listening Automaticity) A MANOVA test was employed to determine whether there were differences between the two groups of subjects on processing speed and processing stability in the three tasks of the listening automaticity test. The result of Box's test did not reject the null hypothesis of equal covariance matrices for the MANOVA test, thus adopting Wilks's lambda for data interpretation. The multivariate tests revealed that there was a statistically significant between-group difference on RT and CV, Wilks's lambda = .711, $F(6, 80) = 5.421$, $p < .001$, partial $\eta^2 = .289$. The tests of between-subjects effects showed that statistically significant differences existed between the two groups on RT of the semantic classification task, $F(1, 85) = 4.57$, $p < .05$, partial $\eta^2 = .051$, and on CV of the sentence verification task, $F(1, 85) = 12.78$, $p < .005$, partial $\eta^2 = .131$.

As the normality assumption of parametric tests is violated in the case of the ACC variable, a Mann–Whitney U test was conducted to compare processing accuracy of the two groups. The results indicated that accuracy rates of the low-proficiency group in the semantic classification task did not differ significantly from those of the high-proficiency group, $U = 871$, $z = -.485$, $p > .05$. However, high-proficiency students could process testing items in the sentence construction task and sentence verification task significantly more accurately, $U = 500.5$, $z = -3.69$, $p < .05$; $U = 598$, $z = -2.84$, $p < .05$.

Processing Through Visually-Presented Stimuli (Reading Automaticity) Similarly, another MANOVA test was conducted to determine whether there were differences between the two groups of subjects on processing speed and processing stability in the three tasks of the reading automaticity test. The result of Box's test was not significant. The multivari-

Table 5 Parallel correlations between indices of automaticity in the low-proficiency group $N = 39$, * $p < .05$; ** $p < .01$

Task	RT	CV	ACC
Semantic classification task	.376*	.309	.114
Sentence construction task	-.050	-.254	.114
Sentence verification task	.573**	.158	.397*

Table 6 Parallel correlations between indices of automaticity in the high-proficiency group $N = 21$, * $p < .05$; ** $p < .01$

Task	RT	CV	ACC
Semantic classification task	.373	.124	-.430
Sentence construction task	.263	.022	.595**
Sentence verification task	.461*	.133	.262

ate tests revealed no significant between-group difference on processing speed and stability, Wilks's lambda = .862, $F(6, 61) = 1.626$, $p > .05$, partial $\eta^2 = .138$. However, the tests of between-subjects effects suggested some differences between the low-proficiency and high-proficiency students on RTs of the semantic classification task and the sentence verification task, $F(1, 66) = 5.01$, $p < .05$, partial $\eta^2 = .071$; $F(1, 66) = 4.34$, $p < .05$, partial $\eta^2 = .062$.

A Mann–Whitney U test showed that the average accuracy rates of the low-proficiency group in the three tasks of the reading test did not differ significantly from that of the high-proficiency group, $U = 399$, $z = -1.45$, $p > .05$; $U = 464.5$, $z = -.56$, $p > .05$; $U = 362.5$, $z = -1.90$, $p > .05$.

Relationships of Word-Level and Sentence-Level Automaticity Across Task Modalities

We conducted parallel analyses of word-level and sentence-level automatic processing between the two task modalities. In other words, we compared the participants' performance across the two modalities based on the same indices of measurement, reaction time, coefficient of variance and accuracy. To begin with, correlational analyses were run to test the participants' performance in the low-proficiency group. The results (shown in Table 5) showed that reaction times of automatic processing at the word-level were significantly correlated between the two task modalities, $r = .398$, $p < .05$. However, the correlations of CV and ACC between the two groups failed to reach the significance level, $r = .309$, $p > .05$; $r = .114$, $p > .05$. In the sentence construction task, none of the indices correlated with each other across the two modalities. Finally, analyses of the sentence verification task indicated that correlations of RT and ACC were significant, $r = .573$, $p < .01$; $r = .397$, $p < .05$.

We ran the same analyses for the high-proficiency group (Table 6). The results found that none of the correlations of RT, CV and ACC at the word level between the two task modalities were significant, $p > .05$. In the sentence construction task, we only found that the correlation of ACC was significant, $r = .595$, $p < .01$ while correlation coefficients of other indices did not reach the significance level, $r = .263$, $p > .05$; $r = .022$, $p > .05$. Finally, pertaining to the sentence verification task, the only significant correlation was found in RT between the two task modalities $r = .461$, $p < .05$ whereas the correlations for the other indices were insignificant, $r = .133$, $p > .05$; $r = .262$, $p > .05$.

Discussion

Automatic Language Processing and Proficiency Level

The first two research questions addressed whether proficiency level affects automatic language processing through aurally and visually-presented stimuli. The MANOVA test and the Mann–Whitney *U* test investigating the three indices of processing through aurally-presented stimuli between the two proficiency levels found (1) a significant difference in RT means but no differences in the other two measures for the lexical semantic classification task; (2) a significant difference in ACC means but no differences in the other two measures for the sentence construction task; (3) significant differences in CV and ACC means but no differences in the RT means for the last task. Furthermore, the MANOVA test and the Mann–Whitney *U* test investigating the three indices of processing through visually-presented stimuli between the two proficiency levels revealed no significant difference in any task. The subsequent tests of between-subjects effects suggested that, though the difference between the two groups failed to reach significance in processing stability and accuracy at both lexical and sentential levels, undergraduate students reacted faster than graduate students in the lexical task and the sentence verification task.

A number of interpretations can be drawn from the findings. First, high-proficiency students did not necessarily react faster than low-proficiency students in automatic processing under visual and aural situations. It is vital to address the role of processing speed in automatic language processing. Automaticity is not just a simple speed-up of performance but a qualitative change of underlying processing mechanisms (Cheng 1985; McLaughlin 1990; Neely 1977; Schneider and Shiffrin 1977; Segalowitz and Segalowitz 1993). Speed-up effects indexed by reaction times may be unreliable predicting increased automaticity. Coefficient variance, computed by the division of standard deviation and mean reaction times, may provide a reliable index of relative variability.

Second, in terms of processing under the aurally-presented situation, there was no salient difference in word-level processing between the two participant groups. However, the graduate students with high proficiency level had stronger automatic processing ability in sentence-level processing stability and accuracy than their undergraduate counterparts did. This is partially because automaticity of spoken word recognition, which occurs at a lower level of linguistic processing, is presumably achieved prior to automatic sentence processing and both graduate and undergraduate students may have achieved the similar level of automaticity in spoken word recognition. The differences of sentential processing found in the current study may support the previous findings that L2 automaticity may increase with students' growing proficiency (Lim and Godfroid 2014; Rodgers 2011). Finally, the findings did not testify to the hypothesis that proficiency level affects automatic language processing through visually-presented stimuli. This part of results may be explained by the flattening-out part in the power law of learning curve (DeKeyser 1997). Under the influence of the law of practice, L2 automaticity of EFL learners increases dramatically at the beginning. However, after a period of rapid increase, it gradually levels off reaching its final or even fossilized state of automaticity for individual learners. Chinese learners of English tend to draw upon visual representations in the process of English reading (Wang et al. 2003; Wang and Koda 2005) and they seemed to develop their ability to retrieve information from visual clues instead of aural cues throughout their learning, therefore they may have developed solid foundations in visual word recognition ability in the initial stage of language learning.

Automaticity Across the Modalities

One of the purposes of the present study was to understand the relationship between automaticity of L2 processing across different modalities. The correlation analyses revealed that there were weak to moderate correlations between automaticity across aural and visual modalities. This study aimed to understand the similarities or differences of automaticity in language processing across modalities. The core component of automaticity is automatic and subconscious processing of linguistic knowledge, which involves procedural knowledge (Anderson 2007; Ellis 2005; Paradis 2009). When certain skills are proceduralized or automatized, they seem to become modal-specific and non-sharable across different skills or modalities (Dekeyser 2007; Rodgers 2011). The current study indicated that language processing under visual stimulation is not necessarily connected with language processing under aural stimulation. However, it is important to ponder why the two modalities involved different processing skills given that both of them seemingly measured receptive skills of the students. Processing through visually-presented stimuli entails the connection between form and comprehension while processing through aurally-presented stimuli undergoes the connection between sound and comprehension. The two comprehension skills are comprised of different linguistic units (sound and form), which involve different cognitive demands underlying these two types of language processing.

Theoretical and Applied Implications

The findings of the present study have some implications for both theory and practice. Concerning the potential theoretical contributions, the present study sought to explore procedural and automatic language processing. In the ACT-R model, Anderson (1983) has clearly identified that knowledge acquisition undergoes two different types of processing, namely declarative and procedural processing. More critically, he argues that procedural language processing necessitates more skill- and module-specific abilities. In contrast, the Instance Theory proposed by Logan (1988) states that language users with automatic processing capacity are able to retrieve stored instances from memory and have equal access to memory in receptive knowledge (e.g., listening and reading). The findings of the present study attested to the Anderson's ACT-R model that language processing under different modalities may require different sets of competencies and sub-skills. Modal specificity found in the present study indicated that L2 students process aural and visual information under relatively different cognitive routes given that low correlations were found across the task modalities. It is noteworthy that the Instance Theory emphasizes the processing of everyday skills rather than language acquisition, therefore, the theory alone may not be applicable to second language acquisition (DeKeyser 2001; Segalowitz 2003).

Additionally, the present study aimed to furnish the current literature with the multi-dimensional conceptualization and operationalization of automaticity which includes processing speed, stability and accuracy. In previous studies, the construct of automaticity has been defined and measured in various ways and the operationalization and measurement of the automaticity remain controversial. The empirical evidence from the present study demonstrated that in addition to speed-up processing, the measurements of stability and accuracy could generate different patterns in language processing across L2 learners with different proficiency levels. Therefore, it is vital to expand our understandings in automatic language processing and an evolving and dynamic conceptualization of automatic language processing is highly valued in the field.

In addition to the theoretical implications, we would also like to discuss some applied implications pertaining to second language learning and instruction. Firstly, as we discussed, there was a mismatch between language processing across different modalities. The results indicated that reading automaticity (indexed by automaticity through visual stimuli) appeared to remain stagnant between the two proficiency levels. However, the learners in the two groups performed differently in listening automaticity (indexed by automaticity through visual stimuli). These findings can possibly relate to teaching practices in EFL classrooms. EFL students in China have extensive learning experiences in print-based learning and they learn English mostly by the medium of written texts. According to the findings, variations in listening competence seemed to be consistent with students' proficiency. In practical instruction and learning, EFL curricula somehow have overlooked the importance of aural abilities among students. It is advisable to lay more emphasis on the development of students' listening skills.

Secondly, the results identified different patterns in word-level processing and sentence-level processing. For L2 learners, word learning is crucial because it builds semantic foundations for oral and written communication. Sentence comprehension is also vital since it incorporates multiple abilities: syntactic knowledge, vocabulary knowledge and context-based understanding. Language educators need to consider the instructional focus in a specific course and balance word-level and sentence-level instruction. Instructors ought to cater to students' needs based on their preexisting competencies. For example, students with strong syntactic knowledge could be instructed to develop in-depth word knowledge in context because word learning could also occur in the sentential level. Students with strong vocabulary knowledge could be instructed to develop sentence comprehension ability based on word-meaning information and contextual clues.

Limitations and Future Research Directions

This study has some limitations that warrant future investigations. Firstly, cross-sectional research design may not provide a clear picture of development of L2 automaticity because single-wave data collection would limit the generalizability of study findings. Future studies could emphasize longitudinal research design which would be able to compare developmental patterns across time through within-group and between-group analyses. Secondly, the participants were L1 Chinese speakers and they were all from the same university. Some systematic confounding variables (e.g., L1 background, instructional approach) may exist within the participating students. Future studies may expand the sampling procedure and recruit participants from different linguistic backgrounds. Finally, the present study only investigated the relation between proficiency and automaticity and the relation between automaticity across task modalities. We did not address the actual process of automatization and the possible route to achieve language automaticity. Future studies could highlight the transformation from controlled processing to automatic processing to provide insight into the procedure of automatization.

Acknowledgements Funding was provided by the National Social Science Fund of China; Title of the Project: A System for Dynamic Diagnostic Assessment of Chinese University Students' English Speaking Ability (Grant No. 15BYY079). We would like to thank the reviewer for providing constructive feedback on our previous manuscript.

Compliance with Ethical Standards

Conflict of interest The authors below certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speak-

ers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?*. New York: Oxford University Press. doi:10.1093/acprof:oso/9780195324259.001.0001.
- Akamatsu, D. (2008). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics*, 29, 175–193. doi:10.1017/S0142716408080089.
- Breznitz, Z. (2006). *Reading fluency: Synchronization of processes*. Mahwah, NJ: Erlbaum.
- Cheng, P. (1985). Restructuring versus automaticity: Alternative accounts of skill acquisition. *Psychological Review*, 92, 414–423. doi:10.1037/0033-295X.92.3.414.
- DeKeyser, R. M. (1997). Exploring automatization processes. *TESOL Quarterly*, 30(2), 349–357. doi:10.2307/3588151.
- DeKeyser, R. M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). New York: Cambridge University Press.
- DeKeyser, R. M. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press. doi:10.1017/CBO9780511667275.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141–172. doi:10.1017/S0272263105050096.
- Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first- and second-language reading of fluent bilinguals. *Memory and Cognition*, 11(6), 565–574. doi:10.3758/BF03198281.
- Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading*. Harlow: Longman.
- Hulstijn, J. H., Van Gelderen, A., & Schoonen, R. (2009). Automatization in second language acquisition: What does the coefficient of variation tell us? *Applied Psycholinguistics*, 30, 555–582. doi:10.1017/S0142716409990014.
- Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, 57(1), 1–33. doi:10.1111/j.1467-9922.2007.00397.x.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Koda, K. (2005). *Insights into second language reading*. New York: Cambridge University Press. doi:10.1017/CBO9781139524841.
- Lim, H., & Godfroid, A. (2014). Automatization in second language sentence processing: A partial, conceptual replication of Hulstijn, Van Gelderen, and Schoonen's 2009 study. *Applied Psycholinguistics*, doi:10.1017/S0142716414000137.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492–527. doi:10.1037/0033-295X.95.4.492.
- McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, 11, 113–128. doi:10.1093/applin/11.2.113.
- Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology*, 106, 226–254. doi:10.1037/0096-3445.106.3.226.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism* (Vol. 18). Amsterdam: John Benjamins. doi:10.1075/sibil.18.
- Paradis, M. (2009). *Declarative and procedural determinants of second languages* (Vol. 40). Amsterdam: John Benjamins. doi:10.1075/sibil.40.
- Rawson, K. A., et al. (2010). Defining and investigating automaticity in reading comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 185–230). Burlington: Academic Press. doi:10.1016/S0079-7421(10)52005-X.
- Rodgers, D. (2011). The automatization of verbal morphology in instructed second language acquisition. *International Review of Applied Linguistics in Language Teaching*, 49, 295–319. doi:10.1515/iral.2011.016.
- Sato, T., Matsunuma, M., & Suzuki, A. (2013). Enhancement of automatization through vocabulary learning using CALL: Can prompt language processing lead to better comprehension in L2 reading? *ReCALL*, 25(1), 143–158. doi:10.1017/S0958344012000328.
- Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. In J.H. Hulstijn & R. Schmidt (Eds.) *Consciousness and second language learning: Conceptual, methodological*

- and practical issues in language learning and teaching. Thematic issue of AILA Review—Revue de l'AILA, 11 (pp. 11–26).
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing. I: Detection, search and attention. *Psychological Review*, 84, 1–66. doi:10.1037/0033-295X.84.1.1.
- Segalowitz, N. (2003). Automaticity and second language learning. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Oxford: Blackwell. 10.1002/9780470756492.
- Segalowitz, N. (2010). *The cognitive bases of second language fluency*. New York: Routledge.
- Segalowitz, N., & Fishman, S. F. (2005). Attention control and ability level in a complex cognitive skill: Attention shifting and second-language proficiency. *Memory and Cognition*, 33(4), 644–653. doi:10.3758/BF03195331.
- Segalowitz, N., & Segalowitz, S. J. (1993). Skilled performance, practice and the differentiation of speed-up and automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, 14(3), 369–385. doi:10.1017/S0142716400010845.
- Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.
- Wang, M., & Koda, K. (2005). Commonalities and differences in word identification skills among learners of English as a second language. *Language Learning*, 55(1), 71–98. doi:10.1111/j.0023-8333.2005.00290.x.
- Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and non-alphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, 87(2), 129–149. doi:10.1016/s0010-0277(02)00232-9.