

Yearbook of Morphology

2005

Edited by
Geert Booij and Jaap van Marle

YEARBOOK OF MORPHOLOGY 2005

Yearbook of Morphology

Editors: Geert Booij
Jaap van Marle

Consulting Editors: Stephen Anderson (Yale)
Mark Aronoff (Stony Brook, N.Y.)
Mark Baker (New Brunswick, N.J.)
Laurie Bauer (Wellington)
Joan Bybee (Albuquerque, N.M.)
Andrew Carstairs-McCarthy (Christchurch)
Greville Corbett (Guildford, U.K.)
Wolfgang Dressler (Wien)
Ingo Plag (Siegen, Germany)
Jack Hoeksema (Groningen)
Rochelle Lieber (Durham, N.H.)
Peter Matthews (Cambridge, U.K.)
Franz Rainer (Wien)
Sergio Scalise (Bologna)
Henk Schultink (Utrecht)
Andrew Spencer (Colchester, U.K.)
Gregory Stump (Lexington, Kentucky)

Editorial address: Prof. Ingo Plag
English Linguistics, Fachbereich 3
Universität Siegen
Adolf-Reichwein-Str. 2
Siegen
D-57068
Germany
e-mail: plag@anglistik.uni-siegen.de

YEARBOOK OF MORPHOLOGY 2005

Edited by

GEERT BOOIJ

*Leiden University
The Netherlands*

and

JAAP VAN MARLE

*Open University
Heerlen, The Netherlands*

 Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-4065-2 (HB)

ISBN-13 978-1-4020-4065-8 (HB)

ISBN-10 1-4020-4066-0 (e-book)

ISBN-13 978-1-4020-4066-5 (e-book)

Published by Springer,
P.O. Box 17 3300 AA Dordrecht, The Netherlands.

www.springeronline.com

Printed on acid-free paper

All rights reserved.

© 2005 Springer

No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the copyright owner.

Printed in the Netherlands.

Table of Contents

Articles

JAMES P. BLEVINS / Word-based declensions in Estonian	1
HANS-OLAV ENGER / Do affixes have meaning? Polarity in the Toten dialect of Norwegian meets morphological theory	27
JANET GRIJZENHOUT and MARTINA PENKE / On the interaction of phonology and morphology in language acquisition and German and Dutch Broca's Aphasia: the case of inflected verbs	49
ANDREW KOONTZ-GARBODEN / On the typology of state/change of state alternations	83
CHRISTIAN LEHMANN / Pleonasm and hypercharacterisation	119
MARY PASTER / Pulaar verbal extensions and phonologically driven affix order	155
ANDREA D. SIMS / Declension hopping in dialectal Croatian: Two predictions of frequency	201
GREGORY T. STUMP / Referrals and morphemes in Sora verb inflection	227

Discussion: Paradigm Function Morphology

ANDREW CARSTAIRS-McCARTHY / Affixes, stems and allomorphic conditioning in paradigm function morphology	253
GREGORY T. STUMP / Some criticisms of Carstairs-McCarthy's conclusions	283

Book reviews

Paul Boucher (ed.), <i>Many Morphologies</i> (Peter Arkadiev)	305
Danielle Corbin, Pierre Corbin and Martine Temple (eds.), <i>Lexique 16</i> (Claudio Iacobini)	309
Laurie Bauer, <i>A Glossary of Morphology</i> (Geert Booij)	313
Piet van Sterkenburg (ed.) <i>Linguistics Today—Facing a Greater Challenge</i> (Geert Booij)	314

Book information

Proceedings of the 4 th Mediterranean Morphology Meeting, Catania, 21–23 September 2004	315
---	-----

Word-based declensions in Estonian*

JAMES P. BLEVINS

1. INTRODUCTION

The declensional system of modern Estonian exhibits a highly uniform word-based structure. The system is essentially tripartite, comprising a set of singular grammatical cases, a set of plural grammatical cases and a set of semantic cases. The subsystem of singular grammatical cases consists of isolable stems and theme vowels. Yet it is particular stem-vowel combinations that are distinctive, as neither stems nor vowels can be assigned case properties in isolation. The plural grammatical cases are in turn ‘parasitic’ formations (Matthews 1972), based purely on the form of the singular cases. The semantic cases are then based on the form of the corresponding genitive. Within each subsystem, case and other grammatical properties are associated with whole word forms, but these word-level properties cannot be apportioned to smaller units.

This paper suggests that a traditional word and paradigm (WP) model offers an illuminating perspective on the organization of this system. Recognizing words as ‘minimal meaningful units’ directly captures the fact that case is consistently associated with words, but not with sub-word units. The traditional view that words ‘are not wholes composed of simple parts but are themselves the parts within a complex whole’ (Matthews 1991:204) likewise brings out the implicational structure of declensions. This structure includes the patterns of stem syncretism in Figure 1, along with interdependencies between grammatical cases, such as the general predictability of the nominative singular from the partitive or genitive singular. The network of relations between the forms of a noun define a paradigmatic context for the interpretation of individual forms that supplies information that is not represented in their syntagmatic structure.

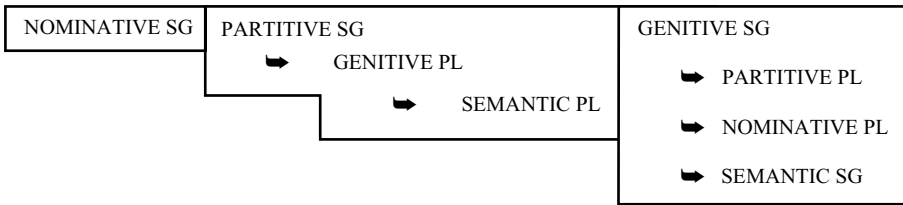


Figure 1. General patterns of parasitic stem syncretism in Estonian declensions

A WP analysis of Estonian declensions also clarifies the challenge that word-based patterns present for post-Bloomfieldian models that assume a bi-unique correspondence between units of content (‘morphemes’) and units of form (‘morphs’). Following Matthews (1972), much of the WP literature has

focussed on cases of ‘cumulative exponence’, in which there are more units of content than units of form, and on cases of ‘extended exponence’, in which there are more units of form than units of content. The emphasis on missing and extra units has fostered the idea that property-form mismatches are the main challenge posed by word-based patterns. However, an examination of the Estonian declensional system shows that a shortfall or excess of units is merely a symptom of a more general problem of morphological ‘overextraction’, in which a property-form relation is extended to units that do not function as Saussurean ‘signs’. In particular cases, notably in flexional languages, overextraction leads to an apparent mismatch between ‘units of form’ and ‘units of content’. Yet, as Estonian shows, overextraction may also characterise analyses in which there is no shortage of form units or content units. Recurrent units of form can be identified at every level in Estonian, from the singular grammatical cases through to the semantic cases. But sub-word units cannot be brought into correspondence with grammatical properties because stems, theme vowels and parasitic bases bear no consistent meaning in isolation. Imposing a morpheme-based description onto such a word-based system simply ‘creates gratuitous problems of analysis and gratuitous problems of explanation’ (Matthews 1991:174).

1.1. Morphomic stem syncretism

Noun paradigms that exhibit productive ‘weakening’ gradation (Erelt et al. 2000:255, Viitso 2003a:27) offer a striking illustration of the challenge posed by word-based patterns of exponence. In the partial paradigms of *HEKK* ‘hedge’ and *KOOL* ‘school’ in Table 1, the nominative singulars *ˈhekk* and *ˈkool*, and the partitive singulars *ˈhekki* and *ˈkooli* are strong, based on the ‘overlong’ or ‘Q3’ (*kolmas välde* ‘third quantity’) stems *ˈhekk-* and *ˈkool-*. The corresponding genitive singulars *heki* and *kooli* are weak, based on the non-overlong stems *hek-* and *kool-*.¹ There are thus three grade-alternating cases in Table 1: nominative, partitive and genitive singular. These cases are realised by means of three distinct formatives: the strong stems *ˈhekk-* and *ˈkool-*, the weak stems *hek-* and *kool-*, and the theme vowel *-i*. There is no ‘mismatch’ between units of form and content, because there are exactly as many case values as exponents. But instead of a biunique correspondence, these elements exhibit the overlapping pattern in Table 1.

Table 1. Overlapping exponence in weakening declensions

Case	Form		Stem	Exponent
Nominative Sg	<i>ˈhekk</i>	<i>ˈkool</i>	STRONG	—
Partitive Sg	<i>ˈhekki</i>	<i>ˈkooli</i>		THEME
Genitive Sg	<i>heki</i>	<i>kooli</i>	WEAK	VOWEL

There is no element that functions as a case marker in this subsystem. There is, in particular, no sub-word formative in Table 1 that signals partitive singular case. The stem of the partitive singular, *ˈhekk-* or *ˈkool-*, also realises the nominative singular. The theme vowel *-i*, which distinguishes the partitive and nominative singular, also marks the genitive singular. In short, ‘partitive singular’ is a word-level property that cannot be associated with any sub-word unit. The weak stem likewise does not signal genitive case in isolation, but only in contrast to a strong partitive, as many nouns are weak through their full paradigm. Moreover, it is precisely the lack of a theme vowel or case exponent that unambiguously identifies *ˈhekk* as a nominative singular. A noun that ends in *-k* can only be nominative singular, because every **other** case form in Estonian ends in a vowel or a case exponent. The interpretation of *ˈhekk* cannot be established by isolating any syntagmatic ‘part’, but only by contrasting *ˈhekk* with the set of alternative case forms.

Weakening declensions are remarkably efficient, yet in ways that defy description in agglutinative terms. The subsystem in Table 1 maximises the case contrasts that can be expressed by means of a strong stem, a weak stem and a theme vowel. From the elements *ˈhekk*, *hek-* and *-i* or *ˈkool-*, *kool-* and *-i*, one can define the forms *ˈhekk*, *ˈhekki*, *heki* and *hek*, and *ˈkool*, *ˈkooli*, *kooli* and *kool*. The first three elements of these series are acceptable words, which realize nominative, partitive and genitive singular in Table 1. But a weak stem in isolation, *hek* or *kool*, is not acceptable, because words are minimally bimoraic in Estonian, and only ‘Q3’ monosyllables contain two moras (cf. Prince 1980 and section 2.2). So the three distinct case forms in Table 1 represent the maximum that can be defined from the formatives in Table 1. The paradigms of HEKK and KOOL thus exploit efficient combinations, which is orthogonal to property-form biuniqueness.

Unlike cases of cumulative or extended exponence that might be described in terms of the ‘empty’ or ‘portmanteau’ morphs in Hockett (1947), the descriptive challenge illustrated in Table 1 is not that particular formatives realize ‘too few’ or ‘too many’ properties. The problem is that the sole function of individual formatives is to distinguish the word forms that realise case values. A morpheme-based analysis of the pattern in Table 1 is confronted with the fact that none of the individual formatives function as case morphemes. Case properties are realised by the word forms in Table 1, and words are characterised by distinctive **combinations** of formatives. But that is as far as one can extend a property-form relation without creating gratuitous problems of analysis.

Apportioning case properties to individual formatives in Table 1 not only misidentifies the locus of the property-form correspondence, but also applies the wrong logic to the analysis of this subsystem. The post-Bloomfieldian model does not seek minimal elements purely for their own sake, but because it assumes that their properties can be ‘summed’ to give the properties of other forms that are composed of the same elements. This analytic technique is based on the

assumption that recurrent elements will make a consistent contribution to the forms that they underlie. But this assumption is patently false in Estonian. One can identify contrastive properties and recurrent formatives in Table 1. However, individual formatives cannot be assigned a grammatical meaning from which the meanings of word forms can be determined. This is, of course, typical of theme vowels. What distinguishes Estonian is that stems are ‘sub-meaningful’ in much the same way, so that a stem ‘can be defined . . . in terms of which forms are built on it, which is to say in terms of its place in the morphological system of the language’ (Aronoff 1994:167).

This ‘morphomic’ pattern is propagated through a system of stems based ultimately on the formatives in Table 1. The chart in Table 2 exhibits the stem syncretism in weakening declensions though, as will become clear below, many of these patterns apply more generally within Estonian.

Table 2. Stem syncretism in weakening declensions

	Singular					Plural		
	Grammatical Cases			‘Fusional’ Cases		Grammatical Cases		
	Nom	Gen	Part	Illa2 Sg	Part2 Pl	Nom	Gen	Part
Form	`hekk	heki	`hekki	`hekki	`hekke	hekid	`hekkide	`hekkisid
	`kool	kooli	`kooli	`kooli	`koole	koolid	`koolide	`koolisid
Stem	Strong	Weak	Strong	Part Sg	Part Sg	Gen Sg	Part Sg	Part Sg

As in Table 1, the nominative and partitive singular are based on the strong stem, and the genitive singular is based on the weak stem. Nouns with vowel-final partitives usually have two additional forms based on the partitive singular, designated ‘ILLA2’ and ‘PART2’ in Table 2. The first is a ‘short’ illative singular, which is identical to the partitive singular in the case of *`hekki* and *`kooli*. The second is a ‘stem’ partitive plural, which preserves the stem of the partitive singular but ends in a different—and generally predictable—thematic vowel, *-e* in the case of *`hekke* and *`koole*.

The singular case forms also provide bases for the plural grammatical cases. The partitive singulars *`hekki* and *`kooli* underlie the partitive plurals *`hekkisid* and *`koolisid*, as well as the **genitive** plurals *`hekkide* and *`koolide*. The genitive singulars *heki* and *kooli* likewise underlie the nominative plurals *hekid* and *koolid*. In general, both the case and number properties of a base: genitive singular *heki/kooli* or partitive singular *`hekki/`kooli*, may differ from the properties of a form that it underlies: nominative plural *hekid/koolid* or genitive plural *`hekkidel/`koolide*. These patterns are strikingly ‘parasitic’ or ‘Priscianic’, in the sense of Matthews (1972), in that the plural grammatical cases are based on the **form**, not the **entry**, of a singular grammatical case.

This Priscianic pattern is propagated through the system of 11 ‘semantic’ cases. Each semantic case form is marked by a number-neutral case exponent.

Yet the base of a semantic case form is a morphomic stem, corresponding to the genitive form. The singular semantic cases in Table 3 are based on the genitive singulars *heki* and *kooli*, and the plural semantic cases are based on the genitive plurals *hekkide* and *koolide*. The allative plural [[[*hekk*]*i*]*de*]*sse* thus contains three levels of morphomic stems. The basic stem *hekk* realises nominative singular in isolation and underlies a second stem, *hekki*. The stem *hekki* realises partitive singular in isolation, and underlies a third stem, *hekkide*. The stem *hekkide* realises genitive plural in isolation and underlies the plural semantic cases. The entire stem system is morphomic, from the simple components of singular grammatical cases in Table 1 through to the complex bases of the semantic cases in Table 3.

Table 3. Parasitic semantic case forms

	Sing	Plur	Sing	Plur	Suf
Genitive	heki	hekkide	kooli	koolide	
Illative	hekiye	hekkidesse	koolisse	koolidesse	-sse
Inessive	hekiye	hekkides	koolis	koolides	-s
Elative	hekiye	hekkidest	koolist	koolidest	-st
Allative	hekiye	hekkidele	koolile	koolidele	-le
Adessive	hekiye	hekkidel	koolil	koolidel	-l
Ablative	hekiye	hekkidelt	koolilt	koolidelt	-lt
Translative	hekiye	hekkideks	kooliks	koolideks	-ks
Terminative	hekiye	hekkideni	koolini	koolideni	-ni
Essive	hekiye	hekkidena	koolina	koolidena	-na
Abessive	hekiye	hekkideta	koolita	koolideta	-ta
Comitative	hekiye	hekkidega	kooliga	koolidega	-ga
		'hedge'		'school'	

1.2. Morphological implication

The declensional system of Estonian conforms to the post-Bloomfieldian model in forming meaningful units from sub-meaningful elements. However, the sub-meaningful elements are formatives and stems, not phonemes, and the meaningful units are words, not morphemes. One cannot assign a determinate morphosyntactic value to the formatives in Table 1 or to the stems in Tables 2 and 3 because the meanings of these elements are context dependent. In Estonian, as in many morphological systems, words function as lexical **constructions**, in the sense that this term has come to be understood in other domains of grammar.² A morphological analysis cannot merely assemble discrete chunks of information associated with individual elements, but must recognise the contribution made by distinctive **combinations** of elements. As with other types of constructions, the whole guides—and may even determine—the selection of its parts, but the properties of the whole are more than the sum of the properties of its parts.

A model that insists on assigning meanings to minimal elements in Estonian declensions will not only create self-inflicted problems of analysis, but also fail to capture the organising principles of this system. An agglutinative analysis is particularly insensitive to the role that morphomic stems play in the **implicational** structure of a paradigm. The paradigm of an Estonian noun is not merely ‘the complete set of surface word forms that can be projected from the members of its stem set’ (Anderson 1992:134). Instead, Estonian declensions comprise networks of word forms which are connected by implicational rather than derivational relations. Each of the declensional types is organised around a set of **principal parts** from which one can predict the other forms of a paradigm. Traditional descriptions tend to list the three singular grammatical case forms: the nominative, genitive and partitive singular. However, in weakening declensions, the partitive singular functions as a ‘leading form’ or **kennform**, which reliably identifies the class of an open-class noun and, moreover, predicts the full paradigm of a noun. The implicational relations between the fusional case forms of HEKK and KOOL are set out in Table 4.

Table 4. Implicational structure of fusional grammatical cases

	Kennform	Nom Sg	Gen Sg	Illa2 Sg	Part2 Pl
Form	`hekki `kooli	`hekk `kool	heki kooli	`hekki `kooli	`hekke `koole
Relation		truncation	weakening	identity	exchange

A strong vowel-final partitive singular, *`hekki* or *`kooli*, implies a weak genitive singular, *heki* or *kooli*, as no open-class declension contains both a strong vowel-final partitive and a strong genitive singular. A strong vowel-final partitive singular also implies a ‘truncated’ nominative singular, *`hekk* or *`kool*, which lacks the theme vowel of the partitive. A vowel-final partitive singular also implies the ‘short’ illatives and stem partitives in Tables 2 and 4. Short illative singulars are minimally **trimoraic**. A strong partitive is already trimoraic, as the Q3 strong stem contributes two moras and the final syllable adds a further mora. Hence a strong partitive, *`hekki* or *`kooli*, implies an identical short illative singular. The stem partitive plural is likewise an ‘exchange variant’ of the partitive singular (cf. Mürk 1997:16). If the partitive singular ends in *-i*, the partitive plural ends in *-e*, and conversely. Thus *`hekki* and *`kooli* imply *`hekke* and *`koole*.

A vowel-final partitive singular also implies a genitive plural in *-de* and a ‘long’ partitive plural in *-sid*. Both genitive forms are predictable from the partitive singular, and these genitives in turn identify the base of semantic case forms. So the full paradigm of HEKK and KOOL is implied by the single kennformen *`hekki* and *`kooli*. Other noun types are based on other principal parts, which are described in section 2.1, but the implicational patterns are similar. Many of these implications are, moreover, ‘reversible’. The kennformen *`hekki* and *`kooli*

are identified by the genitive plurals *˚hekkide* and *˚koolide*, as well as by plural semantic case forms based on the genitive.

These patterns of mutual implication hold the key to the stem syncretism in Tables 2 and 3. The ultimate explanation for many of these patterns is plainly diachronic (Grünthal 2003:51, Viitso 2003b:162–168). But the synchronic function of these stems also cannot be understood in terms of their morphosyntactic ‘content’ alone, and must take into account their role within a system of implications that operates over paradigms, not single forms. It is because patterns at this level are inaccessible to methods of segmentation and classification that the ‘signal’ of a word-based system comes out as the ‘noise’ of non-biunique exponence in a post-Bloomfieldian model.

As this overview of patterns in weakening declensions suggests, property-form ‘mismatches’ are symptoms of a deeper problem. The challenge in Estonian is not that there are incomparable numbers of form and content units, but rather that no correspondence can be established between these units. The lack of a correspondence reflects the fact that the properties of an inflected form in Estonian cannot in general be determined from the properties of its parts. Instead, sub-word units serve to distinguish a word from other members of its inflectional paradigm, and the properties of the word are determined by its place within this larger pattern of forms.

2. WORD-BASED EXPONENCE AND IMPLICATION

This section shows how a WP analysis brings out the organisation of the Estonian declensional system. As recognised in traditional descriptions, the class of an Estonian noun is identifiable from the prosodic and morphotactic structure of one or more principal parts. Matching the principal parts of a noun against the corresponding forms of an exemplary paradigm permits the analogical deduction of other forms. Since form-based analogies use principal parts as the basis for **predicting** forms, they accommodate patterns of morphomic stem syncretism that raise difficulties for analyses that attempt to **build** complex forms from smaller meaningful units.

The basic components of this analysis are set out in sections 2.1–2.3. The first is a set of exemplary paradigms or class patterns. The second is a characterisation of the principal parts that identify the class of a given noun. The third is a system of analogies that allows one to deduce a predictable form of a noun from its principal parts and the forms of an exemplary paradigm.

2.1. Declension classes

Traditional descriptions of Estonian noun paradigms represent regular nouns and adjectives by a set of *põhivormid* (Erelt et al.2000:155), ‘basic forms’ (Viks

1992:39) or ‘principal parts’ (Mürk 1997:12). This set minimally includes the nominative, genitive and partitive singular forms, and may also contain the genitive and partitive plural, and even the short illative singular. Nouns are grouped into classes or *käändkonnad* ‘declensions’ based on the prosodic structure of one or more principal parts, and on patterns of exponence and stem selection within the principal part inventories. Although there is no absolute consensus regarding the number of classes and subtypes in Estonian, most classifications distinguish at least the four productive classes in Table 5.³

Each class in Table 5 exhibits a characteristic prosodic structure and distinctive patterns of exponence. Although it is traditional to list the three singular grammatical cases for each noun, the nominative singular tends to be the least informative, most predictable, and least prosodically consistent form. Class membership is usually predictable from the prosodic structure of the genitive singular and/or the ending of the partitive singular. Matching these kennformen of a noun against the corresponding cells of an exemplary paradigm identifies class-specific forms.

Class 1 nouns have a trochaic genitive singular and a vowel-final partitive singular. The genitive singular exhibits a strong–weak trochaic pattern, with primary stress on the first syllable, and the partitive plural ends in a theme vowel. In class 1a, the genitive and partitive singulars are identical, so that the forms *pesa* and *seminari* realise the genitive and partitive singular. In ‘weakening’ class 1b, the genitive singular is a weak counterpart of the partitive singular, as illustrated by the genitive~partitive pairs *kooli~`kooli* and *tuleviku~tule`vikku*. Since only class 1 nouns have a vowel-final partitive singular, this form alone suffices to identify a class 1 noun.

Class 2 nouns have a non-trochaic genitive singular and a partitive singular in *-t*.⁴ Class 2a contains monosyllables (*`tee~teed*), and iambic bisyllables (*i`dee~i`deed*). Class 2b contains nouns with an initial Q3 foot that remains strong through their entire paradigm (*`aasta~`aastat*), and most nouns with trisyllabic genitive singulars (*raamatu~raamatut*). Class 2c contains nouns with trisyllabic genitive singulars in *-se*, as illustrated by *hobuse~hobust*, and *otsuse~otsust*.

Class 3 nouns combine a penultimate primary stress in the singular grammatical cases with a partitive singular in *-t*.⁵ The pairs *auto~autot* or *kõne~kõnet* ‘speech’ can only belong to class 3: the partitive singulars in *-t* cannot be associated with class 1, and the trochaic structure of the genitive singulars does not conform to the class 2 pattern. In longer genitive singulars, such as *gorilla*, *lauljanna*, *šampánja* ‘champagne’ or even *primadonna* (in which stress is marked by an acute accent), the penultimate stress pattern is an unambiguous marker of class 3 membership.⁶

Class 4 contains nouns whose partitive singulars end in *-st* and whose genitive singulars consist of two feet, the second of which is a trochee. The quadrisyllabic genitive singulars *küsimuse* and *inimese* consist of two trochaic feet, whereas *`alguse* or *ÿõulise* ‘forceful’ consist of an initial Q3 foot, followed by a trochee.

Table 5. Major noun declensions in Estonian

	Nom Sg	Gen Sg	Part Sg	Ill2 Sg	Part2 Pl	Gen Pl	Part Pl
1a	pesa seminar	pesa seminari	pesa seminari	`pessa seminari	pesi seminare	pesade seminaride	pesasid seminarisid
1b	`kool tule`vik	kooli tuleviku	`kooli tule`vikku	`kooli tule`vikku	`koolle tule`vikke	`koolide tule`vikkuude	`koolisid tule`vikkusid
2a	`tee i`dee	`tee i`dee	`teed i`deed	— —	— —	`teede i`deede	`teid/`teesid i`deid/i`deesid
2b	`aasta raamat	`aasta raamatu	`aastat raamatut	— —	— —	`aastate raamatute	`aastaid raamatuid
2c	hobune otsus	hobuse otsuse	hobust otsust	— —	— —	hobuste otsuste	hobuseid otsuseid
3	gorilla lauljanna	gorilla lauljanna	autot gorillat lauljannat	— —	— —	autode gorillade lauljannade	autosid gorillasid lauljannasid
4	`algus küsimus inimene	`alguse küsimuse inimese	`algust küsimust inimest	`algusse küsimusse inimesse	`algusi küsimusi inimesi	`alguste küsimuste inimeste	— — —
							'nest' 'seminar' 'school' 'future' 'road' 'idea' 'year' 'book' 'horse' 'decision' 'car' 'gorilla' 'fem. singer' 'beginning' 'question' 'person'

Given that a partitive singular in *-st* and a genitive singular in *-se* imply one another, class 4 nouns can be identified just from the form of their genitive singular.

The prosodic and morphotactic properties of these kennformen are summarised in Table 6.

Table 6. Prosodic and morphotactic structure of primary kennformen

	Gen Sg	Prosody	Part Sg	1σ	2σ	3σ	4σ
1a	<i>Xv</i>	trochaic	<i>Xv</i>	—	<i>pesa</i>	—	<i>seminari</i>
1b	<i>Xv</i>	trochaic	<i>˘Xv</i>	—	<i>˘kooli</i>	—	<i>tule˘vikku</i>
2a,b	<i>X</i>	non-trochaic	<i>Xt</i>	<i>˘tee</i>	<i>i˘dee/˘aasta</i>	<i>raamatu</i>	—
2c	<i>Xse</i>	non-trochaic	<i>Xst</i>	—	—	<i>otsuse</i>	—
3	<i>X</i>	penult stress	<i>Xt</i>	—	<i>auto</i>	<i>gorilla</i>	<i>primadonna</i>
4	<i>Xse</i>	2 feet, 2nd trochaic	<i>Xst</i>	—	—	<i>˘alguse</i>	<i>küsimuse</i>

2.2. Characterization of principal parts

The chart in Table 6 isolates some of prosodic and morphotactic factors that underlie traditional classifications of nouns and adjectives. A prosodic description of these classes in Table 5 can also be couched in most familiar models of prosodic analysis, such as Selkirk (1980) or McCarthy and Prince (1995). The central descriptive challenge that Estonian presents for any prosodic analysis concerns the treatment of the three-way contrast between short (Q1), long (Q2) and ‘overlong’ (Q3) syllables. As Viitso (2003a:11) observes, this ternary contrast reflects two binary distinctions. The first is a segmental length contrast between short Q1 and long Q2/Q3 syllables. The second is a prosodic weight contrast between light Q1/Q2 syllables and heavy Q3 syllables. This is not the place to review the extensive and largely inconclusive phonological literature on this topic. Of primary importance is the fact that only the weight contrast is relevant for the inflectional system. All quantitative grade alternations involve a binary contrast between a stressed Q3 and a non-Q3 syllable. Estonian reinforces this contrast with a binary split between Q3 syllables, which can function as words in isolation, and Q1 and Q2 syllables, which cannot. Conversely, class 3 nouns may contain an initial Q1 syllable, as in *kõne* ‘speech’ or *ratsu* ‘steed’, or an initial Q2 syllable, as in *auto* or *kiisu* ‘kitty’, but may not contain a Q3 syllable (Erelt et al.2000:247).

In short, the morphological system makes a clear binary split between ‘heavy’ Q3 syllables and ‘non-heavy’ Q1 and Q2 syllables. Moreover, the weight-sensitive processes of Estonian treat a Q3 syllable as the equivalent of two non-Q3 syllables. As noted above, a minimal word may consist of a single Q3 syllable, as in *˘tee* or *˘kooli*, two Q1 syllables, as in *pesa* or *kivi* ‘stone’, or a Q2 and Q1 syllable, as in *kooli* or *kiisu* ‘kitty’. However, a word may not consist of a single Q1 syllable,

such as *pes* or *kiv*, or even a Q2 syllable, such as *kool* or *kiis*. Class 4 nouns exhibit a parallel correspondence, as their genitive singular kennform may consist of an initial Q3 syllable or an initial pair of non-Q3 syllables. To capture the fact that one Q3 syllable corresponds in weight to any two non-Q3 syllables, Ehala (2003) and Blevins (2004) analyse Q3 syllables as bimoraic, and treat both Q1 and Q2 as monomoraic syllables that differ in segmental length.⁷

Provided that the marked member of the opposition is Q2 (historically shortened Q3), one could consider Q1 a normal short and light (monomoraic) quantity and Q3 a normal long and heavy (bimoraic) quantity. Q2 would be something in between: segmentally long, but light by weight (monomoraic). (Ehala 2003:58)

2.2.1. Minimal words and nominative singulars

This treatment of syllable weight permits a straightforward description of the minimal word constraint that applies to open-class word forms in general, and the prosodic constraints that apply to nominative and short illative singular forms in particular. Adapting Prince (1980), feet can be defined as minimally bimoraic, and prosodic words as consisting of at least one foot. Assuming that only Q3 syllables are bimoraic and qualify as feet provides an account for the fact that monosyllables must be Q3, as illustrated by class 2 nouns such as *ʔee* or *ʔoi* ‘moth’.

The minimal word constraint also conditions the form of nominative singulars, as McCarthy and Prince (1986:5) note, though morphotactic structure may play a role as well. The nominative singular of a regular Class 1 noun corresponds to the stem of the partitive singular, provided that the stem constitutes a foot. Thus *KOOL* has the partitive singular *ʔooli* and the nominative singular *ʔool*, and *SEMINAR* has the partitive singular *seminari* and the nominative singular *seminar*. But since *pes-* alone is not bimoraic, hence not a foot, the theme vowel in the partitive singular *pesa* is preserved in the nominative singular *pesa*. The stem of a class 4 noun is always bimoraic, and class 4 nouns with genitive singulars *-use* have nominative singulars in *-s*. This is illustrated by the pairs *ʔalguse~ʔalgus* and *küsimuse~küsimums*. But class 4 nouns with a genitive singular in *-ese* or *-ise* are predominantly adjectives, and have a nominative singular in *-ene* or *-ine*.⁸

Monomorphemic class 2 nouns also tend to have truncated nominative singulars. Thus *raamatu* and *otsuse* alternate with *raamat* and *otsus*. A class of mainly adjectival class 3 forms, including the noun *hobuse*, but more typically *soolase* ‘salty’, have nominative singulars in *-ne*. This represents the one productive nominal subtype in which the form of the nominative singular is genuinely informative, though in the majority of cases, the form of the nominative singular is predictable from the syntactic category of a noun with a genitive singular in *-use*.

Class 2 nouns whose genitive singulars end in a derivational suffix also lack truncated nominative singulars. This group includes nouns with trisyllabic genitive singulars such as *sigala* ‘pigsty’ (cf. *SIGA* ‘pig’), in which *-la* marks ‘nouns expressing a place for a certain action’ (Viitso 2003a:81), caritive adjectives in *-tu* such as *abitu* ‘helpless’ (cf. *ABI* ‘help’), along with nouns, such as *˘söökla* (cf. *söök* ‘meal’) or *˘laulja* ‘singer’ (cf. *LAULMA* ‘sing’), whose bisyllabic genitive singulars contain an initial Q3 syllable that is strong through the entire paradigm.⁹ Given that ‘non-trochaic’ class 2 nouns are already somewhat heterogeneous, one could recognise these morphotactically complex formations as a particular subtype of class 2, or else attribute the lack of truncation to a constraint against resyllabification across morph boundaries, or to some other condition. In either case, the class of these nouns, and the form of their nominative singulars, is predictable from the prosodic and morphotactic structure of the genitive singular kennform.

2.2.2. Principal part deduction

The fusional illative and partitive cases are similarly predictable from kennformen in classes 1 and 4. As noted above, illative singulars are minimally trimoraic. In class 1, short illative singulars are based on the partitive singular. A bisyllabic partitive singular with an initial Q3 syllable, such as *˘kooli*, or a quadrisyllabic partitive singular, such as *seminari*, implies an identical short illative singular. In paradigms with a bimoraic partitive singular, such as *pesa*, the short illative is a Q3 counterpart, here *˘pessa*. The genitive singular forms of class 4 nouns all satisfy the trimoraic length requirement, so there is no ‘strengthening’ in class 4. Instead, short illatives are distinguished by a purely segmental lengthening process in the final syllable. Thus the final syllables *-se* in the genitive singulars *˘alguse*, *küsimuse* and *inimese* all differ in length but not in weight from the final syllables *-sse* in the short illative singulars *˘algusse*, *küsimusse* and *inimesse*.

The ‘stem’ partitive plurals are based on the same kennformen as the short illatives, but exhibit a distinctive pattern of exponence. Stem partitive plurals operate by a system of ‘vowel reversal’ (Matthews 1991:199) or, more generally, ‘vowel exchange’. In class 1, a partitive singular in *-i* implies a partitive plural in *-e*, and conversely. The first of these patterns is illustrated by the pairs *˘kooli*~*˘koole* and *˘hekki*~*˘hekke* and the second by the pairs *˘kukke*~*˘kukki* ‘rooster’ and *˘lille*~*˘lilli* ‘flower’. A partitive plural in *-e* is also implied by a partitive singular in *-u*, as illustrated by *˘lukku*~*˘lukke* ‘lock’, and a partitive plural in *-i* is implied by a partitive singular in *-a*, as illustrated by *pesa*~*pesi* and *mokka*~*mokki* ‘lip’. Since the genitive singular kennformen of class 4 nouns end in *-se*, the corresponding stem partitive plural ends in *-si*. Thus class 4 exhibits the characteristic alternations *˘alguse*~*˘algusi*, *küsimuse*~*küsimusi*, and *inimese*~*inimesi*.¹⁰

The prosodic and morphotactic structure of genitive and partitive singular kennformen thus identify word class, and also determine the form of other principal parts.¹¹ In contrast, the other principal parts are **not** predictable in general

from the nominative singular citation form. Although Estonian preserves traces of a vowel harmony system, reconstructable for proto-Finnic (Viitso 2003b:173) or even Uralic (Laakso 2001:83), theme vowels are not predictable from noun stems. In ‘truncating’ classes 1 and 2 paradigms, the theme vowel is therefore not recoverable from the nominative singular. Without the theme vowel, one cannot define the genitive or partitive singulars, or the case forms that are based, directly or indirectly, on these principal parts.

The deficiencies of the nominative singular are, moreover, shared by any prosodic unit smaller than the genitive and partitive singular. Not only are stems and theme vowels of limited predictive value, but, as noted in section 1.1 above, they do not carry any grammatical properties in isolation. The difference between the forms *heki* and *pesi*, for example, cannot be attributed to any aspect of their syntagmatic structure. No part of the form *heki* signals that it realises genitive singular, rather than, like *pesi*, partitive plural. These forms have parallel morphotactic structures, down to the choice of the final vowel *-i*. Yet they occupy different niches in their respective paradigms, and it is this association that determines their grammatical meaning. The genitive singular interpretation of *heki* reflects its place within the paradigm of HEKK ‘hedge’ and the partitive plural interpretation of *pesi* reflects its place within the paradigm of PESA ‘nest’. The interpretation of the strong forms *hekki* and *kukki* ‘rooster’ is similarly dependent on paradigmatic context. Both forms contain a strong stem and the vowel *-i*. But *hekki* is part of the singular–plural opposition *hekki*~*hekke*, whereas *kukki* is part of the inverse pattern *kukke*~*kukki*.

The meaning of forms such as *heki* and *pesi* or *hekki* and *kukki* is thus context-dependent. But the context is paradigmatic, as the grammatically significant part-whole relations in Estonian declensions hold between words and paradigms, not between morphemes and words. Relations at this level clearly fall outside the descriptive scope of post-Bloomfieldian procedures of segmentation and classification. Even more generally, these declensional patterns challenge any strategy that isolates minimal elements for the purpose of ‘building’ other forms. Techniques of ‘isolation and recombination’ are designed to relate forms by deriving them from recurrent elements that contribute a constant meaning in each of their uses. However, the members of nominal paradigms in Estonian are related in a different way, via patterns of implication.

2.3. Analogical formations

These implicational patterns are traditionally expressed in terms of procedures of **analogy**. Analogy operates at two levels in Estonian declensions. The first level involves cross-paradigm analogies that deduce the principal parts of a noun from the basic kennformen of the noun and an exemplary set of principal parts. The second level involves paradigm-internal analogies that deduce the grammatical and semantic case forms that are based on the principal parts.

Traditional descriptions of Estonian declensions exploit both types of analogical formation. Viks (1992:39–47) identifies six basic forms, and defines ‘rules of analogy’ that apply to the genitives. The *Eesti keele sõnaraamat* (Erelt 1999) gives exemplary principal part inventories for each of the ‘word types’ (*tüüpsõnad*) that it recognises, and then identifies the type of each noun that it lists. Erelt et al. (2000:154) likewise distinguish a class of ‘basic forms’ (*põhivormid*), as in section 2.1, and define a larger class of ‘analysed forms’ (*analoožiavormid*) as ‘forms that can be formed on analogy to some basic form’ (*vormid, mida saab moodustada mingi põhivormi analoogia*).

Traditional analogical models supply the paradigmatic context that resists description in syntagmatic terms. For example, the Q3 partitive singular *hoovi* identifies the noun *hoov* ‘yard’ as a grade-alternating noun of class 1. Given this kennform, and the exemplary forms of *kool* in Table 5, one can deduce the corresponding forms of *hoov*. Each deduction can be expressed as a standard four-part proportional analogy. To determine the short partitive plural of *hoov*, one matches the partitive singular kennform *hoovi* against the exemplary form *kooli*, as in Table 7a, and then ‘solves for *X*’, as in Table 7b. The forms *hoov*, *hoovi* and *hoovi* can be defined similarly.

Table 7. Form-based
analogical deduction

-
- a. $\text{`kooli} : \text{`hoovi} = \text{`koole} : X$
 b. $X = \text{`hoove}$
-

Although traditional accounts characteristically refer to ‘forms’, they tend to mean ‘forms with a given interpretation’, which amount, in effect, to entries. The intended interpretation of Table 7 can be expressed more explicitly in Table 8, using the ‘realisation pair’ notation of Aronoff (1994).

Table 8. Entry-based analogical deduction

-
- a. $\langle [\text{Part Sg}], \text{`kooli} \rangle : \langle [\text{Part Sg}], \text{`hoovi} \rangle = \langle [\text{Part Pl}], \text{`koole} \rangle : \langle [\text{Part Pl}], X \rangle$
 b. $X = \text{`hoove}$
-

The refinement in Table 8 brings out the tacit assumption that analogical deduction involves matching a ‘leading entry’ against the corresponding cell in an exemplary paradigm. The sole function of the properties in a leading entry is to guide this matching. The deductions in Tables 7 and 8 do not ‘construct’ a derived form or entry from the exemplar, but merely use the exemplar as a model for forming the deduced element. A kennform or leading entry is therefore not a kind of ‘basic unit’ that underlies analysed forms, but rather a ‘hook’ into a deductive pattern.

The ‘non-derivational’ character of analogical deductions allows them to capture morphomic paradigm-internal patterns as well. The chart in Table 9

Table 9. Priscianic stem syncretism in semantic and plural grammatical cases

	Singular Grammatical Cases		Plural Grammatical Cases			Semantic Cases		
	Nom	Gen	Part	Nom	Gen	Part	Illa Sg	Illa Pl
1	*kool	kooli	*kooli	koolid	*koolide	*koolisid	koolisse	*koolidesse
2	raamat	raamatu	raamatut	raamatud	raamatute	raamatuid	raamatusse	raamatutesse
3	auto	auto	autot	autod	autode	autosid	autosse	autodesse
4	küsimus	küsimuse	küsimust	küsimused	küsimuste	küsimusi	küsimusesse	küsimustesse

reviews the Priscianic structure of the case forms that are based, directly or indirectly, on the singular grammatical cases.

Forms such as *kooli* or *ʔkooli* make no uniform morphosyntactic contribution to the case forms that they underlie, other than in identifying those elements as forms of the noun *KOOL*. To paraphrase Matthews (1991:200), there is no sense in which the meaning of the genitive plural includes that of the partitive singular, or in which the meaning of the long illative plural includes that of the genitive plural. Rather, the illative plural *ʔkoolidesse* is based on the form of the genitive plural *ʔkoolide* and *ʔkoolide* is based in turn on the form of the partitive singular *ʔkooli* in the same way that the partitive singular is itself based on the sub-meaningful strong stem *ʔkool*.

These pure form correspondences between members of a paradigm can again be expressed as analogical deductions. If one uses ‘*Xv*’ to represent a vowel-final form, the formation of genitive and partitive plurals in class 1 can be expressed by the simple two-part analogies in Table 10.

Table 10. Paradigm-internal analogy
in class 1

-
- a. $\langle [\text{Part Sg}], Xv \rangle = \langle [\text{Gen Pl}], Xvde \rangle$
 b. $\langle [\text{Part Sg}], Xv \rangle = \langle [\text{Part Pl}], Xvsid \rangle$
-

Nominative plural and semantic cases are defined by the more general deductions in Table 11.

Table 11. General
paradigm-internal analogy

-
- a. $\langle [\text{Gen Sg}], X \rangle = \langle [\text{Nom Pl}], Xd \rangle$
 b. $\langle [\text{Gen}], X \rangle = \langle [\text{Illa}], Xsse \rangle$
-

In an analogical pattern, the properties of the antecedent identify the value of the form variable *X*. These properties are explicitly not associated with the consequent, but other properties of the matching entry are preserved. Hence the nominative plural defined in Table 11a differs in case and number from its genitive singular base, whereas the long illatives defined in Table 11b preserve the number features of the corresponding genitives. The illative singular *koolisse* is parasitic on the singular genitive, while the illative plural *ʔkoolidesse* is formed analogically from *ʔkoolide*. But the stems *kooli* and *ʔkoolide* do not contribute number properties to the forms that they underlie, any more than they contribute case properties, because semantic forms are deduced from these stems, not ‘constructed’ from stems through the addition of properties and exponents.

The remaining semantic case forms of *KOOL* can be formed by similar analogies. A combination of general and class-specific analogies will also define the

paradigm of other open-class nouns. In each case, a noun can be identified by its kennformen, and any of its principal parts that are below the frequency threshold for storage can be defined by analogy. The principal parts will in turn imply the form of the plural grammatical cases, and, ultimately, the semantic cases.

Exactly the same strategy applies to irregular nouns. Given that irregularity is largely concentrated in the singular grammatical cases, many non-productive patterns can be represented by associating a noun or noun class with an exceptional set of principal parts, with as many forms as are needed to represent additional irregularities or distinctive patterns of stem syncretism. The nouns TÜVI ‘stem’, TULI ‘fire’ and KÄSI ‘hand’ in Table 12 illustrates successively more irregular patterns, each of which is followed by under a dozen nouns (Erelt et al. 2000:240–241). The irregularity of TÜVI is essentially confined to the exceptional nominative singular in *-i*. To this pattern, the paradigm of TULI adds the exceptional partitive singular *˘tuld*. The plural grammatical cases of both nouns are based on the genitive singular kennform, and the semantic cases in turn on the genitive singular and genitive plural forms. At the other extreme, all of the forms of KÄSI in Table 12 must be stored, apart from the nominative singular. But given these forms, the semantic cases can be formed analogically from the genitive singular *˘käe* and genitive plural *käte*.

Table 12. Degrees of irregularity in non-productive class 1 nouns

No. of nouns	Sing			‘Fusional’ Cases		Plur		
	Grammatical Cases					Grammatical Cases		
	Nom	Gen	Part	Illa2 Sg	Part2 Pl	Nom	Gen	Part
8	tüvi	tüve	tüve	˘tävve	—	tüved	tüvede	tüvesid
6	tuli	tule	˘tuld	˘tulle	—	tuled	tulede	tulesid
9	käsi	˘käe	˘kätt	˘kätte	˘käsi	käed	käte	—

2.4. Summary

The analysis above factors the declensional system of Estonian into two traditional parts. The first is a set of exemplary paradigms (or, alternatively, sets of exemplary grammatical case inventories), and the second is a smaller set of kennformen for each non-exemplary noun or adjective. These components are related by word-based principles of analogy that permit the deduction of novel forms from existing forms or patterns. Since analogical principles are invoked to deduce forms that are not already part of the speaker’s lexicon, they do not apply if the lexicon already contains a suppletive form that is frequent enough to maintain its irregular characteristics.

Principles of word-based analogy also provide a general analysis of overlapping exponence in basic forms and Priscianic stem syncretism in complex case

forms. If one regards stems and exponents as abstractions over whole word forms, as proposed by Kurylowicz (1949), among others, morphomic patterns raise no difficulties of any kind. Full word forms will have determinate properties, and be distinguished by patterns of prosody, exponence and/or stem selection. However, there will be no reason to expect that any individual pattern or sub-word unit will be biuniquely associated with differences in meaning, except insofar as the patterns reflect recently or otherwise transparently grammaticalised sources that retain some morphosyntactic unity.

A speaker must learn that *ʔkooli* is the partitive singular of KOOL, as this information is not predictable from properties that can be associated with the stem *ʔkool-* or with the vowel *-i*. However, classifying *ʔkooli* as a partitive singular permits the analogical deduction of the entire paradigm of KOOL. The form *ʔkooli* predicts the genitive plural *ʔkoolide*, and ultimately the semantic case forms based on *ʔkoolide*, given that each of the semantic case exponents are constant. But this implicational structure cannot be attributed to the derivation of case forms from a set of ‘minimal meaningful’ units. The elements *ʔkool-* and *-i* contribute ‘units of form’ but no ‘units of content’ to *ʔkooli* in the same way that *ʔkooli* contributes form but not content to *ʔkoolide*, and *ʔkoolide* in turn contributes pure form to *ʔkoolidesse* and other semantic case forms. The patterns are transparent and systematic, but they are analogical and deductive, not derivational.

3. EXTENSIONS AND ALTERNATIVES

The traditional perspective adopted in the analysis in section 2 incorporates a number of idealisations and assumptions that are independent of the central claims. One important issue concerns the demarcation between exemplary paradigms and principal part inventories. There is compelling evidence that the forms of highly frequent nouns are stored by speakers, irrespective of their regularity (Baayen et al. 2003). Hence the artificiality of a traditional account is not that it relies on an exemplary paradigm for each inflection class, but that it arbitrarily selects a **particular** lexeme from the many that would be contained in the mental lexicon of most native speakers.

The idealisation that non-exemplary nouns are represented by a minimal principal part inventory requires a similar qualification. The key assumption of a traditional WP account is that an inflectional system can be factored into exemplary patterns, and that one can identify the appropriate pattern for a non-exemplary item from a principal part inventory that contains fewer forms than the pattern itself. An inflectional class system that cannot be factored in this way cannot be analysed in terms of a traditional WP model. However, there is no reason to assume that a speaker’s principal part inventories must therefore contain the **fewest** informative forms, any more than there is a reason to assume that a speaker must identify a unique exemplary paradigm. On a WP

analysis, principal part inventories must contain forms that identify class. But other factors, notably frequency, will determine whether additional forms are stored as well.

The selection of particular kennformen in section 2 also reflects traditional but inessential assumptions. Standard descriptions of Estonian declensions tend to classify nouns on the basis of sets of morphotactically simple forms, to the exclusion of sets of equally informative but more complex forms. This bias is reflected in the traditional choice of genitive and partitive singular kennformen in section 2.1. However, other forms or collections of forms may be equally informative. The short illative singular or stem partitive plural often suffice to identify the class of a noun that contains these forms. Noun class is also generally predictable from the partitive singular and either the partitive or the genitive plural, as discussed in Blevins (2004). Any pair of singular and plural semantic case forms will likewise identify the genitive singular and plural forms of a noun, as well as the base of the genitive plural. So the choice of simple kennformen, the same for each non-exemplary lexeme, serves mainly to identify the properties than any kennformen must have.

3.1. Referrals, families and constructions

On the traditional approach outlined above, an inflectional system consists of a stock of word forms, and a network of analogical principles that extend the patterns exhibited by these forms. This approach casts light on a number of contemporary idealizations. One assumption that is implicit, in one form or another, in many theoretical models, is that the analysis of a given form proceeds in isolation from all other forms of a system. This assumption is reflected in the fact that rules or combinatoric principles apply to a single form in nearly all morphological models, and that these rules define derivations that are fully independent of one another. Rules of ‘referral’ (Zwicky 1985, Stump 1993) are a prominent exception, as they permit one form to be defined with ‘reference’ to another. Yet rules of this type are deprecated or rejected altogether in many approaches, due in part to concerns about their seemingly ‘transderivational’ character.

From a traditional WP perspective, these types of fears are misplaced. A model in which novel forms are deduced by analogy to existing forms places referral-type relations at the centre of a morphological system. It is instead the rules that construct one form at a time from smaller, autonomous parts that represent the marginal and theoretically suspect device. A measure of empirical support for this traditional conception comes from studies of the effects of ‘morphological families’. Schreuder and Baayen (1997) and de Jong et al. (2000), among others, have found that the ‘family size’ of a simple form, i.e. the number of complex forms that it underlies, has a striking effect on lexical decision tasks. By controlling for the token frequency of family members, they are able to

establish that '[t]he larger the morphological family, the faster and more accurate[ly] participants decide whether a word is an existing word' (de Jong 2002:7). While there are questions about how to model these effects, the studies carry a clear implication that the members of a morphological family are not only listed, but activated in processing tasks.

A last general issue concerns the format in which productive patterns are expressed. Analogical relations between forms can be characterised in a variety of different ways, depending on the importance that one attaches to the symbolic representation of patterns. In any analogical model, the deductions that predict novel principal parts will exploit the implicational structure implicit in a set of exemplary forms. But this structure can either remain resident in existing form inventories, or it can be encapsulated in a separate system of rules, constraints, templates or schemas. For example, one could express the truncation of nominative singulars solely in the relation between a kennform and a truncated exemplary nominative singular, or else represent truncation symbolically, as a process that relates a kennform to a nominative singular. Similar remarks apply to the relation between a kennform and the stem partitive plural, etc.

The paradigm-internal analogies in Tables 10 and 11 can likewise be recast in more of a constructional idiom. On this alternative, the genitive plural is construed as a construction that selects a partitive singular stem and the exponent *-de*, the nominative plural as a construction that selects a genitive singular stem and the exponent *-d*, and the long illative as a construction that selects a genitive stem and the exponent *-sse*. Each analogised case form would correspond to a construction that determines the choice of a stem and ending. The result would have more of a 'top-down' structure, in that the whole would more explicitly determine its parts. But it would still be the entire construction, i.e. the full word that would carry grammatical meaning.

3.2. *Stem sets and morphotactic features*

Other types of alternative analyses may differ in what may at least appear to be more substantial respects. One fairly straightforward alternative to a traditional word-based analysis would be to describe Estonian declensions in terms of a system of abstract stems. The generally morphomic character of these stems would then reflect the assumption that they are not associated with any properties other than the 'intrinsic' features of the basic stem or root. Versions of this type of stem analysis are presented in Mürk (1997) and Hughes and Ackerman (2002). The question that arises for this type of account is whether it represents a genuine alternative. Apart from forms such as *hek-* or *hekk-*, which represent stems on nearly any account, the expanded class of 'stems' will turn out to comprise just those word forms that may underlie other word forms. With the exception of the genitive (and nominative) singular none of these stems will participate

in derivational processes or other processes that might confirm their sub-word status. Hence it is not obvious that calling *˚hekkide* a stem achieves anything other than preserving the generalisation that words are based on stems, rather than other words. But if preserving this generalisation involves designating all bases as stems, whether or not they exhibit any other stem-like behaviour, it becomes fairly clear that the only real difference is terminological.

A second, and, on the face of it, more radical alternative would adopt a neo-Jakobsonian perspective, and assign more abstract ‘meanings’ to the exponents in Estonian. The three basic grammatical cases could be described in terms of a pair of binary features, say $\pm F_1$ and $\pm F_2$. Seemingly sub-meaningful elements, such as strong stems and theme vowels could be assigned values for these features, and then the combination of abstract values on a word form would define a conventional case property, such as nominative, genitive or partitive. For example, a strong stem *˚hekk-* could be assigned the value $[+F_1]$, the theme vowel *-i* the value $[+F_2]$, and the property ‘partitive’ defined as $[+F_1, +F_2]$. Then the parts of *˚hekki* could contribute meanings that determine the properties of the whole word. The problem with this alternative is that it merely projects the sub-meaningful character of the formatives *˚hekk-* and *-i* onto a level of morphotactic features. The abstract feature analysis just redundantly mirrors the morphotactic structure. Most importantly, these features will not define any actual case values below the word level, since it is only at this level that the features of stems and exponents are combined. So the level of morphotactic features is not only redundant, but utterly inert. Case remains a word-level property, but the representation of case is mediated through an extraneous level of description. Clearly nothing is achieved by preserving the letter of a post-Bloomfieldian model in this way.

4. CONCLUSION

This brief discussion of alternatives brings the central claims of the paper into sharper relief. The most basic claim is that the declensional system of Estonian is word-based, and exhibits word-level patterns of exponence, stem syncretism and implication that cannot be expressed in terms of sub-word units. A related claim is that the properties of a noun form cannot in general be determined from its syntagmatic structure, but reflect its place in a larger set of forms. Estonian thus accords with a traditional WP perspective, in which form variation is distinctive at the level of words, but interpreted within the larger pattern of a language. The syntagmatic parts of a word distinguish it from other words, but the grammatical meaning of the word depends on the organisation of words into paradigms, and paradigms into inflection classes. Each noun form is distinguished from the other forms in its paradigm by a combination of characteristics that mark the same properties in a class of congruent paradigms. This system-level

congruence provides the basis for the analogical deduction of forms, within and across paradigms. In some cases, the full paradigm of a noun can even be deduced from a single, frequent, *kennform*.

Procedures of segmentation and classification apply at the wrong level of analysis to capture these patterns, and no amount of technical refinement will allow a post-Bloomfieldian approach to characterise the part-whole relation between words and paradigms without conceding the central claims of a WP model. Any model that assumes the biunique correspondence encapsulated in the structuralist morpheme is bound to regard a paradigmatic system 'as an agglutinating system that has somehow gone wrong' (Matthews 1991:204). But the way in which the system goes wrong are symptoms of a model that has been wrongly applied. A model that leads the analyst to ask what a strong stem or theme vowel 'means' in isolation, or to ask which characteristic 'realises' partitive case, is simply ill suited to describe this type of inflectional system.

NOTES

* This paper grew out of a presentation at a workshop on redundancy phenomena at the 11th Annual Morphology Meeting in Vienna. I am grateful to the organisers and participants for helpful and constructive comments and suggestions. I would also like to thank Martin Ehala, Reeli Torn, Dieter Wunderlich, anonymous reviewers, and especially Tiit-Rein Viitso for criticisms and corrections that have led to improvements in the present version.

¹ Since, as these examples show, the contrast between long and overlong forms is not consistently represented in the standard Estonian orthography, strong overlong forms are marked by a preceding grave accent, as in Viks (1992).

² See, for example, Booij (2005) and Shibatani and Thompson (1996), and references cited therein.

³ Viks (1992) and Erelt (1999) provide the most widely accepted descriptions of inflection classes in Estonian. Analyses of the declensional system can be found in Erelt et al. (1995), Ehala (1997) and Peebo (1997). Erelt et al. (2000) usefully distinguish open from closed noun classes. Although the literature in English is sparser, Viitso (2003a) gives an accessible overview. Mürk (1997) and Saagpakk (2000) also offer highly detailed classifications, but their descriptions often differ significantly from Estonian sources, and should be verified against these sources.

⁴ The distinction between *t* and *d* marks a **length** contrast in Estonian. Orthographic *t* represents a long voiceless stop /t:/, and *d* a short counterpart /t/. A long *t* is shortened to *d* following a Q3 syllable, as in *teed* and *i`deed*.

⁵ I am indebted to T.-R. Viitso for drawing this stress pattern to my attention.

⁶ From a representational perspective, the penultimate stress pattern in class 3 can be described by associating class 3 nouns lexically with a single trochee at the right edge of the word.

⁷ This classification of syllable types in Estonian entails that 'the representation of weight and segmental length should be separated' (Ehala 2003:58–59), a position which is advocated on independent grounds in Blevins (1995). See also Lehiste (1971) and Lehiste (1997) for discussion of the phonetic correlates of quantity contrasts.

⁸ There is also a class of adjectives in *-Cse* that have nominative singulars in *-Cne*, partitive singulars in *-Cset*, and otherwise inflect like class 3 nouns. Genitive singulars in this subclass class end in a trochee with a strong first syllable, as illustrated by *viimse* ‘final’, *to taalse* ‘total’, *produktiiivse* ‘productive’ and *territoriiaalse* ‘territorial’.

⁹ The theme vowel is also preserved in the nominative singular of some historically complex class 2 nouns, such as *AASTA*, which corresponds to compound *AAST-AAN* ‘time-from-time-to’ (M. Ehala, personal communication).

¹⁰ Class nouns with an initial Q3 foot may also follow the class 3 pattern, with a partitive plural in *-id* (Erelt et al. 2000:252), a development that Ehala (2003:73) attributes to a shift from moraic to syllabic stress assignment.

¹¹ Although there are some exceptions to the prosodic patterns in Table 6, these exceptions tend to fall under other morphological generalisations. For example, class 1 contains some nominals in *-ik* whose partitive singulars appear to have an odd number of syllables and thus deviate from a trochaic structure. Examples include the trisyllabic genitive singular *hapniku*, ‘oxygen’ or the forms *harjumusliku* ‘habitual’ and *raamatuliku* ‘bookish’, which have five syllables. However, these examples pattern with compounds, in which the class of the final word determines the class of the compound. It is the trochaic genitive singular *kogu* ‘collection’ that determines the class of *RAAMATUKOGU* ‘library’ (‘book collection’) not the non-trochaic genitive singular *raamatukogu*. Moreover, even though *raamatukogu* satisfies the trimoraic length requirement on short illative singulars, the short illative *raamatu'kokku* still contains the Q3 form *'kokku*. The noun *RAAMATUKOGUKAART* ‘library card’ in turn inflects following the pattern of class 1 *KAART*.

REFERENCES

- Anderson, S. R. (1992). *Amorphous Morphology*. Cambridge: Cambridge University Press.
- Aronoff, M. (1994). *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press.
- Baayen, R. H., J. M. McQueen, T. Dijkstra, and R. Schreuder. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In R. H. Baayen and R. Schreuder (eds.), *Morphological Structure in Language Processing*. Berlin: Mouton de Gruyter, 355–390.
- Booij, G. (2005). Compounding and derivation: Evidence for Construction Morphology. In W. U. Dressler, D. Kastovsky, O. E. Pfeiffer, and F. Rainer (eds.), *Morphology and its Demarcations*. Amsterdam: John Benjamins, 109–132.
- Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (ed.), *Handbook of Phonological Theory*. Oxford: Blackwell, 206–244.
- Blevins, J. P. (2004). Inflection classes and economy. In L. Gunkel, G. Müller and G. Zinfonun (eds.), *Explorations in Nominal Inflection*. Berlin: Mouton de Gruyter, 41–85.
- de Jong, N. H. (2002). *Morphological families in the mental lexicon*. Doctoral dissertation, University of Nijmegen. (Published as MPI Series in Psycholinguistics 20).
- de Jong, N. H., R. Schreuder, and R. H. Baayen. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes* 15, 329–365. Reprinted as Chapter 2 of de Jong, N. H. (2002), 17–58.
- Ehala, M. (1997.) Eesti morfoloogia olemus. *Keel ja Kirjandus* 6, 448–461.

- Ehala, M. (2003). Estonian quantity: Implications for moraic theory. In D. Nelson and S. Manninen (eds.), *Generative Approaches to Finnic and Saami Linguistics*. Stanford: CSLI, 51–79.
- Erelt, M. (ed.) (2003). *Estonian Language*. Tallinn: Estonian Academy.
- Erelt, M., R. Kasik, H. Metslang, H. Rajandi, K. Ross, H. Saari, K. Tael and S. Vare (1995). *Eesti keele grammatika: morfoloogia*. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut.
- Erelt, M., T. Erelt and K. Ross. (2000). *Eesti keele käsiraamat*. Tallinn: Eesti Keele Sihtasutus.
- Erelt, T. (ed.) (1999). *Eesti keele sõnaraamat*. Tallinn: Eesti Keele Sihtasutus.
- Grünthal, R. (2003). *Finnic Adpositions and Cases in Change*. Helsinki: Société Finno-Ougrienne.
- Hockett, C. F. (1947). Problems of morphemic analysis. *Language* 23, 321–343. Reprinted in Joos (1957), 229–242.
- Hughes, M. and F. Ackerman. (2002). Words and paradigms: Estonian nominal declension. *Proceedings of the 37th annual meeting of the Chicago Linguistics Society*, Chicago.
- Joos, M. (ed.) (1957). *Readings in Linguistics I*. Chicago: University of Chicago Press.
- Kuryłowicz, J. (1949). La nature des procès dits “analogiques”. *Acta Linguistica* 5, 121–138. Reprinted in Hamp, et al. (1966), *Readings in Linguistics II*, University of Chicago Press, 158–174.
- Laakso, J. (2001). The Finnic languages. In Ö. Dahl and M. Koptjevskaja-Tamm (eds.), *Circum-Baltic Languages. Volume I: Past and Present*. Amsterdam: John Benjamins, 179–212.
- Lehiste, I. (1971). *Suprasegmentals*. Cambridge, MA: MIT Press.
- Lehiste, I. (1997). The search for phonetic correlates in Estonian prosody. In I. Lehiste and J. Ross (eds.), *Estonian Prosody: Papers from a symposium*. Tallinn: Institute of Estonian Language, 11–35.
- Matthews, P. H. (1972). *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge: Cambridge University Press.
- Matthews, P. H. (1991). *Morphology*. Cambridge: Cambridge University Press.
- McCarthy, J. J., and A. Prince. (1986). *Prosodic Morphology 1986*. Massachusetts: University of Massachusetts and Rutgers University.
- McCarthy, J. J. and A. S. Prince. (1995). Prosodic morphology. In J. A. Goldsmith (ed.), *The Handbook of Phonological Theory*. Oxford: Blackwell, 318–366.
- Mürk, H. W. (1997). *A Handbook of Estonian: Nouns, Adjectives and Verbs*. Bloomington: Indiana University.
- Peebo, J. 1997. *Eesti keele muutkonnad*. (Tartu Ülikooli eesti keele õppetooli toimetised 5.) Tartu: Tartu Ülikooli Kirjastus.
- Prince, A. (1980). A metrical theory for Estonian quantity. *Linguistic Inquiry* 11, 511–562.
- Saagpakk, P. F. (2000). *Estonian-English dictionary*. Tallinn: Koolibri.
- Schreuder, R. and R. H. Baayen. (1997). How complex simplex words can be. *Journal of Memory and Language* 37, 118–139.
- Selkirk, E. O. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry* 11, 563–605.
- Shibatani, M. and S. A. Thompson (eds.) (1996). *Grammatical Constructions: Their Form and Meaning*. Oxford: Oxford University Press.
- Stump, G. T. (1993). On rules of referral. *Language* 69, 449–479.

- Viitso, T.-R. (2003a). Phonology, morphology and word formation. In Ereht M. (ed.), *Estonian Language*. Tallinn: Estonian Academy Publishers, 9–92.
- Viitso, T.-R. (2003b). Rise and development of the Estonian language. In Ereht M. (ed.), *Estonian Language*. Tallinn: Estonian Academy Publishers, 130–230.
- Viks, Ülle (1992). *Morphological Dictionary of Estonian: Introduction and Grammar*. Tallinn: Institute of Language and Literature.
- Zwicky, Arnold M. (1985). How to describe inflection. In M. Niepokuj et al. (eds.), *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 372–386.

RCEAL
University of Cambridge
English Faculty Building
9 West Road, Cambridge CB3 9DP, UK

jpb39@cam.ac.uk

Do affixes have meaning? Polarity in the Toten dialect of Norwegian meets morphological theory*

HANS-OLAV ENGER

1. INTRODUCTION

1.1. *Two extreme positions*

In many approaches to morphology, an affirmative answer to the question in the title is more or less taken for granted. Within what one might refer to as ‘morpheme-based’ approaches, a central idea is that affixes and stems are both morphemes. However, in the ‘word-based’ school, including such scholars as Anderson (1992), Beard (1995), Stump (2001) and Spencer (2001, 2003), the answer is not equally obvious. This tradition takes the word—the lexeme—as its major unit, and it has been quite influential among morphologists, for good reasons. Within this tradition, much emphasis has been placed on certain differences between affixes and stems, including their semantics.¹ There are different positions within this school, but the most radical position is that affixes basically are semantically empty. Thus, Beard (1995:20) says quite explicitly that affixes “bear no semantic content” and that bound grammatical morphemes “have no semantic or grammatical content” (Beard 1995:69).

Summing up so far, there are two possible extreme positions:

- (A) an affix has independent properties; affixes and stems have meaning, and they are on a par;
- (B) an affix is nothing but the formal trace of an abstract grammatical process; words have meaning, affixes do not; the two are entirely different. Compare Beard (1995:39): “affixes share none of the properties by which we identify the prototypical lexemes”.

Position B has received some support in recent years. Even if rather few scholars presumably would be prepared to go quite as far as Beard, there is still a trend in recent morphological theory to downplay the semantic contribution made by parts of words, and there are some good reasons why (cf. section 1.2). Nevertheless, the bulk of this paper is devoted to defending a less radical intermediate position. Notably, contrary to first appearances, some affixes in the Toten dialect of Norwegian obey the Principle of Contrast and the Exclusive Disjunction Bar (these concepts will be explained in section 1.3); they can thus be said to have meaning and to have SOME properties in common with lexemes.

1.2. *The appeal of hypothesis B*

Position B may, to some extent, be seen as a reaction against position A.² Although the aim of this paper is to question hypothesis B, there are important ideas behind it. Most importantly, affixes are not as orderly and well-behaved examples of Saussurean signs as are words. The issue is given a lengthy treatment elsewhere (e.g. Anderson 1992, chapter 3; Beard 1995, chapters 1 and 3); a few points will be summarised here.

To begin with, if we ask the proverbial man in the street what the meaning of *dog* is; we will probably get an answer; if we ask what the meaning of *-s* is, we probably will not. Affixes usually presuppose a word in order to be interpreted, but not necessarily the other way around, so that one might say that only the meaning of affixes is relational (compare Wurzel 1989, Beard 1995, chapter 3). In some cases, affixes would at least at first seem rather meaningless. This is, for example, the case for the notorious German *Fugenelement* as in *Schafskopf* 'mutton head', where it is far from trivial what meaning to ascribe to the *s* (see, e.g. Fuhrhop 1996). In many cases, it seems more reasonable to ascribe meaning to a word as a whole rather than to its parts. This holds for familiar English examples such as on the one hand *refer*, *receive*, *defer*, *deceive*, on the other *boysenberry*, *huckleberry*, *cranberry* (Anderson 1992:51, 55).

Finally, a semantic point that is often made in support of hypothesis B in the literature is that affixes often are polyfunctional (compare Beard 1995:33–34), and thereby differ from words. For example, English *-s* can mean either '3. singular' or 'plural' or 'possessive', and this seems to differ from the behaviour of a prototypical lexical item.

1.3. *Reconsidering the polyfunctionality issue; plan and purpose*

One may have much sympathy with the word-based tradition, and I emphatically do not wish to advocate position A. Nevertheless, it remains hard to believe that no affixes have any meaning whatsoever, and some of the arguments in favour of position B are not quite as strong as they may seem. This paper is devoted to the issue of polyfunctionality in particular.

Another way of phrasing the polyfunctionality argument (from section 1.2) would be to say that homonymy is much more common in the case of affixes than in the case of words. This point is important, but it does not necessarily indicate a fundamental difference between affixes and words. Given the fact that affixes are much shorter than are words, there is less room for phonological differentiation among affixes, and it is therefore only to be expected that they should display more accidental homonymy. Thus, the observation that homonymy is more common with affixes does not support a fundamental lexical-semantic distinction between words and affixes. In fact, Beard qualifies many of his claims

on the difference between words and affixes with the word ‘prototypical’ (see for example his pp. 15–16 or the quote in section 1.1 above); this seems (to me, at least) to indicate a difference in degree rather than in kind.

Carstairs-McCarthy (1994, 1998, 2001) has argued that affixes and words (lexical items) are subject to at least PARTLY the same semantic constraints: Both are subject to the Principle of Contrast—the idea that ‘every two forms contrast in meaning’ (cf. Clark 1993)—and to the Exclusive Disjunction Bar; neither can exhibit incompatible disjunctive meaning.³ These generalisations are interesting and promising, and they are not compatible with the extreme hypothesis B, i.e. the view that affixes do not have any meaning whatever and nothing in common with lexical units.

The Principle of Contrast is relevant in many different sub-disciplines of linguistics. These include developmental psycholinguistics and lexical semantics (cf. Clark 1993) and inflectional morphology, as argued by Carstairs-McCarthy. Also syntax may be included, at least in a construction-based approach: Different constructions differ in meaning (cf. Goldberg 1995), so that choosing one construction rather than the other is in some ways similar to choosing one word rather than the other.⁴ Since the Principle of Contrast appears to be relevant to so many different sub-disciplines, we should not abandon this generalisation too easily.

The purpose of this paper is not to discuss in full the question whether affixes have meaning; the question of ‘the meaning of “meaning”’ is known to have filled volumes. Nevertheless, two brief remarks may be made. Whatever the outcome of the discussion of the meaning of “meaning” may be, one is unlikely to find a criterion that can distinguish the meaning of ANY affix from that of ANY word; see the discussion of *commit* in section 2.2 and of clitics in section 3.2. Furthermore, if linguistic theory is assumed to have anything to do with psychological plausibility, one wonders if it is realistic, in terms of language acquisition, to assume that an audible element does not have a meaning, except indirectly as a trace of an ‘invisible’ operation, which is what really carries meaning, according to hypothesis B.

The issue whether affixes have meaning is the theoretical background against which we shall examine two examples that at first sight seem to give strong support to extreme position B. These are also cases of polarity. To the best of my knowledge, polarity is not given special treatment within Beard’s (1995) framework, *Lexeme-Morpheme Base Morphology*, but Beard (1995:33–36) uses polyfunctionality as an argument in favour of hypothesis (cf. section 1.2 above) B, and polarity is a rather extreme case of polyfunctionality. If even cases that appear to give particularly strong support to hypothesis B can be given another interpretation, then some of the motivation for adopting it is weakened.

Section 2 presents these cases, which at first seem to be forceful evidence for hypothesis B. However, on closer inspection it turns out that the examples do not really support hypothesis B after all. Rather than defending extreme

position A, however, we take a less radical in-between position: While affixes may not be just like words, many affixes do have some meaning; they are similar to words in this respect. Notably, they obey the Principle of Contrast and the Exclusive Disjunction Bar. This starting point turns out to be heuristically useful in a way hypothesis B simply cannot be. Section 2 also presents some reflections on the difficult relation between affixal and non-affixal inflection in morphological theory.

In section 3, some points are made on polarity, including clitics. If hypothesis B is argued on the grounds that affixes display polarity, then it should be noted that clitics can exhibit the same pattern. This is an embarrassment for hypothesis B, because presumably nobody would want to claim that clitics have no meaning. An apparent similarity between polarity and gender mixture is examined and found to be relatively insignificant, and some Norwegian cases of polarity in clitics turn out to be counter-examples against an earlier generalisation in the typological literature. After a brief review of converging literature in section 4, the paper is summarised in section 5.

Central claims of the paper, then, is that at least some apparently meaningless suffixes do have a meaning, so that hypothesis B is unhelpful, and that the Principle of Contrast and the Exclusive Disjunction Bar hold even for some cases that at first sight look like blatant violations, so that affixes and words do have something in common.

2. TOTEN *-a* AND *-en*: A CASE IN POINT FOR MEANINGLESS AFFIXES?

2.1. *The empirical problem*

The following example is taken from the noun inflection of the Toten dialect in Norwegian (Lie 1990, Faarlund 2000):

Table 1. Noun inflection in the Toten dialect

	Indefinite sg.	Nominative definite sg.	Dative definite sg.
Masculine	<i>båt</i> 'boat'	<i>båt-en</i>	<i>båt-a</i>
Feminine	<i>øks</i> 'axe'	<i>øks-a</i>	<i>øks-en</i>
Neuter	<i>hus</i> 'house'	<i>hus-e</i>	<i>hus-i</i>

The distinction between the nominative and the dative is neutralised in the indefinite, as Table 1 shows.

The polyfunctionality of *-a* and *-en* is striking. In Table 1, the affix in the nominative definite singular for masculines is identical to the affix in the dative definite singular for feminines, and conversely the affix in the dative definite

singular for masculines is identical to the affix in the nominative definite singular for feminines.

The pattern displayed in Table 1, then, is a case of polarity (although, to the best of my knowledge, nobody has explicitly said so in the Norwegian literature before). The formal definition of polarity, according to Serzisko (1982:183), is as follows:

(The concept of polarity) means that when there exist two grammatical categories (*signifiés*) X and Y, and two corresponding exponents (*signifiants*) A and B, then value X can sometimes be assumed by A, while B denotes Y; and sometimes value X is expressed by B, and then it is necessarily A that represents Y. In this manner we obtain two paradigms: (1) XA vs. YB, and (2) YA vs. XB.

The Toten data in Table 1 illustrate well why scholars might want to argue position B: At least at first sight, it seems pointless to argue that the affixes *-en*, *-a* carry any meaning here. There is no obvious semantic criterion according to which masculine nominative and feminine dative constitute a natural class against feminine nominative and masculine dative. Of course, one might argue that *-en* means ‘either masculine nominative definite singular or feminine dative definite singular’, *-a* means ‘either feminine nominative definite singular or masculine dative definite singular’; compare 1:

- (1) *-en* means ‘EITHER masculine nominative def.sg OR feminine dative def.sg.’,
-a means ‘EITHER feminine nominative def.sg. OR masculine dative def. sg.’

However, this analysis belies Carstairs-McCarthy’s (1998) claim that affixes (like words) do not have genuinely exclusive disjunctive meaning, i.e., his ‘Exclusive Disjunction Bar’; this conclusion is unwelcome (cf. section 1.3). Thus, the Toten example presents an empirical challenge.

Other analyses than that given in (1) are possible, however. Homonymy is much more common in the case of affixes than in the case of words, and it seems to be the easy way out here. After all, *-a* and *-en* are used as affixes also for other word classes than nouns; also in cases where (presumably) few linguists would claim a semantic link. For example, in the Toten dialect *-a* is also the infinitive suffix for a set of verbs, and *-en* is also a derivational suffix for a number of adjectives.

Thus, one possible analysis of the facts in Table 1 might be to posit two distinct, but homophonous *-a* suffixes and two distinct, but homophonous *-en* suffixes. This may seem the easy way out of our problem, probably too easy. Let us examine the criteria for polysemy vs. homonymy. Speaker intuition on semantic relatedness is a common criterion, but it is not easily applicable in the present case. In typological studies, however, it has become common to look at cross-linguistic evidence (see, e.g. Haiman 1985): If parallel patterns are found,

then polysemy is preferred; if not, then homonymy is. Now, the masculine and the feminine arguably constitute a “natural grammatical class” against the neuter in much Norwegian (see, e.g. Torp 1997). From the semantic point of view, the masculine and the feminine are the animate genders—and that is the ‘core semantics’ of gender—against the less animate neuter. As for their form, the masculine and the feminine share personal pronouns in *h-* which the neuter does not have, they share adjective forms against the neuter, they are intimately related in the formation of the plural, as compared to the neuter, which stands out, compare Nynorsk *gutur* ‘boys’—*jenter* ‘girls’—*hus* ‘houses’. They share so-called definite article against the neuter (*den* vs. *det*). Furthermore, masculines and feminines often display patterns of ‘swapping’ in formal marking similar to that in Table 1 in many Norwegian dialects. Typically, while *-ar* is the common plural suffix for masculines and *-er* for feminines, minority groups display an ‘opposite pattern’; a minority group of masculines have *-er*, while a minority group of feminines have *-ar* (see also Enger 2004a). We shall also see in section 2.3 that one may see partial parallels to the Toten pattern in Somali and in the Tromsø dialect of Norwegian. This speaks in favour of a polysemy analysis. There is also a possible methodological argument: In general, homonymy analyses should be our last resort, and it is preferable to maximise polysemy (cf. Lyons 1977:554f).⁵ We may therefore conclude with polysemy.

In order to formulate the polysemy analysis, we first set aside the neuter, the inanimate gender. It stands out both formally and semantically, as shown above. A polysemy analysis is then possible along the following lines: The masculine is often seen as the unmarked gender in Norwegian (e.g. Trosterud 2001:34–35), and the nominative is clearly unmarked in comparison with the dative.⁶ Thus, the suffix *-en* in Table 1 signals that the word-form either is ‘unmarked animate gender, unmarked case’ (i.e. masculine, nominative) or ‘marked animate gender, marked case’ (i.e. feminine, dative). Conversely, the suffix *-a* signals that the word-form either is ‘marked animate gender, unmarked case’ (feminine, nominative) or ‘unmarked animate gender, marked case’ (masculine, dative). This opens for an analysis in which there is no genuinely exclusive disjunctive meaning:

- (2) *-en* means ‘agreeing in markedness values for animate gender and case’
-a means ‘non-agreeing in markedness values for animate gender and case’

Under the analysis in (2), we can maintain the claim that affixes (like words) do not have genuinely exclusive disjunctive meaning, and even the apparently ‘empty’ affixes *-en* and *-a* have meaning. Thus, even what looked as a good example in favour of the idea that affixes are semantically meaningless (hypothesis B), or that they violate the Exclusive Disjunction Bar, turns out on closer inspection not to be such an example, after all.

There is no denying that an analysis of the Toten examples of *-a*, *-en* that is compatible with hypothesis B is feasible. Yet such an analysis would HAVE TO miss

the point that the formal similarity is semantically motivated, and so it would not, in my view, represent any gain.

Admittedly, ‘agreement in markedness values’ is not a kind of meaning that we would normally ascribe to a lexical unit. Thus, while one may question the idea that affixes do not have meaning and argue that they share some semantic constraints with words, it seems impossible to sustain the claim that affixes are exactly like words in every way, semantically (i.e. hypothesis A). But my claim is a more modest one, namely that affixes and words are not quite so radically different as hypothesis B leads us to expect (cf. also Booij 1996).

More radically, some might wish to object that ‘agreement in markedness values’ is not really a kind of meaning at all. This objection raises the question what meaning really is, and we cannot go into that issue in all its complexity (cf. section 1.3), but a couple of points may be made. Firstly, previous research indicates that affixes can have intra-morphological meaning (see Cameron-Faulkner and Carstairs-McCarthy 2000, Carstairs-McCarthy 1994, 1998, 2001, Enger 2004a, Maiden 2001); i.e. they can signal morphological information such as inflectional class. Wurzel (1984b) even suggested that the German *umlaut* means the marked category. In this perspective, the meaning ascribed to *-en*, *-a* in (2) is not particularly surprising.

Secondly, the markedness values for *Toten* above are uncontroversially compatible with the semantics of the respective genders and cases, and this renders the idea that agreement in markedness values has nothing to do with semantics unappealing.⁷

2.2. The issue of compositionality

A theoretical objection against the idea that affixes have meaning, and hence against the concrete analysis in (2), is this: Affixes in themselves do not mean anything; it is in the syntagmatic context, either within the word, i.e. together with a stem, or within the sentence, i.e. the syntactic context, that there is meaning. (Compare Beard 1995:20: “grammatical morphemes are interpretable only contextually”.) While this objection is understandable, it is not quite convincing. The idea that affixes do not have any meaning because only words and larger syntactic phrases do, follows only if we assume that there can be no redundancy, but this assumption is dubious for languages in general, and for morphology in particular, where it is well known that languages often make use of a ‘belt-and-braces-strategy’ (Cahill and Gazdar 1999, Carstairs-McCarthy 2001).

To put it crudely, the fact that meaning may be found in syntax does not mean that there cannot be meaning in morphology. There are many studies indicating that morphology has some autonomy (e.g. Aronoff 1994, Carstairs-McCarthy 2001, in press, Cameron-Faulkner and Carstairs-McCarthy 2000, Maiden 2001, 2005). In fact, there are even cases where the gender of a word—its syntactic

agreement properties—is changed because of its declensional class—its morphological properties (Enger 2004a:57–59). This indicates that syntax does not always dominate morphology; the reverse can happen.

Furthermore, the argument that affixes should not be ascribed any meaning because the meaning depends crucially on the syntagmatic context can in fact be turned on its head: There are many open-class vocabulary items whose meaning ALSO depends crucially on their syntagmatic context.⁸ Consider the following examples, involving the lexical item *commit*:

- (3) The magistrate committed Rocco for murder
 Rocco had committed many serious crimes
 The Greeks had committed themselves to having the stadium ready on time
 Fred is eager to marry, but Sue is reluctant to commit

For most purposes, whether we are dealing with one polysemous verb *commit* or several homonymous ones is not the most important question. The relevant point is that the verb *commit* depends on the syntagmatic context for its meaning, much like the suffixes *-a*, *-en* in the Toten dialect. The difference between words and affixes on this point appears to be rather one of degree than one of kind (compare section 1.3 above).

The question whether affixes have meaning can be seen as a version of the longstanding question of semantic compositionality. Within some morpheme-based approaches, it is assumed that the meaning of a word is simply the sum of the meanings of its parts.⁹ In response to this, scholars have emphasised that the meaning of the word is more than the sum of the meanings of the parts, as is particularly clear in compounds (e.g. Becker 1990:34–35). While this argument is correct, it does not follow that the parts are meaningless.

In short: Affixes do have some meaning. Even Toten *-a* and *-en* obey the Principle of Contrast and the Exclusive Disjunction Bar. This does not mean that affixes necessarily are on a par with words in every respect. In particular, affixes may have meanings of a more general, more abstract kind than words, and “agreement in markedness values” is a case in point. This should not be particularly surprising, however, because affixes are grammatical units, and the meanings of grammatical units are typically more abstract than the meanings of open-class words.¹⁰

It could even be argued that unless we look for meanings of affixes, we are not going to find them. In other words, the idea that affixes do have meaning is heuristically fruitful in a way that hypothesis B just cannot be. If we had assumed from the outset that affixes basically were meaningless, then the analysis in 3 would have been impossible. In yet other words, there is no denying that one might come up with an analysis of *-a*, *-en* that would be compatible with hypothesis B, but this move would not represent any real gain; rather, generalisations would be lost, in that one would have to claim that the cross-dialectal similarities had no semantic reason.

2.3. Polarity in Somali and Tromsø

Polarity is well known from the literature (e.g. Serzisko 1982, Corbett 1991, 2000); we shall only consider two cases here, from Somali and from the Tromsø dialect of Norwegian.

- (4) The definite article in Somali (a Cushitic language) (Corbett 1991:196, Serzisko 1982:185):

	singular	plural
masculine	<i>kii</i>	<i>tii</i>
feminine	<i>tii</i>	<i>kii</i>

First of all, note that the approach to polarity in Toten suggested in (2) carries over straightforwardly to the example from Somali in (4). Also in Somali, the polarity can be accounted for in terms of agreement in markedness. This generalisation is simply lost if one assumes the extreme hypothesis B—that affixes have no meaning. So again, the assumption that affixes do have some meaning is heuristically fruitful in a way that hypothesis B cannot be, as argued in section 2.1. The link between polarity and ‘markedness agreement’ is also observed by Serzisko (1982:195–196), so analyses of the sort in (2) are familiar within typology. It seems promising that the idea can be generalised in this way.¹¹

In Table 1, polarity involves case, gender and definiteness. Cases of polarity that involve number and gender (like (4)) appear to be more common, both in Norwegian and at large. Thus, in Tromsø (Northern Norwegian) we find the pattern in (5). The case is, however, unlike the Somali example in (4) (but like that of Toten) in involving an affix rather than a word:

- (5) Polarity in the Tromsø dialect
- | | | |
|-----------|---------------------|-------------------|
| | Indefinite singular | Indefinite plural |
| masculine | <i>bakke</i> ‘hill’ | <i>bakka</i> |
| feminine | <i>jenta</i> ‘girl’ | <i>jente</i> |

The indefinite masculine singular has the same suffix as the indefinite feminine plural (cf. Table 1, and compare also the Somali example in (4)). In the same vein as in (2), one might analyse this as a case where one suffix (*-e*) indicates agreeing markedness values, the other (*-a*) indicates disagreeing markedness values.

2.4. Polarity, definiteness and case

Run-of-the-mill examples of polarity such as the Somali example in (4) involve gender and number, presumably because the category with which number has the closest relations is gender (Corbett 2000:272, Serzisko 1982:179; see also Janda 2000). The polarity in Toten, however, has to do with gender, case and

definiteness, and their connection is not equally close. It is not so common to find polarity where gender and case interact, in the way they do in Table 1, but at least, an Old French example is known (see Serzisko 1982:197f, or see Serzisko 1982:197–8 for further references).

We know that case interacts with definiteness in Norwegian dialects. As Table 1 illustrates, the case distinction dative/nominative is only found in the definite, never in the indefinite. Another indication that there is a link between definiteness and case is that Scandinavian definiteness distinctions often translate into Finnish case distinctions. It has been argued that gender has to do with the picking out of referents (e.g. by Lehmann 1982 and Claudi 1985), and it seems obvious that definiteness also has to do with the picking out of referents. In other words, the particularly close semantic relations between gender and number may explain why these two categories are particularly often involved in polarity, but there are also semantic relations between gender and definiteness, having to do with individuation, so it should come as no surprise that these categories also can be involved in polarity.

One may wonder why polarity so often involves gender. According to Serzisko (1982:196), the explanation is that gender is ‘pre-destined for secondary uses’ (“für eine uneigentliche Verwendung prädestiniert”). The premise is that since gender is a lexical property of the noun, it cannot change between singular and plural, anyway, so it can be used for other purposes. However, most linguists today argue that gender actually can change between numbers (Carstairs-McCarthy 1994, Corbett 1991). If we accept their view, the basic premise in Serzisko’s explanation does not hold. I am not aware of any alternative suggestions why polarity so often involves gender, so the question is open for future research.

2.5. *The plural—and the relation between affixal and non-affixal inflection*

What makes the pattern in Table 1 so striking is of course that it is a sort of mirror-image. If we go on to the plural, we find an even more striking challenge for the assumption that affixes have meaning. Compare the paradigm in Table 2:

Table 2. Noun inflection in Toten (Faarlund 2000)

Nominative indefinite sg.	Nominative definite sg.	Dative definite sg.	Nominative indefinite pl.	Nominative definite pl.
<i>bât</i> ‘boat’ ¹ bo:t/	<i>bât-en</i> ¹ bo:ten/	<i>bât-a</i> ¹ bo:ta/	<i>bât-er</i> ² bo:ter/	<i>bât-a</i> ² bo:ta/
<i>øks</i> ‘axe’ ¹ øks/	<i>øks-a</i> ¹ øksa/	<i>øks-en</i> ¹ øksen/	<i>øks-er</i> ¹ økset/	<i>øks-en</i> ² øksen/

Here, the suffixes in the definite nominative plural are ‘the opposite’ of what they are in the dative singular. Thus, there has been an additional ‘swapping’, and

even readers who were sympathetic with the analysis in (2) may think that the paradigm in Table 2 is the last straw, and that it is time to throw in the towel by now. However, more interesting options are available. While the affixes are identical in the nominative definite singular of masculines (*båt-en*), dative definite singular of feminines (*øks-en*) and the nominative definite plural of feminines (*øks-en*), the ‘toneme’ (word tone) is not. The singular forms have toneme 1, the plural toneme 2. In other words, the plural forms stand out non-affixally.

This means that we are not forced to abandon an analysis similar to the one in (2) above, but we do have to take toneme into account. This can be done straightforwardly: Toneme 2 on a noun that has toneme 1 in the indefinite singular (the ‘reference form’ in Norwegian, cf. Enger 2004a:64–67) signals that it is in the plural. This may be compared to the role played by the umlaut in German. According to Carstairs-McCarthy (2001), the German umlaut can “be regarded as a clearcut signans for plural Number”.

This move may perhaps look like keeping the letter of the law while breaking its spirit. After all, Cameron-Faulkner and Carstairs-McCarthy (2000) stress ‘the wisdom of a version of morphological theory that differentiates firmly between affixal and non-affixal inflection’. If we really have to drag in non-affixal inflection to save the semantic analysis, one might object that this amounts to accepting the opposite claim, i.e. that there are no theoretically significant differences between affixal and non-affixal inflection (Stump 2001:9). In other words, does not the fact that we have to know the toneme in order to give the right interpretation of *båten* show that it is the COMBINATION OF affixal and non-affixal means that produces meaning—and does not this show that we had better not differentiate too firmly between the two kinds of inflection, after all?

Cameron-Faulkner and Carstairs-McCarthy (2000) ask a related question in their analysis of Polish: “Does the fact that the NBP [No Blur Principle, a corollary of the Principle of Contrast, HOE] ignores stem alternation mean that the distribution of stem alternants is entirely unconstrained, so far as morphological theory is concerned?” There are two reasons why the answer is ‘not necessarily’. For one thing, there may be other restrictions on stem alternation.¹² More relevantly for present purposes, while the No Blur Principle constrains only affixal inflection, this does not mean that stem alternation is entirely irrelevant. As Cameron-Faulkner and Carstairs-McCarthy (2000) point out, stem alternation may be relevant in the sense that there may be implicational links between stem choice and affix choice, and indeed, we find something similar in the Toten examples: The analysis in (2) holds only for word forms having toneme 1. Toneme 2 in combination with the affix *-en* signals definite plural of feminines, toneme 2 in combination with *-a* signals definite plural of masculines.¹³ Thus, it is the combination of affixal and non-affixal means that is relevant for us.

When claiming that there are no theoretically significant differences between affixal and non-affixal inflection, Stump (2001) does not discuss the Paradigm Economy Principle (Carstairs 1987) or the No Blur Principle

(Carstairs-McCarthy 1994)—which is a corollary of the Principle of Contrast. The way in which inflectional sameness and difference needs to be assessed for those purposes presupposes that affixal and non-affixal inflection be kept apart. If we assume that these principles are valid, it follows that for *SOME* purposes, affixal and non-affixal inflection should be kept apart. It does not follow that they should be kept apart for *EVERY* purpose, however. Perhaps morphologists have been too categorical on this point. For example, Wurzel (1989) assumes that there are two options, either what he calls a ‘process morphology’ (inspired by hypothesis B) or what he calls an ‘affix morphology’ (inspired by hypothesis A). There is an alternative view, however, viz. “that some aspects of inflection fit an inferential approach [e.g. Stump’s 2001 model] while others are best described if elements such as affixes are treated as having at least some of the characteristics of lexical items” (Carstairs-McCarthy, in press). This alternative view is taken here.

3. POLARITY, GENDER MIXTURE AND CLITICS

Section 2 has shown how apparently meaningless suffixes—in cases of polarity—can be analysed so as to have semantic content. Hopefully, the section will also have shown that dialectal material is relevant to linguistic typology and vice versa. In principle, nobody would disagree, of course; in practice, the scientific exchange between the two disciplines has been rather meagre, though there are signs of change, cf. Kortmann (2004).

In section 3.1, we turn to a putative difference between polarity and gender mixture, which some Norwegian dialects at first sight seem to prove wrong. In section 3.2, we look closer at polarity in clitics. The main claim there is that if polarity in suffixes is an argument that suffixes are meaningless, then, by the same logic, polarity in clitics, as exemplified in the Toten dialect, indicates that at least some words are meaningless, and this undermines the extremely sharp distinction between affixes and words drawn in hypothesis B.

3.1. *Polarity and gender mixture*

Can there be a parallelism between polarity and so-called gender mixture—e.g. nouns that are feminine in the singular, neuter in the plural—with respect to which categories they can be based on? At first sight, one would not expect so; according to Carstairs-McCarthy (1994:771), “the only morphosyntactic basis for gender mixture is number”. In other words, a case in which a noun has gender X in property A, gender Y in property B, is possible if and only if the category to which properties A and B belong is the category of number. It would then seem that polarity can involve categories other than number, gender mixture cannot. While this undoubtedly holds by and large, there are at least two Western

Norwegian dialects (Bjerkreim in Rogaland, Bud in Møre & Romsdal) and one Northern Norwegian one (Tromsø) in which a few nouns can change gender between the indefinite and the definite (Levang 2003, Helge Sandøy p.c., Curt Rice p.c.). Examples from Levang (2003):

- (6) Gender mixture in Bjerkreim¹⁴
ei avis—*avisen*
 a(fem) paper—paper(def, masc)
 ‘a paper—the paper’
ei gong—*gongen*
 a(fem) time—time(def, masc)
 ‘a time—the time’
ei hevd—*hevden*
 a(fem) shit—shit(def, masc)
 ‘a shit—the shit’

These examples show that gender actually can shift between the indefinite and the definite singular. However, these examples are truly marginal. Example (6) shows the only three lexemes of this kind in the Bjerkreim dialect, and they are diachronically unstable, according to Levang (2003); in Bud, only one lexeme is of this kind. So there are exceptions to Carstairs-McCarthy’s generalisation, but these exceptions are lexical rather than grammatical in nature, and no real cause for worry. Even if the difference is not 100%, it comes close: Gender mixture is different from polarity with respect to its basis.

3.2. Polarity and clitics

So far, we have considered polarity for suffixes. However, the Toten dialect evinces a similar phenomenon in the case of proclitic personal pronoun. Compare (7), where clitics are boldfaced and the relevant affixes are underlined.

- (7) Examples of clitic pronouns and affixes in the Toten dialect (Faarlund 2000)
- (7a) *Der ær **n** far.*
 there is he(nom) father
 ‘There is [my] father’
- (7b) *Dær ær **a** mor.*
 there is she(nom) mother
 ‘There is [my] mother’
- (7c) *Du kænntala ått **a** far, så skar je tala ått **n** mor.*
 you(nom) can speak to he(dat) father, so shall I(nom) speak to
 she(dat) mother
 ‘You’d better speak to Dad, and I’ll speak to Mum.’

- (7d) *Du k^ænn tala åt såna, så skar je tala åt kjæringen*
 you can speak to son (dat, masc, def), so shall I(nom) speak to
 wife(dat, fem, def)
 ‘You’d better speak to the son, and I’ll speak to the wife’
- (7e) *Der ær sånnen og kjæringa*
 there is son(nom, masc, def) and wife(nom, fem, def)
 ‘There is the son and the wife’

In other words, the dative form of the proclitic personal pronoun in the 3. singular masculine (7c) looks like the nominative form of the proclitic personal pronouns in the 3. singular feminine (7b). Conversely, the dative form of the proclitic personal pronoun in the 3. singular feminine (7c) looks like the nominative form of the proclitic personal pronoun in the 3. singular masculine (7a). This is polarity. For the reader’s convenience, the similar patterns in affixes are repeated in (7d and e). Presumably, it is not arbitrary that suffixes and clitics should behave similarly; the two are closely related.

Now, it is customary to define genders as “classes of nouns, reflected in the behavior of associated words” (Corbett 1991, following Charles Hockett). According to this agreement-based definition of gender, the definiteness suffix in Scandinavian does not count as a gender marker, since it occurs on the word itself. It would be rash to dismiss the agreement-based definition entirely, but a number of authors have voiced some concern as to whether agreement really should be given such a decisive role in the definition of gender and whether markers on ‘associated words’ really should count for everything, for the purposes of gender identification, and markers on the word itself for practically nothing (see, e.g. Carstairs-McCarthy 1994, Dahl 2000 and Unterbeck 2000). In many other (dative-less) Norwegian dialects, the suffix in the definite singular indicates gender unequivocally, in the sense that it correlates completely with the choice of form on ‘associated words’, and it would therefore seem rather pointless to deny this suffix status as an indicator of gender (cf. Enger 2004a:65–66). According to a strictly agreement-based definition of gender, one would have to argue that the clitic *a* in the noun phrase *a far* in (7c) is a gender marker, while the suffix *-a* in the corresponding inflected noun *fara* in (7d) is not. This seems to be an undesirable consequence, since the two correlate.

The examples in (7) illustrate that clitic pronouns may display a behaviour that is rather similar to affixes, so that one may think of clitics as well in terms of polarity. Now, Serzisko (1982:196) claims that complete polarity is only found in cases of NP-internal agreement (“daß vollständige Polarität nur in der internen Kongruenz vorkommt”). This may seem reasonable a priori, but it represents a problem if we accept Corbett’s (2003:116) conclusion that there is little empirical evidence for drawing any particular boundary between NP-internal and NP-external agreement. In fact, the correctness of Corbett’s position seems to be borne out also by Norwegian dialect data. Consider the examples in (8), taken

from the Toten dialect (Faarlund 2000 and p.c.):

(8) Clitic pronouns in the Toten dialect (continued)

- (8a) *Je sennte a eitt brev.*
I sent he(dat) a letter.
'I sent him a letter'
- (8b) *Så sennte n eitt brev.*
then sent he(nom) a letter
'Then he sent a letter'
- (8c) *Je sennte n eitt brev.*
I sent she(dat) a letter.
'I sent her a letter'
- (8d) *Så sennte a eitt brev.*
Then sent she(nom) a letter.
'Then she sent a letter'

Unlike the examples in (7a–c), the clitic pronouns *a*, *n* in (8a–d) cannot be considered examples of NP-internal agreement. Furthermore, they conform to the definition of polarity given in section 2.1 above. Thus, Serzisko's hypothesis that complete polarity is restricted to NP-internal agreement does not hold against examples such as (8). This shows again the value of confronting typology and dialectology.

Anyway, (7) and (8) show that we find polarity in the case of clitics, just as we do with affixes. It seems less likely that full words should display similar behaviour.¹⁵ This is presumably not arbitrary: Clitics have grammatical meaning, like affixes. Presumably, it is only elements that have grammatical rather than lexical meaning that can display polarity, for it is only with elements having grammatical meanings that we are dealing with closed classes.

At least since Sapir (1921, chapter 5), a well-known view is that affixal meaning is typically more abstract than root meaning. If the argument above holds, we have found some (additional) evidence for the value of the distinction between lexical elements on the one hand and clitics/affixes on the other.

At the same time, we have found a problem for the extreme hypothesis B: If polarity is an argument that affixes are meaningless, the way hypothesis B leads us to believe, then the Toten examples must constitute an argument that clitics also are meaningless. This consequence is surely undesirable, however. Furthermore, since clitics are usually understood as syntactic words, at least SOME words are much like affixes in terms of meaning, after all.

4. CONVERGING STUDIES

The purpose of this paper is not to advocate the extreme position A. The 'word-based' tradition has produced too many forceful arguments against this view, and

they need not be rehearsed here (see, e.g. Anderson 1992, Spencer 2001, 2003, Stump 2001). Nevertheless, there are reasons why one might want to argue that at least some affixes do have some kind of meaning. Some such reasons have been put forward by a number of authors, so the present study is not the only one questioning hypothesis B.

Maiden (2001) presents diachronic cases of homonymy avoidance for derivational affixes in Romance. While such cases probably are infrequent, they can be found, and they make good sense only if affixes are assumed to have some kind of meaning.

In a study of derivational affixes in English, Plag (1999:237) points out that an account within Lexeme-Morpheme Base Morphology is unable to account for the differences in meaning that he finds between different affixes. Plag (1999:240) also finds that putatively synonymous affixes on closer inspection turn out to have subtle differences in meaning; he argues that this parallels the situation of lexemes, which in general can be nearly but not totally synonymous.

Another study on English derivational affixes is presented by Lehrer (2000:152), who concludes that there is “a cline rather than a clear-cut division” between open-class lexemes and affixes. Nevertheless, Lehrer argues that affixes display antonymy and polysemy, just like lexemes do. Lehrer also points out that Beard has used as his primary examples “morphemes with relatively little meaning”.

The findings of the present study should (at least to some extent) be compatible with these studies. Plag’s study presents additional support to the point made in section 2: the assumption that affixes are meaningless is heuristically unhelpful; it is productive to assume that they do have meaning. Like Lehrer, I have tried to show that the distinction between words and affixes is not clear-cut (at least not with respect to the Principle of Contrast and the Exclusive Disjunction Bar); notably, clitics complicate the picture for hypothesis B. Like Maiden, I have tried to show how Carstairs-McCarthy’s idea of intra-morphological meaning can be relevant to morphologists, while also drawing on Serzisko (1982).

The present study is not a mere repetition, however. It has presented evidence from a different language, and concentrated on inflectional affixes, which at first glance may seem even more “meaningless” than derivational affixes (on which Maiden, Plag and all Lehrer focus).

5. CONCLUSIONS

The main points of the paper can now be summarised:

- The claim that affixes, like words, do not display genuinely exclusive disjunctive meaning can be upheld even in the face of some apparent counterexamples. The affixes examined in this study obey the Principle of Contrast and the Exclusive Disjunction Bar.

- Cases of polarity do not necessarily indicate that affixes are meaningless.
- If we do not look for meaning in affixes, we run the risk of losing valuable generalisations.
- While polarity is usually exemplified by affixes, polarity in clitics is also found. This represents a problem for extreme hypothesis B.
- Serzisko's (1982) explanation why gender so often is involved in polarity does not quite hold.
- Complete polarity can be found also in NP-external agreement.
- Gender mixture can be found with other categories than number, but the examples are lexical and diachronically unstable.

The claim that affixes, like words, are subject to a prohibition against genuinely exclusive disjunctive meaning appears to hold, even when confronted with examples of polarity such as those from Toten. In this respect, affixes are like words; they have meaning. It does not follow that affixes are like words in every semantic or every other respect, however. In short, if we return to hypotheses A and B from the outset, there is room for many in-between positions (cf. also Carstairs-McCarthy, *in press*), and this point should be acknowledged more often in the literature. In this perspective, it is understandable that Maiden (2001:45) criticises Beard (1995) for “dubious procrusteanism”.

One may wonder if the value of theoretical consistency sometimes has been overstated in linguistics. For example, against the claims made in this paper, it has been objected that, if one assumes that the distribution of the affixes *-a* and *-en* is influenced by economy, there is no need to invoke semantics as well. In my view, such an objection is misplaced. Within Optimality Theory, for example, one would not argue that one constraint makes all other superfluous. In other words, perhaps this objection is a case of what Langacker (1987:28–30) dubs ‘the exclusionary fallacy’; the idea that one explanation necessarily precludes another. Finally, I would argue that, if we really have to choose, empirical generalisations—such as the one that affixes like lexemes do not have exclusively disjunctive meaning—are more valuable than conceptual economy in the description.

In short, we should try to find a reasonable position between the two extreme views that (A) affixes are just as lexical units, (B) affixes and lexical units are worlds apart. The right question to ask is perhaps not ‘are affixes like words?’—although the initial hypotheses A and B are both answers to this question. Perhaps it is better to phrase the question as follows: ‘In what ways are affixes (un)like words? What properties are common, what properties are different?’ The idea that affixes and lexemes can be subject to SOME but not ALL the same constraints opens for this question.

NOTES

* This paper is dedicated to Andrew Carstairs-McCarthy on the occasion of his 60th birthday. Thanks to audiences in Oslo, Tromsø, Stavanger and Vienna, and to T. Trosterud, who first

brought the problem to my attention. For comments on previous versions, particular thanks to A. Carstairs-McCarthy, J.T. Faarlund, T. Kinn, T. Nessel, A. Spencer, A. Torp, and to three very constructive referees.

¹ An anonymous referee points out that the difference is downplayed in other respects. Thus, Stump (2001:9) explicitly denies any important difference between concatenative and non-concatenative inflection. (Compare also Wurzel 1989.)

² This is quite clear in the way Beard introduces his framework.

³ The observation that lexical units are not likely to have genuinely exclusive disjunctive meaning seems to have a long and respectable tradition, cf. Bruner et al. (1956, chapter 6).

⁴ In fact, Goldberg (1995:4, 23) explicitly draws a parallel between constructions and morphemes; Spencer (1999:67) points out that the parallel is unfortunate, since the classical morpheme concept is so controversial. Nevertheless, the basic idea—a parallel between lexical units and constructions—seems valid (and to be sure, Spencer does not dispute it), and it opens for the Principle of Contrast to have a role also in syntax. The idea that “in syntax, as in morphology, doublets that are semantically and functionally non-distinct are disallowed” has been suggested in variationist approaches to syntactic change too, according to Pintzuk (2003:525).

⁵ An additional, if perhaps weaker argument, is that we do not find polarity of the kind in Table 1 in the most ‘conservative’ Norwegian dialects (Setesdalen, South Sunnmøre). This seems to indicate that the pattern is not due to regular phonological development from Old Norse. This in turn indicates that it is due to the workings of morphology, and not arbitrary.

⁶ Apart from nominative and dative, there are no other cases on the nouns in this dialect. A truly vexed question is what to understand by the word ‘marked’. Firstly, note that the word is here used in the sense of marked with respect to lexical gender assignment rules, and not with respect to agreement. In other words, “marked” is here used in the sense Fraser and Corbett (2000) refer to as ‘normal case default’. Secondly, some linguists understand markedness as an independent structural property (e.g. Wurzel 1984a); others take it to be an epiphenomenal by-product of other factors, as frequency (e.g. Haspelmath, in press). Be that as it may; the markedness ranking above remains the same.

⁷ Another possible objection against the analysis in (2) is that affixes have a limited phonological inventory anyhow, so that there is no need to find a semantic rationale behind their distribution. This objection seems unconvincing to me, however, simply because the category of affixes is delimited on morphological, and not on phonological, grounds in the first place. Thus, the fact that affixes have a limited phonological inventory is not an argument indicating that phonology ‘governs’ morphology or that the patterns are morphologically uninteresting; if anything, it is rather an argument to the opposite effect.

⁸ Thanks to an anonymous referee for this point.

⁹ In Langacker’s (1987) terms, this is the ‘building-block metaphor’.

¹⁰ In a way, the examples discussed from Norwegian are somewhat similar to Steins’ (2000) “flip-flop” examples from Kiowa. However, while Steins aims at defending the ‘morpheme’, my aim is a more modest one, in that I only try to argue that at least some affixes have some meaning. In other words, much like Plag’s (1999:234) analysis, the present one is inspired more by the notion of ‘sign’ than by that of ‘morpheme’.

¹¹ The basic idea has to do with markedness reversal (Serzisko 1982:196 and references), which means that it does not stand in isolation.

¹² For example, there may be a Stem Alternation Constraint (Enger 2004b).

¹³ Similarly, it is well known in the study of Norwegian that a toneme can have an ‘intra-morphological’ meaning; for example, it is well known from the verb inflection that toneme 1 in the present tense can signal that a verb probably is strong (e.g. Enger 1998:106–111). Again, there is a link between stem choice and affix choice.

¹⁴ In the examples, *-en* may be taken as a marker of the masculine gender, cf. Dahl (2000), Enger (2004a). Be that as it may, if they are in the definite, the nouns in question select a masculine form of ‘associated words’, so the point is solid enough, anyhow.

¹⁵ An anonymous referee objects that at least one author would disagree; Lutzeier (1997) takes antonymy to be a special case of polysemy. What that means, however, is that on Lutzeier’s view, also this putative difference between full words and affixes disappears; thus, this objection cannot be said to support hypothesis B in any way that is immediately relevant to present concerns. (Besides, interesting though Lutzeier’s view may be, it seems to be a minority view.)

REFERENCES

- Anderson, Stephen R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Aronoff, Mark (1994). *Morphology by Itself*. Cambridge, MA: MIT Press.
- Beard, Robert (1995). *Lexeme-Morpheme-Base Morphology*. Albany, NY: SUNY Press.
- Becker, Thomas (1990). *Analogie und morphologische Theorie*. Munich: Fink.
- Booij, Geert (1996). [Review of] Beard, R. 1995: Lexeme-morpheme-base morphology. *Language* 72, 812–817.
- Bruner, Jerome S., Jacqueline Goodnow and George A. Austin 1956. *A Study of Thinking*. New York: Wiley.
- Cahill, Lynne and Gerald Gazdar (1999). German noun inflection. *Journal of Linguistics* 35, 1–42.
- Cameron-Faulkner, Thea and Andrew Carstairs-McCarthy (2000). Stem alternants as morphological signata: Evidence from Blur avoidance in Polish. *Natural Language and Linguistic Theory* 18, 813–835.
- Carstairs, Andrew (1987). *Allomorphy in Inflexion*. London: Croom Helm.
- Carstairs-McCarthy, Andrew (1994). Inflection classes, gender and the principle of contrast. *Language* 70, 737–788.
- Carstairs-McCarthy, Andrew (1998). How lexical semantics constrains inflectional allomorphy. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1997*. Dordrecht: Kluwer, 1–24.
- Carstairs-McCarthy, Andrew (2001). Umlaut as signans and signatum: Synchronic and diachronic aspects. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1999*. Dordrecht: Kluwer, 1–23.
- Carstairs-McCarthy, Andrew (in press). Affixes, stems and allomorphic conditioning.
- Clark, Eve E. (1993). *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.
- Claudi, Ulrike (1985). *Zur Entstehung von Genusssystemen*. Hamburg: Buske.
- Corbett, Greville G. (1991). *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville G. (2000). *Number*. Cambridge: Cambridge University Press.
- Corbett, Greville G. (2003). Agreement: terms and boundaries. In William Griffin (ed.), *The role of agreement in natural language: Proceedings of the 2001 Texas Linguistic Society Conference*. Retrieved from <http://uts.cc.utexas.edu/~tls/2001tls/Corbett.pdf>

- Dahl, Östen (2000). Elementary gender distinctions. In Barbara Unterbeck et al. (eds.), *Gender in Grammar and Cognition*. Berlin: Mouton de Gruyter, 577–593.
- Enger, Hans-Olav (1998). *The Classification of Strong Verbs in Norwegian with Special Reference to the Oslo Dialect* (Acta Humaniora 26). Oslo: Scandinavian University Press.
- Enger, Hans-Olav (2004a). On the relation between gender and declension: A diachronic perspective from Norwegian. *Studies in Language* 28, 51–82.
- Enger, Hans-Olav (2004b). A possible constraint on non-affixal inflection. *Lingua* 114, 59–75.
- Faarlund, Jan Terje (2000). *Totenmålet*. Østre Toten kommune: Kulturetaten.
- Fraser, Norman and Greville G. Corbett (2000). Default gender. In Barbara Unterbeck et al. (eds.), *Gender in Grammar and Cognition*. Berlin: Mouton de Gruyter, 55–97.
- Fuhrhop, Nanna (1996). Fugenelemente. In Ewald Lang and Gisela Zifonun (eds.), *Deutsch—typologisch. Institut für deutsche Sprache—Jahrbuch 1995*. Berlin: Mouton de Gruyter, 525–550.
- Goldberg, Adele (1995). *Constructions*. Chicago: University of Chicago Press.
- Haiman, John (1985). *Natural Syntax: Iconicity and Erosion*. Cambridge: Cambridge University Press.
- Haspelmath, Martin (in press). Against markedness (and what to replace it with). *Journal of Linguistics*.
- Janda, Laura (2000). From number to gender, from dual to virile: Bridging cognitive categories. In Ellen Contini-Morava and Yishai Tobin (eds.), *Between Grammar and Lexicon*. Amsterdam: John Benjamins, 73–87.
- Kortmann, Bernd (2004). Introduction. In Bernd Kortmann (ed.), *Dialectology Meets Typology*. Berlin: Mouton de Gruyter, 1–10.
- Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar, Volume I: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Lehmann, Christian (1982). Universal and typological aspects of agreement. In Hansjakob Seiler and Franz Josef Stachowiak (eds.), *Apprehension: Das sprachliche Erfassen von Gegenständen, Teil II: Die Techniken und ihr Zusammenhang in Einzelsprachen*. Tübingen: Gunter Narr, 201–267.
- Lehrer, Adrienne (2000). Are affixes signs? In Sabrina, Bendjaballah, Wolfgang Ullrich, Dressler, Oskar E. Pfeiffer and Maria D. Voeikova (eds.), *Morphological Analysis in Comparison*, 143–155. Amsterdam: John Benjamins, 143–155.
- Levang, Synnøve (2003). Variasjon i bjerkreimsdialekten. Unpublished paper, Stavanger University College.
- Lie, Svein (1990). Indre Østlandet. In Ernst Håkon Jahr (ed.), *Den store dialektboka*. Oslo: Novus, 101–119.
- Lutzeier, Peter Rolf (1997). Gegensinn als besondere Form lexikalischer Ambiguität. *Linguistische Berichte* 171, 381–395.
- Lyons, John (1977). *Semantics, Volume 2*. Cambridge: Cambridge University Press.
- Maiden, Martin (2001). What sort of thing is a derivational affix? In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1999*. Dordrecht: Kluwer, 25–52.
- Maiden, Martin (2005). Morphological autonomy and diachrony. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2004*. Dordrecht: Kluwer, 137–175.

- Pintzuk, Susan (2003). Variationist approaches to syntactic change. In Brian D. Joseph and Richard D. Janda (eds.), *The Handbook of Historical Linguistics*. Oxford: Blackwell, 509–529.
- Plag, Ingo (1999). *Morphological Productivity*. Berlin: Mouton de Gruyter.
- Sapir, Edward (1921). *Language*. New York: Harcourt Brace.
- Serzisko, Fritz (1982). Numerus/Genus-Kongruenz und das Phänomen der Polarität am Beispiel einiger ostkuschitischen Sprachen. In Hansjakob Seiler and Franz Josef Stachowiak (eds.), *Apprehension: Das sprachliche Erfassen von Gegenständen. Teil II: Die Techniken und ihr Zusammenhang in Einzelsprachen*. Tübingen: Gunter Narr, 179–200.
- Spencer, Andrew (1999). Gender as an inflection category. *Essex Research Reports in Linguistics*, 25 (August 1999), 35–72.
- Spencer, Andrew (2001). The paradigm-based model of morphosyntax. *Transactions of the Philological Society* 99, 279–313.
- Spencer, Andrew (2003). Putting some order into morphology: Reflections on Rice (2000) and Stump (2001). *Journal of Linguistics* 39, 621–646.
- Steins, Carsten (2000). How to account for non-concatenative phenomena in a morpheme-based theory. In Barbara Stiebels and Dieter Wunderlich (eds.), *Lexicon in Focus* (Studia Grammatica 45). Berlin: Akademie-Verlag, 105–122.
- Stump, Greg T. (2001). *Inflectional Morphology*. Cambridge: Cambridge University Press.
- Torp, Arne (1997). Har bokmål to, tre eller fire genus? Unpublished manuscript, University of Oslo.
- Trosterud, Trond (2001). Genustilordning i norsk er regelstyrt. *Norsk Lingvistisk Tidsskrift* 19, 29–59.
- Unterbeck, Barbara (2000). Gender: New light on an old category. In Barbara Unterbeck et al. (eds.), *Gender in Grammar and Cognition*. Berlin: Mouton de Gruyter, xvi–xlvi.
- Wurzel, Wolfgang Ullrich (1984a). *Flexionsmorphologie und Natürlichkeit* (Studia Grammatica XXI). Berlin: Akademie-Verlag.
- Wurzel, Wolfgang Ullrich (1984b). Was bezeichnet der Umlaut im Deutschen? *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 37, 647–663.
- Wurzel, Wolfgang Ullrich (1989). Von der Inadäquatheit einer Affixmorphologie: weshalb morphologische Kategorienmarker nicht als eigene Einheiten im Lexikon repräsentiert sein können. *Linguistische Studien, Reihe A*, 194, 277–298.

*Institutt for Kultur- og Språkvitenskap
Universitetet i Stavanger
NO-4036, Stavanger*

hans.o.enger@uis.no

On the interaction of phonology and morphology in language acquisition and German and Dutch Broca's Aphasia: the case of inflected verbs

JANET GRIJZENHOUT and MARTINA PENKE*

1. INTRODUCTION

In this paper we first investigate how children acquire the German inflectional endings *-t/-d* and we subsequently consider how the inflectional endings *-t* and *-d* are realised by German and Dutch aphasics, respectively. By comparing the acquisition of an inflectional ending and the realisation of that ending in impaired language, and by contrasting the use of an inflectional ending in two different but related languages, we hope to reveal properties of the grammar that speakers use when they want to produce an inflected word.

A central issue with respect to learnability of a language is how to account for the fact that children produce structures that do not occur in the target adult language and which are considered to be universally 'unmarked'; for example, English children may produce word-final obstruents which are voiced in adult speech at a stage when they correctly voice word-initial obstruents (e.g., Bernhardt and Stemberger 1998:424, Smith 1973). With respect to language impairments, it has often been suggested that aphasic speech is also characterised by the fact that it displays unmarked patterns (e.g., Jakobson 1941/1971, Blumstein 1973, Dressler 1991, Smolensky 1999, Waugh and Lafford 2000). This paper investigates an interesting case of the emergence of the unmarked in child speech as well as in the speech of Broca's aphasics. In particular, we discuss how the realisation of inflectional endings in child language and in language disorders is affected by unmarked phonological structures. We will show that even though a child may have successfully acquired morphological markers, she may refrain from using them in a particular context when the phonology that she acquired thus far forces her to do so. Similarly, aphasics may successfully use morphological markers in most contexts, except in certain phonological environments. The languages under investigation here are German and Dutch since these languages display elaborate rhyme structures and allow consonants in word-final position which are not part of the syllable's rhyme. Moreover, these languages have segmental suffixes whose appearance results in interesting phonological phenomena at the right edge of words which—we will argue—are the result of the different strategies that German and Dutch adopt to comply with universal markedness conditions.

It is generally accepted that phonologically unmarked syllables are those without a final consonant and children initially tend to produce open syllables where the target language has a closed syllable. Our findings show that in German

child speech, both VVC and VCC rhymes are first realised with either a long vowel without a following consonant or a short vowel plus one consonant. Subsequently, we find more marked structures, i.e. rhymes with three positions but without an appendix. At this stage, segmental morphemes such as the participle ending *-t* are realised only if they fit the syllable structure that the child assumes. Finally, even more segments are realised in word-final position, but at this acquisitional stage we see the emergence of a markedness constraint ($*C_i\partial C_i$) which affects the realisation of inflectional endings and which was at that point 'hidden' in the child's grammar. Booij (1998, 2002) suggests this particular constraint—which says that schwa cannot be preceded and followed by the same consonant—to explain the fact that Dutch suffixes which have the structure schwa-consonant do not attach to stems in which the final consonant is identical to the consonant of the affix. Consider as an example the derivational affix *-er* (meaning 'someone who Vs') which attaches only to verbal stems that do not end in [r] (e.g., *dans+er* → *danser* 'dancer'). Stems that end in [r] select another allomorph to avoid the sequence *rər* (e.g., *huur+er* → *huurder*/**huurer* 'someone who rents'). We will make use of the same constraint for German inflection and argue that in order to explain the development of the inflectional ending *-t* for words that end in [t], the German child has to learn to demote the constraint $*C_i\partial C_i$ below the corresponding faithfulness constraint.¹

Interestingly, the effect of this particular markedness constraint, which appears relatively late in child language, also plays a role in the emergence of the unmarked in language disorders. Whereas it has often been stated that language impairment results in the emergence of the unmarked (Jakobson 1941/1971, Smolensky 1999, Waugh and Lafford 2000 and others), not much evidence has been provided until now to support this statement. In section 4 of this paper we will present a study confirming that Broca's aphasia often implies the loss of marked structures. We compare one particular case of the emergence of the unmarked in early child language with the emergence of the unmarked in Broca's aphasia, i.e. the lack of schwa between two identical consonants. By looking at child language and language disorders, we hope to be able to better understand the role of markedness in morphology. At the end of the paper, we will show that while in the first language acquisition unmarked structures emerge due to highly ranked markedness constraints, unmarked structures in aphasic speech emerge due to the loss of ranking between markedness constraints and faithfulness constraints.

The paper is organised as follows. In section 2 we outline major steps in the acquisition of German rhyme structure based on longitudinal data on German language acquisition. Section 3 analyses the data within the framework of Optimality Theory and shows that the acquisition process involves structure building as well as gradual demotion of markedness constraints. Section 4 presents a case for the emergence of the unmarked in participle formation in Broca's aphasia. Section 5 discusses the different strategies that are applied in first language

acquisition and language disorders regarding the emergence of unmarked prosodic structures at the right edge of words.

2. THE ACQUISITION OF CONSONANTAL POSITIONS WITHIN THE GERMAN RHYME

Very early German child speech is characterized by the fact that at least one consonant and one vowel appear in every word. Below we show that the next step in the acquisition of German rhyme structure involves sensitivity for syllable weight—in the sense that a two-positional rhyme is produced—but not for segmental faithfulness (cf. Grijzenhout and Joppen-Hellwig 2002). At this stage, we find syllables with an onset followed by a long vowel or a diphthong. The third step also involves a two-positional rhyme, but now more consonants per word are realised, so that words may consist of a consonant in onset position followed by a short vowel plus a different consonant. Subsequently, we find rhymes with three positions, but without an appendix. At this stage, segmental morphemes such as the inflectional ending *-t* should fit the syllable structure that the child assumes.

In section 2.1 we discuss the sources of the German child data on which our analysis is based. Section 2.2 briefly discusses German syllable structure. Section 2.3 illustrates the early acquisition of word-final consonants by three German children and discusses their realisations of inflectional suffixes after rhyme structure has been acquired successfully.

2.1. *The database for German child speech*

The data of German child speech used in this paper are taken from three sources: two diary studies (Annalena, see Elsen 1991, and Eva) and a longitudinal corpus based on spontaneous speech recordings (Naomi, see Grijzenhout and Joppen 1998a,b and Grijzenhout and Joppen-Hellwig 2002). Elsen (1991) is a diary study on the speech of Annalena who grew up in München. We consider the data collected from 28 January 1988 until 29 June 1989 (i.e. from age 0;11.29 until 2;05.30),² because the stages in the acquisition of German consonants we find during this time closely resemble the corresponding developmental stages we find in Eva and Naomi's speech.

The data from Eva are based on a diary study by the second author of this paper. Eva grows up in Bonn and is exposed to standard German. Data were collected from January 2002—when Eva was aged 1;1—until December 2002, when she was 2;0. For the time period between January and June, the data consist of a list of every new word Eva uttered. From July onwards, when Eva produced the first two-constituent utterances, the data consist of short transcripts of

spontaneous speech that were collected on a weekly basis. All speech utterances and participle forms were transcribed during or immediately after the respective form or utterance was produced. Apart from the weekly sessions when Eva's speech was transcribed, her acquisition of participles was monitored at home, so that all the participles Eva used at home between July and December are listed.

Data from the third child under investigation, Naomi, are based on a diary by her mother Sandra Joppen-Hellwig, who is a trained linguist, and on audiotape recordings. The diary covers utterances from age 1;2.06, when Naomi produced her first words, to 1;4.19. The audiotape recordings were made every week from age 1;4.26 until 2;1.20 at the child's home near Düsseldorf. Sandra Joppen-Hellwig phonetically transcribed all recordings the same day as the recordings took place and the transcriptions were checked by at least one other linguist.

Before discussing the developmental stages of German children in the acquisition of final consonants and the loss of inflectional endings in German Broca's aphasia, it is useful to first look at the structure of rhymes in adult German.

2.2. German rhyme structure

The minimal requirement for a German word is that it contains at least one syllable which has either a long vowel, as in the word *Kuh* [k^hu:] 'cow', or a diphthong, as in the word *Bau* [bau] 'building, burrow', or a sequence of a short vowel plus a consonant, as in *Ball* [bal] 'ball' and *mit* [mit] 'with'. There are no words in adult speech that consist of one syllable with a full short vowel without a following segment, which indicates that the rhyme has minimally two positions in German.³ This restriction on the phonology of German has to be learned by the child (cf. section 2.3).

In word-final position, the rhyme of a syllable may be followed by an obstruent (1a–c), or a sonorant consonant (2a–c):

- | | | | | | | | |
|-----|----|------|--------|------------------|--------|---------|---------|
| (1) | a. | Bad | [ba:t] | 'bathroom' | Schlaf | [ʃla:f] | 'sleep' |
| | b. | Neid | [nait] | 'envy' | reif | [raif] | 'ripe' |
| | c. | bald | [balt] | 'soon' | Elch | [ʔelç] | 'moose' |
| (2) | a. | Bahn | [ba:n] | 'railroad, tram' | | | |
| | b. | Bein | [bain] | 'leg' | | | |
| | c. | Halm | [halm] | 'straw' | | | |

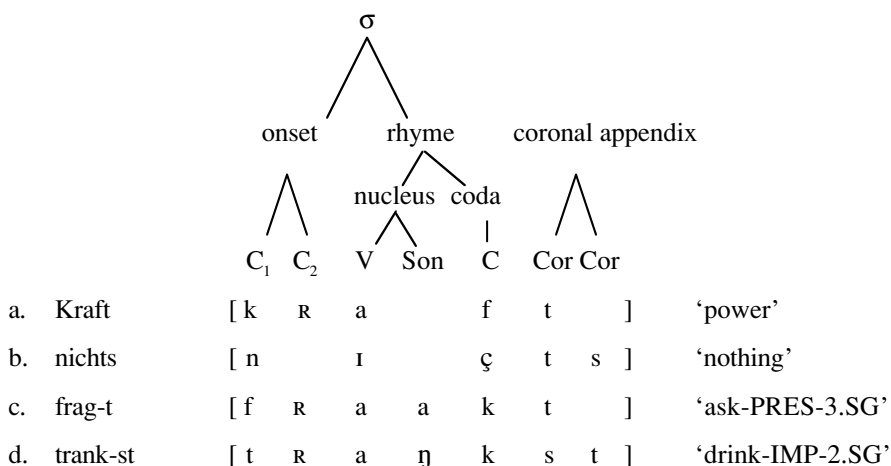
The data presented above are compatible with the so-called 'sonority sequencing principle' which requires post-vocalic consonants to fall in sonority (e.g., Jespersen 1904). To ensure that the sonority sequencing principle is not violated,

Grijzenhout (2001) proposes that the two nucleus positions may be occupied by a long vowel, or a short vowel plus a sonorant (i.e. a glide, a liquid or a nasal) and that obstruents never occupy a nucleus position. Segments in the nucleus may be followed by a less sonorant consonant in the coda position of a rhyme.⁴ Notice, however, that some post-vocalic consonant clusters have an equal level of sonority. In these clusters, the second member is always an alveolar obstruent:

- (3) a. Gips [gɪps] ‘plaster’
- b. nicht [nɪçt] ‘not’
- c. Keks [ke:ks] ‘biscuit’

Vennemann (1988) proposes that a so-called ‘coronal-appendix’ outside the rhyme accommodates word-final alveolar obstruents in German. The appendix is not part of the rhyme and the sonority sequencing principle—which holds within the syllable—is therefore not violated by final coronal obstruents.⁵

(4) *German Syllable Structure*



In German, the coronal obstruents /t/ and /s/ often function as morphemes (e.g., -t functions as the 3rd person singular morpheme, the 2nd person plural morpheme or it marks the past participle form, -st marks the 2nd person singular, and -s marks genitive singular masculine and neuter forms or the plural). These segments are often realised after a VVC- or VCC-rhyme where they occupy the appendix (cf. 4c and d).

In summary, the rhyme structure that we assume for German is characterised as follows: (i) the nucleus may be filled by vocalic segments and sonorant consonants and (ii) the rhyme is minimally bipositional and maximally tripositional and

may be followed by at most two coronal obstruents. The next section examines the realisation of post-vocalic consonants in early German child speech.

2.3. Stages in the acquisition of German rhyme structure

At the earliest stage of word production (stage I) the three German children in our study tend to produce words with an initial consonant followed by a short or a long vowel (e.g., 5a and b), as well as words without an initial consonant, but with a final consonant (e.g., 5c and d).

(5) *Very early German child speech (Stage I): one consonant and one vowel*

	Spelling	Adult form	Child's form		Gloss
a.	Ball	[bal]	[ba]	(Naomi 1;2)	'ball'
b.	Löffel	[løfl]	[lɔ:]	(Eva 1;1.7)	'spoon'
c.	an	[ʔan]	[an]	(Naomi 1;2)	'at'
d.	auf	[ʔauf]	[af]	(Annalena 1;2.10, Naomi 1;3)	'on'

Gradually, words with an initial consonant and a single short vowel become rare and most monosyllabic words have a long vowel (6a–d) or a short vowel plus a consonant (7a–c). In addition, an initial consonant is sometimes followed by a diphthong (8a–d).

(6) *Long vowels in early German child speech (Stage II)*

	Spelling	Adult form	Child's form		Gloss
a.	Bahn	[ba:n]	[ba:]	(Naomi 1;4.26)	'railway'
b.	Buch	[bu:x]	[bu:]	(Naomi 1;4.26)	'book'
c.	Mond	[mo:nt]	[mo:]/[mɔ:]	(Naomi 1;5.21/Eva 1:2.7)	'moon'
d.	Bär	[be:ʁ]	[ba:]	(Eva 1;2.21)	'bear'

(7) *Short vowels and coda consonants in early German child speech (Stage II)*⁶

	Spelling	Adult form	Child's form		Gloss
a.	ab	[ʔap]	[ap ^h]	(Annalena 1;2.15–1;5.30)	'off, away'
b.	ab	[ʔap]	[ap ^h]	(Naomi 1;4.26)	'off, away'
c.	satt	[zat]	[at ^h]	(Annalena 1;2.19) ⁷	'satisfied'

(8) *Diphthongs in early German child speech (Stage II)*

	Spelling	Adult form	Child's form		Gloss
a.	Bauch	[baux]	[bauw]	(Annalena 1;1.29)	'belly'
b.	drauf	[drauf]	[daw]	(Annalena 1;4–1;6)	'on it'
c.	Schwein	[ʃvain]	[vaɪ]	(Eva 1;3.07)	'pig'
d.	nein	[naɪn]	[naɪ]	(Naomi 1;5.01)	'no'

At this acquisitional stage, German children thus realise a rhyme with two positions filled by either a long vowel, or a diphthong, or a short vowel followed by one consonant.⁸ We attribute this to a so-called 'minimality requirement' which says that a rhyme must branch (i.e. a rhyme has at least and at most two moras, see section 3.1 below).

At the first two stages, there is a strong tendency for one consonantal place of articulation per word, but soon more and more exceptions to this generalisation begin to emerge. The examples below each have two consonants with different places of articulation. At stage III, coda consonants are realised even when another consonant is present somewhere else in the word (9a–c), and—in Naomi's case—sometimes at the expense of losing vowel length (10a–c).

(9) *Two consonantal places of articulation in German child speech (Stage III)*

	Spelling	Adult form	Child's form		Gloss
a.	Licht	[liçt]	[liç]	(Eva 1;3.25)	'light'
b.	Mann	[man]	[man]	(Naomi 1;5.01, Eva 1;3.14)	'man'
c.	kaputt	[ka'put]	[butʰ]	(Naomi 1;6.05) ⁹	'broken'

(10) *Two consonantal places of articulation in German child speech (Stage for Naomi in which she realises a coda consonant after a short vowel only)*

	Spelling	Adult form	Child's form		Gloss
a.	warm	[va:m]	[bam]	(Naomi 1;6.05)	'warm'
b.	Bahn	[ba:n]	[ban]	(Naomi 1;6.12)	'tram'
c.	Buch	[bu:x]	[buχ]	(Naomi 1;6.19)	'book'

The next stage is characterised by the fact that most adult words which have three positions in the rhyme are realised by all three children with three positions in the rhyme:

(11) *Three positions in the syllabic rhyme in German child speech (stage IV)*

	Spelling	Adult form	Child's form		Gloss
a.	Zaun	[tsaʊn]	[daɪn]	(Annalena 1;6.27)	'gate'
b.	Hund	[hʊnt]	[hu:tʰ]	(Naomi 1;6.12)	'dog'
c.	Milch	[mɪlç]	[mi:ç]	(Naomi 1;7.02)	'milk'
d.	Baum	[baʊm]	[baʊm]	(Eva 1;3.28)	'tree'

At this stage, we do not find extra-syllabic elements yet, i.e. the rhyme may consist of three elements and the appendix in (4) is not yet available to accommodate an extra coronal obstruent. When the adult form consists of a long vowel plus two consonants, the child realises either the long vowel followed by one of the two consonants, or a short vowel followed by two consonants. The selection of the consonant following the long vowel varies among children; they sometimes realise the consonant immediately adjacent to the vowel (12a), or they realise the final consonant of the word (12b):

(12) *At most three positions in the syllabic rhyme in German child speech (stage IV)*

	Spelling	Adult form	Child's form		Gloss
a.	Mond	[mo:nt]	[mo:n]	(Eva 1;3.28)	'moon'
b.	Mond	[mo:nt]	[mo:t]	(Naomi 1;6.12)	'moon'

During this stage, the first inflected forms with final *-t* appear. In early German child language, *-t* is either used as a finiteness marker on verbs (cf. Clahsen and Penke 1992, Clahsen, Penke and Parodi 1994)¹⁰ or it appears on past participles. For Eva, the inflectional ending *-t* appears for the first time at 1;4.07 and for Naomi at 1;7.27. Note that for both children, the rhyme still has at most three positions: a long vowel may be followed by at most one consonant and a short vowel by a sequence of a sonorant and an obstruent.¹¹

(13) *At most three positions in the rhyme and realisation of inflectional -t in German child speech (stage IV)*

	Spelling	Adult form	Child's form		Gloss
a.	mal-t	[ma:lt]	[ma:t]	(Eva 1;4.7)	'paint-3.SG'
b.	mal-t	[ma:lt]	[malt]	(Eva 1;5.7)	'paint-3.SG'
c.	stimm-t	[ftɪmt]	[tɪnt]	(Naomi 1;7.27)	'be right-3.SG'
d.	schwimm-t	[ʃvɪmt]	[vɪmt]	(Naomi 1;9.11)	'swim-3.SG'

The next development (stage V) is characterised by the emergence of the coronal appendix, i.e. a long vowel or a diphthong may be followed by two consonants of which the last one is a coronal obstruent ([t] or [s]):

(14) *The rhyme and a coronal appendix in German child speech (stage V)*

	Spelling	Adult form	Child's form		Gloss
a.	wein-t	[va:nt]	[va:nt]	(Eva 1;6.14)	'cry-3.SG'
b.	schlaf-t	[ʃle:ft]	[la:ft]	(Naomi 1;10.10)	'sleep-3.SG'

We thus find that the use of the inflectional suffix *-t* may result in the deletion of a stem-final consonant (13a) or vowel shortening (13b) at stage IV due to the fact that the child does not realise a syllabic appendix yet, whereas both the stem-final consonant and the inflectional ending *-t* are realised at stage V (14a and b).

However, matters are more complex: some verbal stems already end in a coronal stop which poses a challenge for the child who wants to realise both the stem-final coronal stop and the suffix. Generally, languages avoid sequences of two adjacent identical elements (e.g., Leben 1973, Goldsmith 1976, McCarthy 1986, Yip 1998). The strategy adopted in adult German is the following: the inflectional suffix *-t* is realised as $[-\text{ət}]$ when the stem ends in a coronal stop. Thus, the stems *rett-* and *find-* are followed by $[-\text{ət}]$ in the inflectional forms *rettet* 'save-3.SG' and *fin.det* 'find-3.SG', respectively.¹²

An inflectional affix *-t* also occurs in past participle forms; in German weak verbs, past participle forms are built by prefixing *ge-* and suffixing *-t* to the stem of a verb (15a and b). In strong verbs, the stem vowel usually changes and the ending is *-en* rather than *-t* (15c and d). When a stem of a weak verb ends in a coronal stop (/t/ or /d/), schwa is inserted between the final stem consonant and the inflectional ending (15e and f).¹³

(15) *German infinitives and past participles*

	infinitive	past participle	Gloss
a.	lachen	ge-lach-t	'to laugh'
b.	machen	ge-mach-t	'to make'
c.	kommen	ge-komm-en	'to come'
d.	gehen	ge-gang-en	'to go'
e.	retten	ge-rett-et	'to save'
f.	landen	ge-land-et	'to land'

From stage V onwards, Eva produces *-t* inflected forms regularly, except for stems with a final [t].¹⁴ We found that between 1;6 and 1;9 Eva regularly realises inflectional *-t* after stems not ending in [t] (e.g., 16a) and avoids building 3rd person singular forms for verbs with stem final [t]. This is remarkable, because Eva knows verbs with stem-final coronal stops well and uses them in 6.9% of the contexts of *n*-inflection (19 of 277 cases), but strikingly avoids them in a context where they would get the ending *-t*: of the 93 times that she uses *-t* inflected 3rd singular forms, only 2 cases involve stems with a final coronal stop (2.1%). During this stage, we find only six contexts for a 3rd singular form of a verb with stem final [t]. In these few contexts, Eva omits the inflectional ending once (17%) (16b), produces incorrect *-n*-inflected forms three times (50%) (e.g. 16c) and realises the target form with schwa-insertion twice (33%) (e.g. 16d).

(16) *Avoidance of inflectional -t after stems ending in coronal stop in German child speech (stage V)*

	Spelling	Adult form	Child's form		Gloss
a.	komm-t	[kɔmt]	[dɔmt]	(Eva 1;5)	'come-3SG'
b.	find-et	[fɪndət]	[fɪnt]	(Eva 1;8.11)	'find-3SG'
c.	bad-et	[ba:dət]	[ba:dɪ]	(Eva 1;9.0)	'bath-3SG'
d.	pust-et	[pustət]	[pustət]	(Eva 1;8.01)	'blow-3SG'

Figure 1 illustrates that this distribution differs from the realisation of verb forms for verbs without a final coronal stop in 3rd singular contexts during stage V: column 5 shows that 79% of these verbs are correctly inflected with the affix *-t*

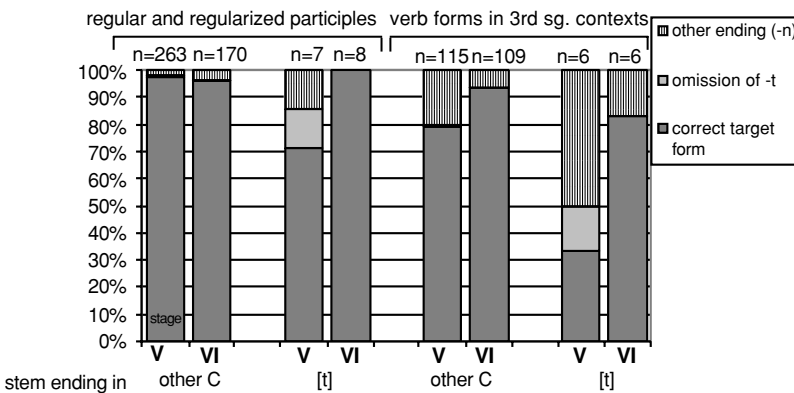


Figure 1. Different realisations of regular and regularised participles and of verb forms in 3rd singular contexts for Eva in acquisition stages V and VI

(i.e. 91 of 115 cases), the affix *-t* is never omitted for these verbs and in only 21% of the 3rd singular contexts incorrect *-n*-inflected forms are produced (i.e. 24 of 115 cases). The incorrect *-n*-inflected forms are so-called ‘root infinitives’, i.e. a non-finite verb form appears in a root clause instead of the target finite form. A comparison between columns 5 and 7 in Figure 1 reveals that Eva makes more use of this means of avoiding a finite 3rd singular form for verbs with a stem-final coronal stop (50%, i.e. 3 of 6 cases) than for verbs without a stem-final coronal stop (21%, i.e. 24 of 115 cases).

Although the database is rather small, columns 7 and 8 in Figure 1 show that a change takes place from 1;9 onwards (stage VI). Whereas the first realisation of the verb *bad-* ‘bath’ in a 3rd singular context on this day is again a root infinitive (as in 16c), all following realisations of this verb later that day result in the correct target form *badet*. From this day onwards until the end of the data collection period (at 2;0) all 3rd singular forms for verbs with stem final coronal stops are correctly inflected with $[-\text{ət}]$.

A similar development can be observed for *-t* inflected participles in Eva’s data. As with 3rd person singular forms, Eva carefully avoids the realisation of *-t* inflected participles for verbs whose stem end in a coronal stop in stage V. The first participles occur in Eva’s speech at 1;6.28. Until the age of 1;9.23 she produces *-t* inflected participle forms for 257 of 263 weak verbs (=98%) which do not end in a coronal stop (cf. column 1 in Figure 1). During the same time, however, she strikingly avoids participles when the verb’s stem ends in $[t]$, even though she knows these verbs well. Stage V only comprises seven participle forms for two verb types with stem final $[t]$ (cf. column 3 in Figure 1). The participle of the verb *bad-* ‘bath’ is first realised by her as $[\text{g}\text{ə}\text{b}\text{a}\text{d}\text{n}]$ (17a).¹⁵ From 1;7.20 onwards, all other five instances of this verb are correctly produced as *gebadet* (17b). The only other participle of a weak verb with stem final $[t]$ that appears during this stage is the form *gepust* ‘blow’, where the target $[-\text{ət}]$ inflection is omitted (17c).

(17) *Eva’s realisation of German participles from 1;6.28 until 1;10.5: avoidance of $[\text{t}\text{ət}]$*

	Stem	Spelling	Adult form	Child’s form	Gloss
a.	<i>bad-</i>	<i>gebadet</i>	$[\text{g}\text{ə}\text{b}\text{a}\text{d}\text{ət}]$	$\text{g}\text{ə}\text{b}\text{a}\text{d}\text{n}$ (Eva 1;7.14)	‘bathed’
b.	<i>bad-</i>	<i>gebadet</i>	$[\text{g}\text{ə}\text{b}\text{a}\text{d}\text{ət}]$	$\text{g}\text{ə}\text{b}\text{a}\text{d}\text{ət}$ (Eva 1;7.21)	‘bathed’
c.	<i>pust-</i>	<i>gepustet</i>	$[\text{g}\text{ə}\text{p}\text{u}\text{st}\text{ət}]$	$\text{g}\text{ə}\text{p}\text{u}\text{st}$ (Eva 1;9.23)	‘blew’

From age 1;10.5 onwards, Eva faithfully realises *-t* inflection with schwa insertion (cf. columns 2 and 4 in Figure 1) and her productions suddenly contain a number of different verb types with stem-final $[t]$ that are inflected with $[-\text{ət}]$ in participle forms (eight different verb types, e.g., 18a–c).

(18) *Eva's realisation of German participles from 1;10.5: regular realisation of final [tət]*

	Stem	Spelling	Adult form	Child's form	Gloss
a.	hust-	hustet	[gəhustət]	[gəhustət] (Eva 1;10.5)	'coughed'
b.	auspust-	rausgepustet	[rausgəpustət]	[rausgəpustət] (Eva 1;10.10)	'blown out'
c.	reit-	geritten	[gərɪtən]	[gərɪtət] (Eva 1;10.22)	'rode' ¹⁶

Whereas Eva notably avoids *-t* inflected forms for verbs ending in [t] at one developmental stage, Naomi's data reveal a slightly different picture. Note first of all that Naomi pronounces velar [g] and [k] as [t]. The verb *guck-* 'look', for example, which is very frequent in child speech, is pronounced as [tut] by Naomi. Hence, not only verbal stems that end in a coronal stop, but also verbal stems that end in a velar stop in adult German are pronounced with a coronal stop by Naomi. As a result, she has a higher number of productions of verbs with a stem-final coronal stop. This fact might possibly explain why for Naomi—in contrast to Eva—we do not find a stage when she avoids producing forms which involve a stem-final coronal stop and the affix *-t*. In her speech, the first *-t* inflected forms for verbal stems which do not end in a coronal stop appear at 1;7.27 and the first *-t* inflected forms for verbal stems ending in [t] appear at 1;9.18. From that age onwards until the last recording available to us at age 2;1.20, Naomi varies between three realisations for inflection after [t]-final stems: (i) she omits *-[ət]* (19a,d) in the majority of cases (71.4%, i.e. 65 out of 91 contexts for *-[ət]* inflection in 3rd singular contexts and participle forms), (ii) she produces a sequence of two identical elements (19c,f) in a few cases (11%, 10 of 91) and (iii) she produces correct target forms with *-[ət]* (19b,e) in very few instances (6.6%, 6 of 91).

(19) Naomi's realisation of German 3rd singular forms and past participles

	Spelling	Adult form	Child's form	Gloss
a.	schneid-et	[ʃnaɪdət]	[nɪt] (Naomi 2;1;03)	'cut-3.SG'
b.			[nɪdət] (Naomi 2;1;03)	
c.	geschnitten	[gəʃnɪtən]	[nɪtt] (Naomi, 2;1;03)	'cut-PART' ¹⁷
d.	angeguckt	[angəkʊkt]	[tət] (Naomi 1;11;24–2;0;26)	'look at-PART'
e.			[tətət] (Naomi 1;9;18)	
f.			[anətutt] (Naomi 2;1;03)	

Figure 2 indicates that this distribution is in marked contrast to inflected forms Naomi produces for verbs not ending in stem final [t]. For these verbs, the *-t*

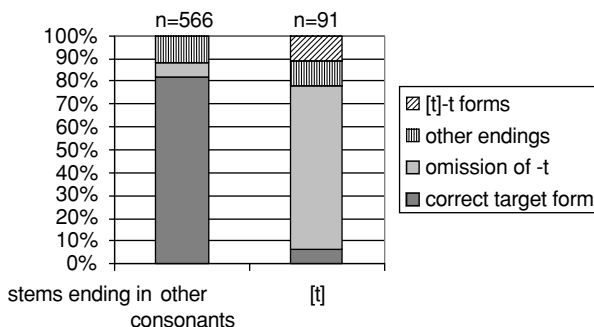


Figure 2. Different realisations of regular and regularised participles and of verb forms in 3rd singular contexts for Naomi

inflection is correctly realised in 82.1% (465 of 566) of the relevant contexts compared to 6.6% of correct $[-\partial t]$ inflection after stem final $[t]$. Also, whereas the omission rate of $[-\partial t]$ inflection after stem-final $[t]$ is 71.4% in 3rd singular and participle contexts, the omission rate for the affix $-t$ in verbs without stem-final $[t]$ is only 6.4% (36 of 566). This difference between correctly and incorrectly produced forms for verbs with and without stem-final $[t]$ in 3rd singular and participle contexts is highly significant ($\chi^2 = 218.838$ Yates corrected, $p < .000$).

At 2;1.20 (i.e. her age at the last recordings available to us), Naomi still has not reached a stage where she only realises $[-\partial t]$ after verbal stems that end in $[t]$.

Our observation that the affix $-t$ is often omitted after verbal stems ending in a coronal stop is also confirmed by other data on German child language: in the published data of Annalena's speech (Elsen 1991), the first $-t$ -inflected forms of verbal stems ending in $[t]$ appear at 1;7.19 when the child varies between the realisations *duft* and *duftet* ('smell-3.SG'). From 1;7.19 until 1;11.30, Annalena omits the inflectional ending for stems ending in $[t]$ in 5 cases (24%) and realises $[\partial t]$ 16 times (76%).

We will now briefly account for the observed stages in the development of rhyme structure and the inflectional ending $-t$ within the framework of Optimality Theory.

3. ACQUISITIONAL STAGES AND CONSTRAINT DEMOTION

In OT literature, the emergence of the unmarked in adult grammars is attributed to markedness constraints that militate against marked structures. Gnanadesika (1996) and others propose that at the initial stage in language acquisition, markedness constraints outrank faithfulness constraints and Tesar and Smolensky (1993) propose that learning a grammar involves step-by-step

demotion of single constraints below corresponding faithfulness constraints based on available positive evidence. Under these proposals, children thus go through a stage where the markedness constraint *NoVoicedCODA*, for example, outranks the faithfulness constraint *IDENTITY (VOICE)*. During this stage, children will devoice all word final obstruents—even those which are voiced in the target adult language. Based on positive evidence provided by word pairs such as *bet-bed* English children will demote the markedness constraint *NoVoicedCODA* below the faithfulness constraint *IDENTITY (VOICE)* resulting in voiced final obstruents. In contrast to English children, Dutch and German children never encounter evidence to demote *NoVoicedCODA* to a position below *IDENTITY (VOICE)*. We will now briefly turn to an OT account of the observed stages in the acquisition of German rhyme structure.

3.1. Rhymes and inflectional *-t* in German child language

The steps in the development of rhymes described above can be characterised as follows:

(20) Summary of first steps in the acquisition of rhyme structure:

- Stage I: Vowel length is not distinctive; One consonant per word;
Inflection has not been acquired yet
- Stage II: Rhyme exactly two positions; One consonant per word;
Inflection has not been acquired yet
- Stage III: Rhyme exactly two positions; More than one consonant per word;
No instances of the inflectional ending *-t* are attested in our data.
- Stage IV: Rhyme may have three positions; More than one consonant per word;
First instances of the inflectional ending *-t* appear.

Recall from section 2.3 that adult CV(C)C- and CV:C-words which had a CV:-structure in stage II (cf. examples 6a–d and 8a–d) are more often realised with a CVC-structure in stage III (cf. examples 9a–c and 10a–c). This effect can be attributed to the fact that the children realise more consonantal places of articulation per word in stage III. In OT-terms, this means that the child reranks markedness constraints and faithfulness constraints. In particular, we see that a markedness constraint penalising consonantal places of articulation within the rhyme (i.e. *NoCODA*) is demoted to a position lower than a faithfulness constraint which says that segments may not be deleted (i.e. *MAX-IO*).

Table 1. Structure of Eva's and Naomi's rhymes for adult -VVC and -VCC rhymes

<i>Child</i>	Input Rhyme CVXC	Rhyme one position (-V), <i>n</i> (in %)	Rhyme two positions (-VV, -VC), <i>n</i> (in %)	Rhyme three positions (-VVC, -VCC), <i>n</i> (in %)
<i>Eva</i>	Stage III (1;3.0–1;3.28)	2 (5%)	31 (72%)	10 (23%)
	Stage IV (1;4.0–1;5.14)	0	5 (17%)	25 (83%)
<i>Naomi</i>	Stage III (1;5.01–1;6.05)	3 (6%)	35 (71%)	11 (23%)
	Stage IV (1;6.12–1;7.27)	0	19 (24%)	60 (76%)

Table 1—based on collected data from Eva and Naomi—shows the gradual transition from a stage in which a bipositional rhyme is preferred to a stage where -VVC and -VCC rhymes are mostly realised with three positions.

As illustrated in Table 1, Eva realises 72% of words which have a three positional rhyme in adult German with maximally two positions in the rhyme until 1;3.28. From 1;4.0 onwards, words with a long vowel plus a consonant and words with a short vowel plus two consonants start to emerge more frequently. The same development can be observed in Naomi; from 1;5.01 she most often produces words with two positions in the rhyme. From 1;6.12, the rhyme structure is more complex (i.e. a rhyme may have three positions). We tentatively assume that the step from stage III to IV involves the minimal demotion of a markedness constraint which says that a rhyme is minimally and maximally bimoraic (i.e. BRANCHING RHYME or “BR”) to a stratum below the faithfulness constraint which says that the output should be faithful to the number of moras in the input, i.e. MAX- μ . This constraint ranking accounts for the fact that at stage IV, the rhyme has at most three positions: a long vowel may be followed by at most one consonant and a short vowel by two:

(21) Input form /ba:n/; Naomi's and Eva's outputs are [ba:n] at stage IV

	ONS	MAX-IO _i	MAX- μ	BR	NoCODA
a. an	*!	*	*		*
b. ba:		*!	*		
c. ban			*!		*
d. ba:n				*	*

At stage IV, the rhyme has at most three positions: a long vowel may be followed by at most one consonant and a short vowel by two. At this stage, the first inflected forms with final *-t* appear. For Eva this stage starts at 1;4.7, for Naomi at 1;7.27. The reason why a stem-final consonant is not realised or a stem vowel is shortened when the inflectional ending *-t* follows a stem, may be the fact that the child does not have a coronal appendix yet, i.e. the markedness constraint NOAPPENDIX (Sherer 1994) is never violated and has not been demoted to a position in the ranking below the faithfulness constraint MAX-IO yet:

- (22) *Input form /mo:nt/; Eva's and Naomi's outputs are [mo:n] and [mo:t], respectively, at stage IV (Eva preserves the linear order between the vowel and a following consonant, whereas Naomi seems to prefer to realise the final segment of a word)*

	ONS	NOAPPENDIX	MAX-IO	MAX- μ	BR	NoCODA
a. mo:			**!	*		
b. mon			*	*!		*
c. Ⓢ mo:n			*		*	*
d. Ⓢ mo:t			*		*	*
e. mo:nt		*!			*	*

The next development is the emergence of the coronal appendix, i.e. a long vowel or a diphthong may be followed by two consonants of which the final one is a coronal obstruent. This implies that the markedness constraint NOAPPENDIX has been demoted to a position in the ranking below the faithfulness constraint MAX-IO at stage V. Hence, the grammar of the child now allows for the inflectional ending *-t* after a tripositional rhyme in words like *schläft* [ʃle:ft] ‘sleeps’ and *weint* [vaɪnt] ‘cries’. The reason why inflectional *-t* sometimes still does not appear after stems ending in [t] may be attributed to the constraint $*C_1\text{ə}C_1$ that Booij (1998:156–160) proposes for Dutch and which penalises the realisation of schwa between two identical consonants.¹⁸ We conclude from the observation that the inflectional ending *-t* is not realised after verbal stems ending in [t] that initially, the markedness constraints $*C_1C_1$ —which prohibits a sequence of two identical consonants—and $*C_1\text{ə}C_1$ —which prohibits the occurrence of schwa between two consonants with identical manner and place of articulation—outrank MAX-IO.

- (23) *Input stem /find/ or /nait/ + suffix -t; Naomi's outputs are [fɪnt] or [nait] at stage V*

	*C _i C _i	*C _i əC _i	MAX-IO	NOAPPENDIX
a. \mathcal{F} fɪnt / nait			*	
b. fɪntt / naitt	*!			*
c. fɪndət / naidət		*!		

From 1;7.19 to 1;11.30 Annalena seems to employ two grammars: one grammar like (23) which she uses in the minority of cases (24%) in which *C_iC_i, *C_iəC_i ≫ MAX-IO (for the output [duft] rather than [duftt] or [duftət] for adult [duftət] ‘smells’) and one which she uses more often (76%) in which the markedness constraint *C_iəC_i is demoted below MAX-IO (so that the verbal stem *duft* plus the inflectional ending *-t* is realised as [duftət]).

For Naomi, demotion of the markedness constraints apparently involves a stage in which she has not determined the ranking exactly. First the markedness constraints outrank the faithfulness constraint MAX-IO as in (23). Subsequently, reranking takes place so that variation is possible (see examples in (19) above). We assume that Naomi tries out different rankings (mostly *C_iC_i, *C_iəC_i ≫ MAX-IO for the optimal output *nait*, but also MAX-IO ≫ *C_iC_i for *naitt* or MAX-IO ≫ *C_iəC_i for *naidet*) and finally the markedness constraint *C_iəC_i is demoted to a position below the faithfulness constraint MAX-IO as it is for Eva (see 24), who seems to skip the intermediate stage in which we find variation for Naomi.¹⁹

- (24) *Input stem /find/+ suffix -t; Eva's output is findet at stage VI*

	*C _i C _i	MAX-IO	*C _i əC _i
a. fɪnt		*!	
b. fɪntt	*!		
c. \mathcal{F} fɪndət			*

According to the analysis presented here, German children have to demote *C_iəC_i below the faithfulness constraint which says that segments—and, hence, inflectional segmental suffixes—in the input should be realised in the output.

Note that in this respect the German ranking is more marked than, for instance, the Dutch ranking. In both German and Dutch, regular past participles are formed by adding a prefix (written as *ge-*) and a suffix (i.e. a coronal stop) to a verbal stem. When the stem ends in a coronal stop, the two languages display different strategies; in German, schwa appears between the stem and the affix (*ge + rett + t* → [gəretət] ‘rescued’), whereas in Dutch only one of the two stops is realised and devoiced at the end of a syllable (*ge + red + d* → [xəret]).

(25) *Input stem /rɛd/+ suffix -t; children’s and adult Dutch output [xəret] ‘saved’*

	$*C_i C_i$	$*C_i \emptyset C_i$	MAX-IO
a. ☞ <i>xəret</i>			*
b. <i>xərett</i>	*!		
c. <i>xəredət</i>		*!	

Thus, in contrast to German children, Dutch children do not have to demote the markedness constraint $*C_i \emptyset C_i$ below MAX-IO.

(26) *Markedness constraint outranks faithfulness constraint in early child speech until the child demotes the markedness constraint (based on positive evidence)*

- a. Initial state (German and Dutch): $*C_i \emptyset C_i \gg \text{MAX-IO}$
- b. Final state for German: $\text{MAX-IO} \gg C_i \emptyset C_i$

Next we will consider the role of the constraints discussed in this section in the speech of German and Dutch aphasics.

4. THE EMERGENCE OF UNMARKED STRUCTURES IN GERMAN AND DUTCH BROCA’S APHASICS

In this section, we will investigate the emergence of unmarked structures in Broca’s aphasia. Broca’s aphasia is an acquired language disorder that is typically caused by strokes affecting anterior parts of the left hemisphere. The speech production of agrammatic Broca’s aphasics is characterised by the omission of function words and by a reduction of sentence length and syntactic complexity.

This leads to the preponderance of one or two word utterances in spontaneous speech. Bound inflectional morphology can be omitted or substituted (e.g. Menn and Obler 1990, Grodzinsky 1990).

It has sometimes been suggested that language impairments due to aphasia mostly affect marked structures and that unmarked structures are maintained (e.g., Jakobson 1941/1971, Blumstein 1973, Dressler 1991, Smolensky 1999, Waugh and Lafford 2000). However, until now, evidence for this assumption has been scarce. We will present data that support this assumption focussing on realisations of regular past participles in German and Dutch. Here we report data from three elicitation tasks performed with 6 German and 12 Dutch agrammatic Broca's aphasics.

4.1. Past participles in German Broca's aphasia

The data on German Broca's aphasia was collected in an elicitation task on regular and irregular German participle inflection (see Penke et al. 1999). All tested subjects were right-handed native speakers of German, who had no language problems prior to aphasia, and who evinced aphasia secondary to a left hemispheric brain damage at least 3 years before our investigation. Broca's aphasia was classified by the standard *Aachen Aphasia test-battery* (Huber et al. 1983).

Subjects had to complete a sentence by transforming a given 1st person singular present tense form into a participle (27). Sentences were presented on cards placed before the subjects and were read out aloud.

- | | | |
|------|----------------|-------------------------|
| (27) | Ich schreibe | (<i>I am writing</i>) |
| | Ich habe ____. | (<i>I have</i> —) |

Per subject we elicited 39 regular and 39 irregular participles which were controlled for lemma and participle frequency. Regular test items consisted of six verbs with stems ending in [t] (e.g., *heft-* 'to staple') and 33 verbs with stems ending in other segments (e.g., *leb-* 'to live'); 8 of the irregular test items had a stem ending in [t] (e.g., *flecht-* 'to braid', *tret-* 'to kick') and the stems of the other 31 irregular verbs ended in other segments.

Here we will focus on the data of 6 of the originally 11 tested subjects, namely on those subjects who committed errors where the participle suffix *-t* was omitted on regular verbs (for the original data, see Penke et al. 1999).

Figure 3 presents an overview of the results for *-t*-inflected regular participles, distinguishing between verbs with a stem-final [t] (columns on the left-hand side

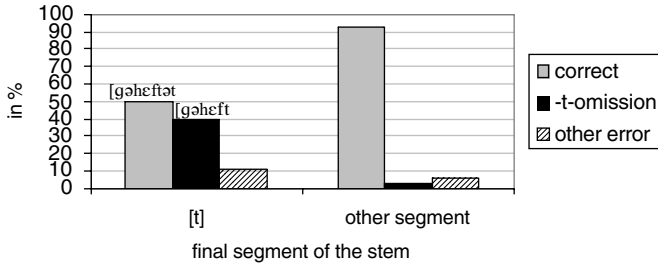


Figure 3. Results for regular German participles

of the figure) and verbs with other stem final segments (columns on the right-hand side).

The figure indicates a clear correlation between the segmental ending of the verbal stem and errors in participle production. Whereas 92.3% of the participle forms for regular verbs with stem final segments other than [t] are correctly inflected with *-t*, the correctness scores for regular verbs ending in [t] drop significantly to only 50% (grey columns) (Wilcoxon, $df = 5$, $p = .028$). A look at the incorrectly produced forms reveals that for verbs with stem ending in [t] the German aphasic subjects produce a total of 39.5% forms in which the affix *-t* is omitted (e.g., *geheft* instead of *geheftet*). For verbs ending in other segments, participle forms without overt realisation of the suffix (*geleb* instead of *gelebt*) occur in only 2.4% of the cases (cf. Janssen and Penke 2002), a significant difference (Wilcoxon, $df = 5$, $p = .028$). A comparison between the number of correctly produced forms such as *gerettet* and the number of affix omissions (i.e. *geret*) for verbs with a stem-final segment [t] reveals that the distribution of these two forms in the data of the six subjects is compatible with a chance selection between these two forms.

The significant difference in error rates between the two types of regular *-t*-inflected participles and the high correctness score for participles of verb stems ending in segments other than [t] indicate that occurring errors are not caused by a deficit with regular inflection per se (cf. Ullman et al. 1997),²⁰ but are due to the phonological shape of the verb stem. Omissions of the participle affix *-t* do only occur in those cases where the verbal stem already ends in [t].

A similar relationship between omission errors and phonological shape can be observed in the data on irregular participles. Irregular German participles often show a modification of the stem vowel and take the participle ending *-en* (e.g., *berst-geborsten*-PART 'burst'). German Broca's aphasics have been shown to frequently overgeneralise regular participle inflection to irregular verbs. In these cases, the regular participle suffix *-t* is attached to the verb's base stem (e.g. *berst-gerestet*-PART instead of *geborsten* 'burst', *pfeif-gepfeift*-PART instead of *gepfiffen* 'whistled') (cf. Penke et al. 1999). For the German aphasics under

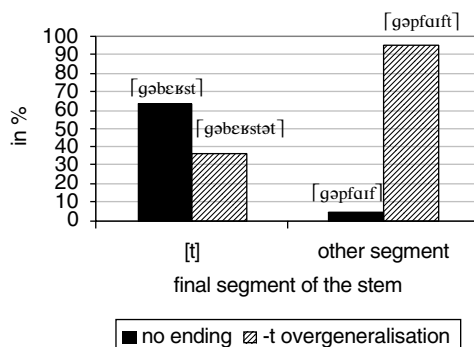


Figure 4. Results for irregular German participles—distribution of omission and overgeneralisation errors

study here, 32.7% of the participle forms for irregular verbs are incorrectly produced (=81 errors). Of these 81 errors, 97.5% (=79 cases) are either (i) overgeneralisations of the regular participle ending *-t* to the basic verb stem of an irregular verb (e.g., *geberstet* instead of *geborsten* ‘burst’, *gepfieft* instead of *gepfiffen* ‘whistled’), or (ii) errors where no affixal ending occurs on the basic verb stem. In the latter cases, the subjects produce participle forms such as *geberst* or *gepfief* instead of overregularised forms such as *geberstet* or *gepfieft*. Figure 4 illustrates the distribution of these two error types for verbs with stem on [t] (leftmost columns in Figure 4) and for verbs with other stem-final segments (columns on right-hand side of Figure 4).

The figure reveals a strong correlation between the final segment of the verb’s stem and the distribution of the two error types. Of the 19 errors that affect irregular verbs with stem final [t] 63.2% result in forms such as *geberst* where no ending is affixed to the basic verb stem. In contrast, such omission errors can only be observed for 3 of the verbs which end in other segments than [t] (=5%). The difference in the occurrence of uninflected participle forms that shows up between verbs ending in stem final [t] and those verbs ending in other segments is significant (Wilcoxon, $df = 5$, $p = .028$). Again, the distribution between *-t* overgeneralisations (i.e. *geberstet*) and *-t* omissions (i.e. *geberst*) for verbs with stem final segment [t] is compatible with a chance selection between these two forms.

We want to argue that errors such as *geberst* instead of the correct irregular participle form *geborsten* constitute errors of overgeneralisation where the regular pattern of participle formation is overapplied to build participles for irregular verbs. Figure 4 reveals that the overgeneralisation of the regular participle affix *-t* to the basic verb stem of an irregular verb is the dominant error type in the aphasic subjects’ data. Ninety-five percent of the incorrectly inflected irregular verbs that end in a stem final consonant other than [t] are inflected

with the participle affix *-t* (i.e. *gepfeift* instead of *gepfiffen*). Only in those cases where the verb's stem already ends in the segment [t] do we find a substantial number of cases where the affix *-t* does not occur in incorrectly inflected forms (i.e. *geberst*). Moreover, in these cases, the participle is not built with the correct ablauting participle stem (e.g., *borst-*), but rather with the basic verb stem (e.g., *berst-*) that is used in regular participle formation. Thus, errors such as *geberst* are overgeneralisation errors in which the confrontation of the stem final segment [t] with the (overgeneralised) affix *-t* results in the omission of the affix.

4.2. Past participles in Dutch Broca's aphasia

To compare the results for German Broca's aphasics to data from aphasic speakers of a language where schwa-epenthesis between stem final [t] and affix *-t* is not an option, data from 12 Dutch subjects with Broca's aphasia were collected.²¹ Comparable to the German experiment, the Dutch subjects were asked to transform a given 3rd plural present tense form into a participle form.²²

- (28) *Zij redden.* *Zij hebben* — ([χəreɪt])
 ('They rescue') ('They have rescued')

Five of the aphasic subjects were tested with a long version of this experiment that consisted of 67 regular verbs (40 with stem final segment [t]) and 51 irregular verbs (28 with stem final segment [t]). Seven of the aphasics could only be tested with a shorter version of this experiment, which included 30 regular verbs (20 ending with stem final [t]) and 30 irregular verbs (20 ending with stem final [t]) of the longer version. To control for frequency effects the material consisted of both frequent and infrequent participles and frequency was carefully matched within and across experimental conditions. Table 2 gives a short overview of the material of these two experiments.

Table 2. Overview on test materials

	Stem ending in [t]	Stem ending in other segment
Regular verbs	<i>plant-</i> ('plant'), <i>red-</i> ('rescue') Long version: 40/short: 20 items	<i>til-</i> ('lift') Long version: 27/short: 10 items
Irregular verbs	<i>vreet-</i> ('feed') <i>lijd-</i> ('suffer') Long version: 28/short: 20 items	<i>help-</i> ('help') Long version: 23/short: 10 items

Figure 5 presents the results for regular *-t*-inflected participles for the 12 Dutch aphasic subjects tested. The figure draws a distinction between verbs with a stem

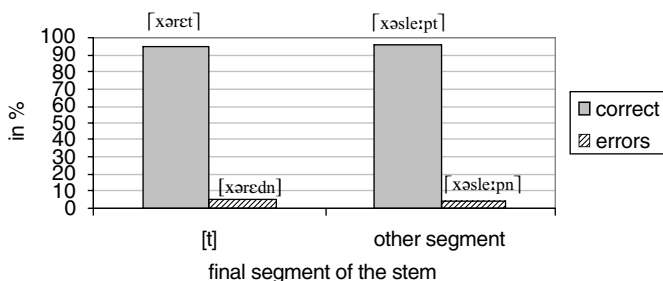


Figure 5. Results for regular Dutch participles

final segment [t] (columns on the left-hand side of the figure) and verbs with other stem-final segments (columns on the right-hand side).

Like German Broca's aphasics, Dutch aphasic subjects perform well for the production of regular inflected participles with a mean correctness score over all regular participles of 95%. In contrast to the German aphasics, however, this good performance holds for both types of verbs—verbs ending in stem final [t] and verbs ending in other segments. Statistical testing revealed no significant differences between the two verb types (Wilcoxon, $df = 11$, $p = .26$, n.s.). In total the Dutch aphasic subjects make 26 errors in the production of regular participles. All but two of these errors are irregularisations where the irregular pattern of participle formation is generalised to regular verbs (i.e. [xəredn] instead of [xəret] 'rescue', [xəsle:pn] instead of [xəsle:pt] 'drag'). None of the errors display epenthesis of Schwa between a stem final [t] and the participle affix *-t* as is required in German, i.e. participle forms such as [xəretət] do not occur in the Dutch data.

With respect to the production of irregular participles, the Dutch aphasics—like the German aphasic subjects—obtain a mean correctness score that is considerably lower than the correctness score for regular participles (mean correctness score for irregular participles is 79.3%). In total the Dutch subjects produce 86 incorrectly inflected participles for irregular verbs; 85 of these errors are overgeneralisations of the regular pattern of participle formation to irregular verbs. These overgeneralisations of the regular participle pattern result in two different error types that are dependent on the stem-final segment of the verb stem. For verbs with a stem-final [t] overgeneralisation of the regular pattern of participle inflection results in omission of the participle affix *-t* (i.e. *fluit-gefluit*-PART instead of *gefloten* 'whistle') in accordance with the Dutch phonological system. For verbs ending in other stem final segments, overgeneralisation of regular participle inflection results in affixation of *-t* to the verb stem (i.e. *spin-gezpint*-PART instead of *gesponnen* 'spin'). All overgeneralisation errors that occur with irregular verbs respect this phonological pattern: the 39 overgeneralisation errors on verbs ending in other stem-final segments than [t] are affixed with the

participle ending *-t* (i.e. *gespint*); and the 46 overgeneralisation errors of verbs with stem-final [t] result in participle forms in which the affix is omitted (i.e. *gefluit*). None of the 47 errors that in total occur with irregular verbs on stem-final [t] show epenthesis of Schwa between the stem-final [t] and the affix [t], i.e. forms such as *gefluitet* are not observed in the Dutch data.

4.3. OT account of the emergence of the unmarked in Broca's aphasia

In section 3, we concluded that the markedness constraint $*C_1\text{ə}C_1$ —which says that a schwa cannot be preceded and followed by the same consonant—outranks MAX-IO at an early stage. German (but not Dutch) children have to demote $*C_1\text{ə}C_1$ below this faithfulness constraint (see the tableau in 29). In this respect, the German ranking is more marked than the Dutch ranking.

(29) Unimpaired German

Input prefix [gə], stem [rɛt] + suffix -t

	MAX-IO	$*C_1\text{ə}C_1$
a. ☞ [gərətət]		*
b. [gərət]	*!	

According to the findings reported on in section 4.1, forms such as [gərətət] and [gərət] co-occur with about equal frequency in the production data of the German Broca's aphasics under study here. Therefore, we propose that both forms are optimal and that the constraints MAX-IO and $*C_1\text{ə}C_1$ are not hierarchically but equally ordered in these German Broca's aphasics. We thus suggest that Broca's aphasia can result in the loss of the marked constraint ranking for German (MAX-IO \gg $*C_1\text{ə}C_1$).

(30) Ranking for German aphasic subjects

Input prefix [gə], stem [rɛt] + suffix -t

	MAX-IO	$*C_1\text{ə}C_1$
a. ☞ [gərətət]		*
b. ☞ [gərət]	*	

In contrast, Dutch aphasics never vary between the two possible realisations for regular participles of verbs with stem final segment [t]: Schwa-epenthesis as in e.g. [xərətət] or *-t* omission as in the correct Dutch form [xəret]. This suggests that the Dutch unmarked ranking ($*C_1\partial C_1 \gg \text{MAX-IO}$) is not lost in Dutch Broca's aphasics. Consequently, no errors of incorrectly applied schwa epenthesis occur in the Dutch data.

- (31) *Ranking for Dutch: initial stage, unimpaired adults and aphasic subjects*
 Input prefix [xə], stem [rɛd] + suffix *-t*

	$*C_1\partial C_1$	MAX-IO
a. [xərədət]	*	
b. [☞] [xəret]		*

Summarising, we propose that the marked constraint ranking for German ($*C_1\partial C_1 \gg \text{MAX-IO}$) is lost in Broca's aphasia, whereas the Dutch unmarked ranking $*C_1\partial C_1 \gg \text{MAX-IO}$ is not lost. The data reported here indicate that agrammatic Broca's aphasia can lead to a deranking of constraints giving rise to the emergence of the unmarked in agrammatic language production.

Articulatory or morphosyntactic explanations for the occurrence of forms such as *geheft* instead of *geheftet* 'staple' in the data of the German aphasic speakers can be ruled out for the following reasons. First, all of the subjects under study here produced forms such as *geheftet* where Schwa is epenthised between stem final [t] and the participle affix *-t*. This shows that the subjects do not have articulatory problems preventing them from uttering the sequence [tət]. Second, the occurrence of forms like *geheft* cannot be attributed to the fact that agrammatic Broca's aphasics generally omit inflectional endings. Omissions of inflectional affixes have indeed long been seen as one of the core symptoms of agrammatic Broca's aphasia (cf. Grodzinsky 1990, Menn and Obler 1990). Accounts of such omission errors in Broca's aphasia mainly focus on syntactic explanations. According to these accounts, the grammatical features of functional heads which ensure correct inflectional markings are either lost (Ouhalla 1993) or not specified (Grodzinsky 1990) in phrase-structure representations. Recently, Ullman et al. (1997) have suggested that the omission of regular inflectional affixes in Broca's aphasia is due to a deficit of the rule component of the grammar. However, these and other morphosyntactic accounts predict that omissions of the participle affix *-t* occur irrespective of the verb stem's final segment. In our data, this does not hold true. As the data of our German aphasic speakers show, omissions of the participle affix *-t* only occur in those cases where the verb stem already ends in

the segment [t]. The data thus reveals a striking correlation between the final segment of the verb stem and the occurrence or non-occurrence of omission errors which cannot be captured by morphosyntactic approaches.

To summarise this section, neither purely articulatory impairments nor morphosyntactic explanations can account for the occurrence of forms such as *geheft* instead of *geheftet* in the data of German Broca's aphasics. The finding that omissions of the participle affix *-t* will only occur when the verb stem ends in the segment [t] strongly asks for a phonological explanation: forms such as *geheft* result from the deranking of the constraints MAX-IO and *C_i∅C_i as proposed in our analysis.

5. CONCLUSION

For the three German children Annalena, Eva and Naomi, we found that their very early speech is characterised by exactly one consonant and exactly one vowel, which may occur in a consonant–vowel sequence, and also in a vowel–consonant sequence and vowel length is not distinctive. At the next stage, Annalena, Eva and Naomi realise each monosyllabic word with either a long vowel, or a short vowel followed by one consonant. We concluded from this that they learned that the rhyme matters and they suppose that rhymes have minimally and maximally two positions. Under the theory assumed in this paper, this means that the constraint BRANCHINGRHYME outranks MAX-μ at this stage.

Gradually, the three German children begin to realise more than one consonant per word. We attributed this development to a stage in which they realise consonantal places of articulation more faithfully, so that, for instance, *Bahn* ‘railroad, tram’ and *warm* ‘warm’ are no longer realised by the same form (viz. as [ba:]), but distinctively (i.e. as [ban] and [bam], respectively). We attribute this development to the minimal demotion of a constraint against places of consonantal articulation in the rhyme (NoCODA).

The next step in the acquisition of German phonology is a distinction between rhymes with two positions and rhymes with three positions. We accounted for this development by assuming that at this stage, the child demotes the constraint BRANCHINGRHYME below MAX-μ. Thus, the first steps in the acquisition of German consonants are best accounted for by minimal constraint demotion.

At the stage when children start to realise three-positional rhymes, they also start to realise the inflectional ending *-t* in coda positions. The wish to realise the inflectional ending, sometimes leads to vowel-length reduction or deletion of the stem-final consonant in coda position (e.g., /ma:l/ + /t/ → [malt] or [ma:t] ‘paints’). When the German child begins to realise a coronal appendix, vowel-length reduction or deletion of stem-final consonants no longer take place (/ma:l/ + /t/ → [ma:lt]). At this acquisitional stage, the only context in which a child has problems realising the ending *-t* is when the stem ends in the sound [t].

According to Bernhardt and Stemberger (1998:642), English speaking children face similar problems with the realisation of past tense forms for verbal stems ending in /t/ or /d/ and with the realisation of plural forms for nominal stems ending in /f/, /s/ or /z/. We argued that these problems should be attributed to highly ranked universal markedness constraints which prohibit a sequence of two identical segments (OCP) or the occurrence of schwa between two identical segments (*C_iəC_i).

The avoidance of a sequence of consonant-schwa-identical consonant (*C_iəC_i in our terminology) is not only a prominent feature in first language acquisition, but also in language impairment. We showed that the German Broca's aphasics do not omit inflectional endings randomly. Rather, the omission of the inflectional ending *-t* in aphasic speech heavily depends on the final segment of the stem: participles of stems ending in [t] are more prone to omission of the inflectional ending *-t* than participles of stems whose final segment is not a coronal stop. To capture these errors, we suggested that in aphasic speech, the ranking of constraints no longer decides whether the wish to faithfully realise the inflectional ending *-t* is—or is not—preferred to the wish to realise unmarked structures (in this case to avoid a schwa preceded and followed by an identical segment). Whereas in child language, the emergence of the unmarked can be attributed to the fact that markedness constraints outrank faithfulness constraints, we proposed that in agrammatic Broca's aphasia, the emergence of the unmarked is due to a deranking of constraints. Our proposal offers a new and interesting way to study and capture aphasic language disorders. It would, for instance, be interesting to see whether English aphasics skip the inflectional ending *-s* more often after stems ending in a strident sound than after a non-strident sound and whether Dutch aphasics realise the plural ending *-en* more often after stems ending in a non-nasal sounds than after stems ending in a nasal. We leave the issue whether other occurrences of erroneous forms in language impaired aphasic speech can be attributed to the loss of a once fixed constraint ranking to future research.

NOTES

* Previous versions of this paper were presented at the 11th Manchester Phonology Meeting, May 2003 and the 11th International Morphology Meeting, Vienna, February 2004. We are grateful to Roelien Bastiaanse, Peter Hagoort, Claudia Hegenscheidt, Ulrike Janssen, Sandra Joppen-Hellwig, Hermann Kolk, Eva Neuhaus and Esther Ruijgendijk for their help in collecting the data. Michaela Zander helped us to store Naomi's data in computer files. We thank Annemarie Kerkhoff, Joseph Stemberger, Dieter Wunderlich and an anonymous reviewer for helpful comments. The research presented in this paper is supported by German Science Foundation Grants SFB 282/C10 and C8 to the respective authors.

¹ One anonymous reviewer suggests that adults and children do not have problems adding the plural ending *-en* to a stem that ends in an alveolar nasal (e.g. Dutch *banaan* ~ *bananen*

‘banana ~ bananas’). This would suggest that a constraint $*n\bar{n}$ is not active in the languages under discussion. We tentatively attribute this to the fact that the constraint $*C_i\bar{C}_i$ takes into account place of articulation features, which—judging from the fact that /n/ is the least marked consonant and functions as the default consonant in Dutch (Booij 1998)—/n/ does have. Also consider the fact that Dutch has two plural suffixes: *-en* and *-s*. The choice between them depends on a prosodic and a segmental factor: (i) stems that end in a stressed syllable prefer the ending *-en* (resulting in a final trochee, cf. van Haeringen 1947, Booij 1998 and others) and (ii) stems that end in a short vowel followed by a sonorant usually select *-s* (de Haas and Trommelen 1993). In a case study with 40 participants, two pseudowords that end in a stressed syllable with a final short vowel followed by a liquid were realised with the plural suffix *-en* in 85% of the cases and pronounced with a stress shift or the ending *-s* in the remaining 15% of the cases. In contrast, six nonsense words that end in a stressed syllable with a final nasal consonant were realised with the plural suffix *-en* in only 55% of the cases (van Ginkel 2005). We here suggest that it is plausible that this significant difference is due to the effect of the $*n\bar{n}$ constraint.

² Details on the child’s age given as year;month.day.

³ Apparent counterexamples are syllables headed by schwa (e.g., the final syllable in *Sprache* [ˈʃpʁaːχə] ‘language’) and syllables headed by sonorant consonants (e.g., the final syllables in *sprechen* [ˈʃpʁɛçŋ] ‘to speak’ and *Atem* [aːtm̩] ‘breath’). Such syllables do not require an onset, they are never stressed, and they never form a word on their own, i.e. they do not constitute words.

⁴ Note that in the phonologies of German and Dutch, it is generally assumed that plosives and fricatives have the same level of sonority (e.g. Wiese 1996 for German). In section 3 we will assume without further discussion that elements in the rhyme contribute to the syllable’s weight and are moras.

⁵ In (4) and below, ‘V’ = vowel, ‘Son’ = sonorant (vowel, glide, liquid or nasal) and ‘C’ = consonant. The symbol ‘σ’ denotes a syllable. Hall (2002) argues against extrasyllabic consonants in German and English. According to him, German allows complex codas such as final *pt*, *kt*, *ps*, *ks*, *ft*, *çt*. Sequences of a coronal stop followed by a non-coronal stop are ruled out by markedness constraints. Even though the details of the analysis presented here differ from his analysis, nothing crucial hinges on this.

⁶ We do not find initial glottal stops in vowel-initial words at this stage in German child speech.

⁷ Omission of a fricative in onset position and realisation of a post-vocalic consonant in early child speech has also been observed in, for instance, Portuguese (Costa and Freitas 1998), Dutch (Fikkert 1994a,b) and English (Velten 1943, Menn 1971).

⁸ For Naomi, there is still some variation at 1;4.26 and 1;5.01 and some words are realised with a final short vowel. For instance, at 1;5.01 *Hund* [hʊnt] ‘dog’ is realised once as [hu], once as [hu:] and once as [hʊtʰ] and *zu* [tsu:] is realised twice with a short vowel and once with a long one. From 1;5.08 onwards, we no longer find words which end in a short vowel.

⁹ Naomi regularly produces a fricated release after word-final /t/.

¹⁰ Note that children produce the suffix *-t* to indicate finiteness in early stages of acquisition. Once the German subject–verb agreement system has been acquired, the suffix expresses 3rd person singular.

¹¹ A short vowel cannot be followed by two obstruents at this stage, because the appendix is not available yet.

¹² German syllable-final obstruents are devoiced. Hence, a stem-final obstruent as in *find* ‘find’ is realised with voicing in onset position (e.g., when followed by a vowel in the inflected form *fin.det* ‘find-3.SG’), whereas it is realised without voicing in syllable-final position (e.g., *findling* [fɪnt.lɪŋ] ‘foundling’).

¹³ According to Wunderlich (1992) and Wunderlich and Fabri (1995), *-t* inflected participles are productively built by a process of affixation that combines the verb stem with the participle suffix *-t*, whereas, the *-n* marking of strong participles is memorised together with information on the particular change in the stem vowel. Both *-t* and *-n* are available for participle formation in German child language. There is, however, a difference in the productivity of the two endings. In her analysis of longitudinal data of nine German children aged between 1 and 3 years, Weyerts (1997) found that in 93% of all occurring errors the participle affix *-t* is overapplied to build a participle of a strong verb (e.g. *gesingt* instead of *gesungen* ‘sung’). In contrast, irregularisations, where the *-n*-ending is overapplied to a weak verb (e.g. *gelachen* instead of *gelacht* ‘laughed’) are rare (7% of the occurring errors); for an analysis of a subset of this data see also Clahsen and Rothweiler (1993). As we show in section 3, German children produce such irregularisations to avoid affixation of the participle affix *-t* to weak verbs that end in a coronal stop.

¹⁴ Bernhardt and Stemberger (1998:642) make a similar observation for English speaking children, i.e. these children have difficulties with the realisation of past tense forms when the stem ends in a /t/ or /d/ and with plural and present tense forms when the stem ends in /s/ or /z/. According to them, English-speaking children—just like the German children in this study—tend to use the stem form rather than inserting schwa between the stem and the inflectional ending.

¹⁵ As noted above, irregularisations in the participle formation of weak verbs are rare in German child language. It is therefore an interesting observation that Eva uses the *-n* form instead of [-ɔ̃]-inflection in the formation of participles for verbs with a stem final coronal stop (see 17a). Her irregularisation rate for verbs not ending in a stem final coronal stop is only 1.9% (5 of 263). In contrast, her irregularisation rate for weak verbs that end in a stem final coronal stop is 14.3% (1 of 7). This suggests that irregularised *-n*-marked participle forms are used to avoid *-t*-inflection of verbs already ending in a coronal stop. This observation seems to be in line with the data published by Weyerts (1997) and Clahsen and Rothweiler (1993). Note that one instance of irregularisation of *-en* in their corpus is found for the verb *schlachten* ‘to slaughter’ whose stem ends in [t]; *geschlachten* instead of *geschlachtet*.

¹⁶ Note that this form is an overregularisation where the regular participle affix *-t* is overapplied to the basic verb stem of the strong verb *reit*- ‘cut’ (correct participle *geritten*).

¹⁷ Note that this form is an overregularisation where the regular participle affix *-t* is overapplied to the basic verb stem of the strong verb *schneid*- ‘cut’ (correct participle *geschnitten*).

¹⁸ Note that this phenomenon cannot be accounted for by reference to DEP-IO, because other evidence suggests that this constraint is ranked lowly at this stage.

¹⁹ The constraint *C_iɔ̃C_i refers to identical manner and place of articulation; it does not take voicing differences into consideration.

²⁰ Based on data from 5 English speaking Broca’s aphasics who showed a deficit with regular past tense inflection, Ullman et al. (1997) propose that regular inflection is subserved by Broca’s area in the brain, i.e. a brain region that is regularly lesioned in Broca’s aphasia. The data on participle formation in German and Dutch Broca’s aphasics presented here (see also Penke et al. 1999, Penke and Westermann submitted for publication), however, provides no evidence

for the claim that regular inflection is selectively impaired in Broca's aphasia. For a discussion of Ullman's proposal see Penke (2002), Penke and Westermann (submitted for publication).

²¹ We thank Claudia Hegenscheidt for constructing the material. We are also very grateful to Esther Ruijgendijk, Herman Kolk, Peter Hagoort and Roelien Bastiaanse for running the experiment with their Dutch aphasic subjects.

²² For the German experiment we presented 1st person singular forms because these forms are neither inflected with *-t* nor with *-n*. However, 1st person singular contexts seemed inadequate for the Dutch experiment, since the singular forms resemble the stem form in Dutch. As we were interested in the occurrence of such forms in participle formation and wanted to avoid priming for these forms, we decided to use 3rd person plural contexts instead.

REFERENCES

- Bernhardt, Barbara Handford and Joseph P. Stemberger (1998). *Handbook of Phonological Development from the Perspective of Constraint-Based Non-Linear Phonology*. San Diego: Academic Press.
- Blumstein, S. E. (1973). *A Phonological Investigation of Aphasic Speech*. The Hague: Mouton.
- Booij, Geert (1998). Phonological output constraints in morphology. In Kehrein, Wolfgang and Richard Wiese (eds.), *The Phonology and Morphology of Germanic Languages*. Tübingen: Niemeyer, 143–164.
- Booij, Geert (2002). Prosodic constraints on affixation in Dutch. In *Yearbook of Morphology 2001*. Kluwer: Dordrecht.
- Clahsen, Harald (1999). Lexical entries and rules of language: A multi-disciplinary study of German inflection. *Behavioral and Brain Sciences* 22, 991–1060.
- Clahsen, Harald and Martina Penke (1992). The acquisition of agreement morphology and its syntactic consequences: New evidence on German child language from the Simone-Corpus. In J. Meisel (ed.), *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Acquisition*. Kluwer: Dordrecht, 181–223.
- Clahsen, Harald, Martina Penke and Theresa Parodi (1994). Functional categories in early child German. *Language Acquisition* 3(4), 395–429.
- Clahsen, Harald and Monika Rothweiler (1993). Inflectional rules in children's grammars: Evidence from the development of participles in German. *Yearbook of Morphology 1992*. Kluwer: Dordrecht, 1–34.
- Costa, João and M. João Freitas (1998). V and CV as Unmarked Structures: Evidence from the Acquisition of Portuguese. Paper presented at the conference on *The Syllable: Typology and Theory*, Tübingen, July 1998.
- de Haas, Wim and Mieke Trommelen (1993). *Morfologisch Handboek van het Nederlands*. 's Gravenhage: SDU Uitgeverij.
- Dressler, Wolfgang U. (1991). The sociolinguistic and patholinguistic attrition of Breton Phonology, Morphology, and Morphophonology. In H. Seliger and R. Vago (eds.), *First Language Attrition*. Cambridge: CUP, 99–112.

- Elsen, Hilke (1991). *Erstspracherwerb: der Erwerb des deutschen Lautsystems*. Wiesbaden: Deutscher Universitäts Verlag.
- Fikkert, Paula (1994a). *On the Acquisition of Prosodic Structure*. PhD dissertation, Leiden: HIL, Leiden University.
- Fikkert, Paula (1994b). On the acquisition of rhyme structure in Dutch. In Bok-Bennema, Reineke and Crit Cremers (eds.), *Linguistics in the Netherlands*. Amsterdam: John Benjamins, 37–48.
- Gnanadesika, Amalia E. (1996). The Acquisition of Phonology as the Promotion of Faithfulness Constraints: A case study in onset selection. Paper presented at WCHSALT, OTS, Utrecht University, June 1996.
- Goldsmith, John (1976). *Autosegmental Phonology*. PhD dissertation, Cambridge, MA. New York: Garland Press, 1979.
- Grijzenhout, Janet (2001). *West-Germanic Sound Structure: A Contrastive Study*, Habilitationsschrift, Heinrich-Heine-Universität Düsseldorf. Amsterdam: John Benjamins.
- Grijzenhout, Janet and Sandra Joppen (1998a). German Child Language and Phonological Features: A Case Study. Paper presented at the Third Utrecht Biannual Phonology Workshop, Utrecht, the Netherlands, 11 June 1998.
- Grijzenhout, Janet and Sandra Joppen (1998b). First Steps in the Acquisition of German Phonology: A Case Study. *Theory des Lexikons; Arbeiten des Sonderforschungsbereichs 282*, Nr. 110.
- Grijzenhout, Janet and Sandra Joppen-Hellwig (2002). The lack of onsets in German child phonology. In I. Lasser (ed.), *The Process of Language Acquisition*. Frankfurt am Main: Peter Lang Verlag, 319–339.
- Grodzinsky, Yosef (1990). *Theoretical Perspectives on Language Deficits*. London: MIT Press.
- Hall, T. Alan (2002). Against Extrasyllabic consonants in German and English. *Phonology* 19, 33–75.
- Huber, W., K. Poeck, D. Weniger and K. Willmes (1983). *Der Aachener Aphasie Test*. Göttingen: Hogrefe.
- Jakobson, Roman (1941/1971). *Kindersprache, Aphasie und allgemeine Lautgesetze*. Frankfurt a.M.: Suhrkamp. Also published as *Kindersprache, Aphasie und allgemeine Lautgesetze. Roman Jakobson: Selected Writings 1: Phonological Studies*. The Hague: Mouton.
- Janssen, Ulrike and Martina Penke (2002). Phonologically conditioned omissions of inflectional affixes in German Broca's aphasia. *Brain and Language* 83, 99–101.
- Jespersen, Otto (1904). *Lehrbuch der Phonetik*. Leipzig: Tuebner.
- Leben, William (1973). *Suprasegmental phonology*. PhD dissertation, Cambridge, MA.
- McCarthy, John (1986). OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17, 207–263.
- Menn, Lise (1971). Phonotactic rules in beginning speech: A study in the development of English discourse. *Lingua* 26, 225–251.
- Menn, Lise and L. K. Obler (1990). Cross-Language data and theories of agrammatism. In Menn Lise and L.K. Obler (eds.), *Agrammatic Aphasia: A Cross-Language Narrative Sourcebook. Vol. II*. Amsterdam: Benjamins, 1369–1388.

- Ouhalla, John (1993). Functional categories, agrammatism and language acquisition. *Linguistische Berichte* 143, 3–36.
- Penke, Martina (2002). *Flexion im Mentalen Lexikon: Eine neuro- und psycholinguistische Perspektive*. Postdoctoral thesis, University of Duesseldorf. Retrieved from http://www.phil-fak.uni-duesseldorf.de/sfb282/C8/Habil_Martina_Penke.pdf. Tübingen: Niemeyer.
- Penke, Martina and Gert Westermann (submitted for publication). Broca's area and inflectional morphology: New evidence from Broca's aphasia and computer modelling. *Cortex*.
- Penke, Martina, Ulrike Janssen and Marion Krause (1999). The representation of inflectional morphology: Evidence from Broca's aphasia. *Brain and Language* 68, 225–232.
- Sherer, Tobias (1994). *Prosodic phonotactics*, PhD dissertation, University of Massachusetts, Amherst.
- Smith, Neil V. (1973). *The Acquisition of Phonology, A Case Study*. Cambridge, England: Cambridge University Press.
- Smolensky, Paul (1999). Grammar-based connectionist approaches to language. *Cognitive Science* 23(4), 589–613.
- Tesar, Bruce and Paul Smolensky (1993). The learnability of Optimality Theory: An algorithm and some basic complexity results, Technical Report CU-CS-678-93, University of Colorado, Boulder.
- Ullman, M., S. Corkin, M. Coppola, G. Hickok, J. Growdon, W. Koroshetz and S. Pinker (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience* 9(2), 266–276.
- van Ginkel, Babbe (2005). Default? Onderzoek naar de defaultstatus van de meervoudsuffixen *-en* en *-s*. BA thesis, Utrecht University.
- van Haeringen, C. B. (1947). De meervoudsvorming van het Nederlands. *Mededelingen der Koninklijke Nederlandse Academi van Wetenschappen, Afdeling Letterkunde, Nieuwe Reeks, Deel 10, No. 5*.
- Velten, H. (1943). The growth of phonemic and lexical patterns in infant speech. *Language* 19, 281–292.
- Vennemann, Theo (1988). *Preference Laws for Syllable Structure and the Explanation of Sound Change*. Berlin: Mouton.
- Waugh, L.R. and B.A. Lafford (2000). Markedness. In Geert Booij, C. Lehmann, J. Mugdan, W. Kesselheim and S. Skopetas (eds.), *Morphology: An International Handbook on Inflection and Word-Formation. Vol. 1*. Berlin; de Gruyter, 272–281.
- Weyerts, Helga (1997). *Reguläre und Irreguläre Flexion: Psycholinguistische und Neurophysiologische Ergebnisse zu Erwerb, Verarbeitung und Mentaler Repräsentation*. Dissertation University of Duesseldorf.
- Wiese, Richard (1996). *The Phonology of German*. Oxford: Clarendon Press.
- Wunderlich, Dieter (1992). A minimal analysis of German verb morphology. *Theory des Lexikons; Arbeiten des Sonderforschungsbereichs* 282. Heinrich-Heine-Universität Düsseldorf.
- Wunderlich, Dieter and Ray Fabri (1995). Minimalist Morphology: An approach to inflection. *Zeitschrift fuer Sprachwissenschaft* 14, 236–294.

Yip, Moira (1998). Identity avoidance in phonology and morphology. In S. LaPointe, D. Brentari and P. Farrell (eds.), *Morphology and its Relation to Phonology and Syntax*. Stanford, CA: CSLI Publications, 216–246.

(Janet Grijzenhout)
English Language Studies
Utrecht University
Trans 10
3512 JK Utrecht, The Netherlands

janet.grijzenhout@let.uu.nl

(Martina Penke)
Institut für Sprache und Information
Heinrich-Heine-Universität Düsseldorf
Universitaetsstr. 1
40225 Düsseldorf, Germany

martina.penke@phil-fak.uni-duesseldorf.de

On the typology of state/change of state alternations*

ANDREW KOONTZ-GARBODEN

1. INTRODUCTION

Words denoting non-causative and causative change of state (COS) predicates often are morphologically related to words denoting the related state predicates, though the relationship sometimes differs for different types of states. For the state of ‘brokenness’, for example, in English the word denoting the state in (1c) is derived from the word denoting the change of state. In contrast, the word denoting the state of ‘looseness’ in (2c) is morphologically basic, with the words denoting the changes of state being derived from it.¹

- (1) a. The glass broke. (non-causative change of state)
- b. Sandy broke the glass. (causative change of state)
- c. The glass is broken. (state predicate is deverbal)

- (2) a. The knot loosened. (non-causative change of state)
- b. Sandy loosened the knot. (causative change of state)
- c. The knot is loose. (state predicate is morphologically basic adjective)

The morphological typology of words denoting non-causative (e.g. (1a), (2a)) and causative (e.g. (1b), (2b)) COS predicates has been relatively well studied (Nedjalkov and Silnitsky 1973; Haspelmath 1993), with one important finding being that for certain types of COS events, languages tend to have morphologically basic words denoting the causative predicates, morphologically deriving the corresponding word denoting the non-causative COS predicate. For other types of events, the opposite direction of derivation is favored. This pattern of behavior is observed in Tongan (Polynesian), as shown in (3) and (4).

- (3) Tongan
 - a. pelu ‘cause become bent’ (causative change of state)
 - b. ma-pelu ‘become bent’ (non-causative change of state)

- (4) Tongan
 - a. lahi ‘become big’ (non-causative change of state)
 - b. faka-lahi ‘cause become big’ (causative change of state)

Certain types of events are lexicalized with the causative as the morphologically basic form, deriving the word denoting the non-causative change of state, as seen in (3) for the word for ‘bend’. Other events, in contrast, have the non-causative change of state lexicalized as the morphologically basic form, deriving the word denoting the causative change of state as in (4) for the word for ‘big’. Haspelmath (1993) argues that the direction of morphological derivation correlates with the likelihood that the event can occur spontaneously—events more likely to occur spontaneously are lexicalized in their morphologically basic form as words denoting non-causative COS predicates (e.g. *melt*), while those less likely to occur spontaneously are lexicalized in their morphologically basic form as words denoting causatives (e.g. *break*). The leading idea behind his research program is that the morphological direction of derivation, within and across languages, is suggestive of how non-causative and causative COS predicates are conceptually related to one another.

I take Nedjalkov and Silnitsky’s and Haspelmath’s ideas further by examining how the non-causative and causative COS predicates they are interested in are related to their associated states. Specifically, for a given state such as ‘broken’ or ‘wide’, there has been no systematic investigation of the morphological relationship between words denoting the state, a non-causative change into the state, and a causative change into the state. In this paper I take the first steps in such an investigation, considering not only whether the meaning of all states is comparable, but also the well-known fact that languages lexicalize states as members of different lexical categories. While some languages have a large open class of adjectives, others do not, instead lexicalizing notions that turn up as adjectives in other languages, as nouns or verbs (Dixon 1982). This point of cross-linguistic variation turns out to have a significant impact on other areas of grammar (cf. Lehmann 1990), as is seen below.

The study yields two principal empirical findings. First, as noted previously by Dixon (1982), not all states are the same. States entailing some prior event giving rise to the state, result states, are derived in many languages from the verb denoting that event as in *break/broken*. States that do not presuppose such a change, for example *red*, are never derived from the corresponding change of state verb. Instead, the COS verb is derived from the word denoting the state (e.g., *red*), the opposite direction of derivation as that for predicates like *break/broken* (cf. (1)–(2)). These facts follow, I argue, from a principle of monotonic composition, whereby meaning is added in the construction of word meaning, but never deleted. The second major empirical finding of the study is that the derivational relationship between words denoting states and words denoting the associated change of state is affected by the lexical category of the associated state. In particular, I find that it is only in languages where states are lexicalized as verbs that the same word is polysemous between a state and a non-causative change of state meaning. I argue that this follows from a

constraint on the type of lexical category that can denote changes of state—only verbs may do this. For there to be a polysemy between state and change of state, the state must be lexicalized as a verb, otherwise this principle would be violated.

I begin by laying out what I believe to be some of the more important empirical questions in this domain. I follow this with discussion of some suggestive data culled from reference grammars and native speakers of relevant languages, and then move on to the explanation of the observed typological facts. I then discuss some potential challenges for the analysis and conclude by discussing the implications of the findings for theories of event structure representation.

2. TWO QUESTIONS ABOUT CHANGE OF STATE ENCODING

The question of how words denoting states are related to their non-causative and causative COS counterparts is prefigured in the work of Hale and Keyser (2002) and Baker (2003). Theories such as theirs predict a very specific type of relationship between states and their causative and non-causative COS counterparts, namely, causative and non-causative COS predicates are predicted to be derived from their state counterparts.

Hale and Keyser especially, give suggestive data supporting the idea that words denoting non-causative and causative COS predicates are morphologically derived from words denoting the corresponding state.

(5) O'odham (Hale and Keyser 1998:92, (31))

- a. (s-)moik 'be soft'
- b. moik-a 'become soft'
- c. moik-a-(ji)d 'cause to become soft'

(6) Warlpiri (Hale and Keyser 1998:92, (31))

- a. wiri 'be big'
- b. wiri-jarri- 'become big'
- c. wiri-ma- 'cause to become big'

In the O'odham data in (5) the word denoting the causative is derived from the word denoting the non-causative, which is in turn derived from the word denoting the state. In the Warlpiri data in (6), on the other hand, the words

denoting the causative and the non-causative COS predicates are derived from the word denoting the state. In both cases the state is morphologically basic, an observation Hale and Keyser use to argue for the derivation of the changes of state from the state itself. Though it is clear that this sort of relationship holds in many instances, the results of Nedjalkov and Silnitsky (1973) and of Haspelmath (1993) suggest that it should not be taken for granted that the relationship will be identical for all types of states and all types of languages. There are two main issues to be considered. First, Dixon (1982) highlights a distinction between two kinds of states; it could be that the relationship between words denoting states and words denoting changes of state is different for these two different types of states. Secondly, Dixon (1982) also shows that languages differ in the lexical category assigned to states. Both of these issues, I argue in the following sections, have an impact on how words denoting states and words denoting changes of state are morphologically related to one another. Before discussing the data illustrating these points, however, I consider Dixon's observations.

2.1. *Two kinds of states*

In contrast to what is suggested by the theories of Hale and Keyser (2002) and Baker (2003), Dixon shows that "... certain states, naturally described by adjectives, contrast with states that are the result of some action" (1982:50). Dixon refers to the class of states naturally described by adjectives—in languages that have that lexical category—as *property concepts* (e.g. predicates denoting states related to speed, age, dimension, color, value, etc. and that presuppose no prior event). Contrasting with the class of property concepts is the class of states "that are the result of some action," *result states*. These two classes of states differ from one another in fundamental ways. The defining distinction between them is in their entailments—while result states entail that there was an event giving rise to the resulting state, this is not the case for property concept states. This contrast is illustrated by the data in (7) and (8).

- (7) a. #The glass is broken, but it never broke.
 b. #Kim prefers his barbecued chicken uncooked.
 c. #Sandy is dressed, but neither she nor anyone else dressed her.
- (8) a. The dirt is red, but nobody reddened it.
 b. Mount Chimborazo is tall and has always been so.
 c. The rainy season is and has always been a bad time to plan a picnic.

The data in (7) show that there must be a prior event leading up to the states of *broken*, *barbecued*, and *dressed*. When the entailment that there was a prior event leading up to the result state is contradicted, the sentence is quite odd (contradictory). For example, (7a) shows that if something is *broken*, it must be the case that that something underwent a breaking event. To deny that there was a prior breaking event leads to contradiction. Likewise for *barbecued*—meat cannot be inherently barbecued. As shown by the data in (7b) it must undergo some sort of cooking process (i.e., barbecuing) in order to reach the result state *barbecued*. The same holds for being *dressed*—one cannot be dressed without putting on clothes, as shown by (7c). These facts contrast with the facts for property concept states, as illustrated by the data in (8). In naive physics, at least, there is no sense in which red dirt becomes red (8a), a mountain must become tall (8b), or a time of year has to become bad (8c). Because these states do not entail a prior event, sentences that deny that there was such an event leading up to the state, like those in (8), are not at all contradictory.

The contrast in entailment behavior between property concept states and result states is also illustrated quite clearly in English by adjectives, which denote property concepts, and their corresponding deverbal adjectives.

- (9) a. Look at the bright picture on your left. (=camera took a bright picture)
 b. Look at the brightened picture on your left. (=camera took a bad picture, brightened with, e.g. software)
- (10) a. Kim ate a red apple.
 b. Kim ate a reddened apple.
- (11) a. Sandy's shirt has long sleeves.
 b. Sandy's shirt has lengthened sleeves.

The data in (9)–(11) illustrate a minimal contrast between property concept states and result states. While the picture in (9a) is inherently *bright*, the picture in (9b) had to undergo some sort of brightening process in order to reach the result state *brightened*. Things are similar for the apple in (10)—whereas in (10a), the apple is just inherently red, the apple in (10b) had to become (more) red via some sort of reddening process. The data in (11) illustrate a similar contrast. While the sleeves of Sandy's shirt are just plain long in (11a), they had to become long via some sort of lengthening process in (11b).

In the following sections, I show that this semantic difference between two types of states is reflected morphologically as well—while result states are often derived from a verb denoting the event leading to the result state (cf. *bright* versus *brightened*), words denoting property concepts in all languages I have

examined, with one possible exception I discuss in section 5, are morphologically basic, whether lexicalized as nouns, verbs, or as adjectives.

2.2. Lexical category encoding of states varies crosslinguistically

An additional relevant question in this domain of study is what effect a language's lexical category inventory has on the relationship between words denoting states and words denoting their associated changes of state. It is well known that not all languages have adjectives. Property concepts show up as nouns in some of these languages, and as verbs in others (Dixon 1982). Given that derivational morphology is often sensitive to lexical categoryhood, it seems quite possible that crosslinguistic variation in lexical category inventory might contribute to different types of relationships between words denoting states and their related changes of state. This is a point that has not thus far been seriously considered in the literature (though see Hale and Keyser (1998) for suggestions that variation in lexical category encoding may be a source of crosslinguistic variation). In section 3, I show that indeed, variation in the lexical category of states across languages correlates with different types of relationships between words denoting states and words denoting changes of state.

3. SOME SUGGESTIVE DATA

Having laid out these questions regarding the relationship between states and changes of state, I turn to some preliminary data suggesting answers and further areas for research. I begin by addressing the question in section 2.1 and then move on to the question in section 2.2. I follow the empirical discussion here with theoretical analysis in section 4.

3.1. Are all states conceptually and morphologically basic?

The lexical semantic considerations discussed above already show, at least on the basis of semantic intuitions from English, that not all states are conceptually basic—while some are (the so-called “property concepts”) others are not (so-called “result states”). Data from a variety of languages, such as English, Quechua, Eastern Armenian, and Tongan suggest that this conceptual difference correlates with a morphological difference.

3.1.1. English

As already suggested by the data discussed above, in English words denoting property concepts are morphologically basic in their stative denotation, with

words denoting the related changes of state being derived. Words denoting result states, in contrast, are often derived from the word denoting the related change of state. The data in (12) and (13) illustrate this point.

Words whose denotation includes a property concept are morphologically basic in their stative denotation, as shown in (12) for *loose*, while the words denoting the changes of state are derived from the word denoting the property concept state with the *-en* suffix.

- (12) a. The knot is loose.
 b. The knot loosened.
 c. Kim loosened the knot.

The same sort of relationship between states and changes of state holds for other adjectives in English, such as *bright, broad, cheap, coarse, damp, dark, deep, fat, flat, fresh* and others (Levin 1993). In other instances, the word denoting the change of state and the associated state are morphologically identical (e.g. *clear*), but I assume that the COS predicates are again derived, as represented by the category change. I attribute the absence of the affix to a failure to meet the phonological conditions governing its appearance, an observation dating at least to Jespersen (1939).²

This contrasts with the situation for words whose denotation includes a result state—for these types of words in English, the word denoting the state tends to be the one that forms English past participles, derived with the *-en* suffix (and its allomorphs) from the word denoting the change of state, as shown in (13).

- (13) a. The glass is broken.
 b. The glass broke.
 c. Alex broke the glass.

The same sort of relationship holds for other verbs denoting an action giving rise to a result state, such as *bend, break, crease, crinkle, crumple, fold, rumple, wrinkle, break, chip, crash, crush, fracture, rip, shatter, smash, snap, splinter, split, tear*, and others (Levin 1993).

Additionally, as Dixon (1982:50ff.) observes, a contrast can be observed in English between states with “action oppositions.” The observation is that there is a subclass of property concept states for which the antonym is a result state. For example, in order for something to no longer be *raw*, it must be *cooked*, i.e., undergo some cooking event. Some other examples of property concept/result state antonyms are given in (14).

- (14) Some basic state/result state opposites
- | | |
|--------------|---|
| raw | cooked |
| whole | broken, split, torn, crushed, snapped, smashed etc. |
| same/similar | changed |
| clear | clogged (e.g. pipes) |
| brilliant | faded |
| live | detonated (e.g. bomb) |
| nude, naked | dressed |
| steady | increasing (e.g. prices) |
| solid | melted |

The data in (14) show that in each case when there is a property concept/result state antonym pair, while the word denoting the property concept is morphologically basic, the word denoting the result state (negation of the property concept state) is derived. Similarly, as was already seen in (9)–(11), there is a morphological contrast correlating with a lexical semantic contrast between property concepts and deverbal adjectival forms. For example, it was seen that while *bright* in (9a) does not entail a prior event leading to the state, *brightened* in (9b) does. The differences in entailments are reflected in the morphological shapes of the words—the word denoting the property concept *bright* is morphologically basic while the word denoting the result state *brightened* is deverbal (i.e., derived from a deadjectival verb).

3.1.2. Cuzco Quechua

This asymmetry between property concepts and result states is observed in other languages as well. In Quechua for instance, words whose denotation includes a property concept have a morphologically underived form that denotes a state. This is illustrated by the data in (15) from the Cuzco dialect.

- (15) a. wasi-qa *hatun*-mi (ka-sha-n)
 house-TOP big-EVIDENTIAL be-PROG-3P
 ‘The house is big’ (Martina Faller, p.c.)
- b. *hatun*-ya-y
 big-TRANSFORMATIVE-INF
 ‘become big’ (agrandarse) (Cusihuaman 1976:195)
- c. wasi-ta *hatun*-ya-*chi*-rqa-n
 house-ACC big-TRANSFORMATIVE-CAUS-PAST-3P
 ‘(s)he made the house big.’ (Martina Faller, p.c.)

The data in (15) show that the word for ‘big’ in Quechua is morphologically basic in its stative denotation. The word denoting the associated non-causative change of state is then derived from the word denoting the state, as shown in

(15b). The word denoting the causative change of state for its part, is derived from the word denoting the non-causative change of state via additional affixation, as is illustrated by the data in (15c). Other words denoting property concepts seem to pattern similarly. According to Weber, describing the related Huallaga dialect, *-yā* is an inchoative marker and "... seems to be completely productive ..." occurring with property concept words with meanings such as 'big', 'crazy', 'white', 'rich', 'red', 'sickness/sick person', 'curly', 'hard', 'deaf', etc. (Weber 1989:30–31). Words denoting causative changes of state can then be derived from the *-yā* marked non-causative changes of state with the *-chi* causative suffix (Weber 1989:166; Cusihuaman 1976:194), Martina Faller, p.c.; compare (15b)–(15c)).

This direction of derivation from state to non-causative change of state to causative change of state contrasts with the direction of derivation for states that entail an event giving rise to a result state. This is illustrated by the data in (16).

- (16) a. *Tela qhasu-sqa ka-sha-n.*
 cloth tear-PAST.PART be-PROG-3P
 'The cloth is torn.' (Martina Faller, p.c.)
- b. *tela qhasu-ku-n.*
 cloth tear-REFL.-3P
 'The shirt tore/got torn.' (Martina Faller, p.c.)
- c. *tela-ta qhasu-sha-n.*
 cloth-ACC tear-PROG-3P
 'She/he tore the shirt./She tears/is tearing the cloth.'
 (Martina Faller, p.c.)

In these cases, the word denoting the state is a participle derived from a verb (Weber 1989:282–283; Cusihuaman 1976:225), as illustrated by the data in (16a). The word denoting the non-causative change of state, for its part, is derived from the word denoting the causative change of state with the reflexive marker *-ku*, as seen in (16b)–(16c).

3.1.3. Eastern Armenian

Megerdooian (2002:96) observes the same sort of contrast in Eastern Armenian that I have documented in other languages. In this language, there is a class of change of state verbs derived from morphologically basic adjectives. This class of change of state verbs contrasts with another class for which there exists no corresponding morphologically basic adjective. The first class is what she calls the Category 1 class, which consists of words with canonical property concept meanings. Words with these meanings are morphologically basic in their stative denotation. Words with the corresponding COS meaning are then derived from the word denoting the property concept state; while the morpheme *-anal* marks

the non-causative COS, *-ats-* marks the causative COS. This is illustrated for various property concept words in (17).

(17) Category 1 (Megerdoomian 2002:98)

Adjective	Non-causative	COS Causative COS
layn (wide)	layn.anal (widen)	layn.ats.nel (widen)
čor (dry)	čor.anal (dry)	čor.ats.nel (dry)
metz (big)	metz.anal (grow)	metz.ats.nel (grow, bring up)
arag (fast, quick)	arag.anal (quicken)	arag.ats.nel (accelerate)
čaq (fat)	čaq.anal (become fat)	čaq.ats.nel (fatten)
sev (black)	sev.anal (blacken)	sev.ats.nel (blacken, darken)

Contrasting with the situation for words whose denotation includes a property concept meaning are words whose meaning includes a result state. For words in this class, there is no basic adjectival form.³ The data in (18) illustrate this point and also show that while the causative COS form is morphologically basic, the non-causative COS form is derived from the latter with the *-v-* morpheme, a marker of the passive according to Megerdoomian (2002:91).

(18) Category 2 (Megerdoomian 2002:98)

Adjective	Causative COS	Non-causative COS
–	k’ot’Rel (break)	k’ot’R.v.el (break)
–	epel (cook)	ep.v.el (cook)
–	poxel (change)	pox.v.el (change)
–	šarjel (move)	šarj.v.el (move)
–	xort’ak’el (sink, drown)	xort’ak’.v.el (sink, drown)

In Eastern Armenian, then, just as has been seen for English and Quechua, there is a contrast in the relationship between words denoting states and words denoting their change of state counterparts. While the morphologically basic form is stative for property concepts, it is the change of state for words whose denotation includes a result state.

3.1.4. Tongan

Tongan is yet another language that distinguishes two classes of states and changes of state. Again, words whose denotation includes a property concept have a morphologically underived form that denotes a state. How the states are

related to changes of state is somewhat more complicated in Tongan, however. Property concepts are lexicalized as verbs in this language, and the same word can have either a state meaning or a non-causative COS meaning, depending upon the aspectual context it appears in.⁴ I discuss these and similar data below. The important point for now is simply that the word denoting the property concept is morphologically basic, as shown in (19a). Note too that the word denoting the causative is derived from the word denoting the state/non-causative change of state via prefixation with *faka-*, as in (19c).

(19) Tongan (property concepts)

- a. Ko e hala 'oku *lahi*.
PRSTNL the road PRESS wide
'The road is wide.'
- b. Hili pe 'uluaki fo'i'akau', kuo *lahi* ia.
after only first medicine, PRFCT big him
'After only one pill, he became big.'
- c. Na'e *faka-lahi* e he puleanga 'a e hala.
PAST faka-wide ERG the government ABS the road
'The government widened the road.'

As in other languages discussed, the direction of derivation for words whose denotation includes a result state is the reverse, as shown by the data in (20). For such words, the word denoting the causative change of state is morphologically basic as in (20c), while the word denoting the state and the non-causative change of state is derived from the word denoting the causative change of state via the prefix *ma-*, as in (20a,b).

(20) Tongan

- a. Ko e hele 'oku *ma+pelu*.
PRSTNL the knife PRES bent.
'The knife is bent.'
- b. Ko e hele kuo *ma+pelu*.
PRSTNL the knife PRFCT bent
'The knife became/got bent.'
- c. Na'e pelu 'e Mele 'a e hele.
PAST bend ERG Mele ABS the knife
'Mele bent the knife.'

There are two important points to take away from the Tongan data. First, as observed for the other languages under discussion, there is an asymmetry in

direction of derivation—for words whose denotation includes a property concept, states are basic. For words whose denotation includes a result state, however, the causative change of state is basic and result states are derived. Secondly, in Tongan a single word takes on both a state and a non-causative COS meaning. I argue below that this is possible only because Tongan has states lexicalized as verbs.

3.2. *Which states are derived and which are basic?*

In English, Quechua, Eastern Armenian, and Tongan, then, while the direction of derivation for words whose denotation includes a property concept meaning is from state to change of state, this is not the case for words whose denotation includes a result state. For these types of states, in many cases the word denoting them is derived from the word denoting the change of state. These data, taken alongside Dixon's study of languages without adjectives, suggest that property concepts are denoted by morphologically basic forms. They may be lexicalized as either stative verbs, nouns, or adjectives, depending on the language, but are morphologically basic whatever their lexical category encoding. This generalization is stated in (21).⁵

- (21) **Generalization 1:** If X is a property concept meaning, then the lexeme Y denoting X is morphologically basic.

Given (21), if there is any overt derivational relationship between words denoting states, non-causative and causative changes of state, then, the words denoting the changes of state will be derived from the word denoting the state, as illustrated in (12) for English, (15) for Quechua, (17) for Eastern Armenian, and (19) for Tongan. The generalization also holds in other languages I have looked at, such as Central Alaskan Yup'ik (Jacobson 1984, 1995), Cora (Vázquez Soto 2001), Spanish, and other Polynesian languages like Maori (Bauer 1993).⁶

3.3. *What is the impact of crosslinguistic variation in lexical category inventory?*

Diversity in lexical category encoding of property concepts turns out to have an interesting impact on the relationship of words denoting property concept states to words denoting their associated non-causative changes of state. I have observed two types of languages so far as this relationship is concerned. The more familiar kind of language is exemplified by O'odham, Spanish, and Warlpiri in (22)–(24). These are languages in which the word denoting the non-causative change of state is derived from the word denoting the property concept through some sort of morpholexical process overtly marked by morphology. In O'odham, as shown in (22), where property concepts are said to be lexicalized as adjectives, the addition of a suffix derives a non-causative change of state from the property

concept state, and the causative change of state is, in turn, derived from the non-causative change of state. In Spanish, as shown in (23), where property concepts are also lexicalized as adjectives, this is done by some combination of prefixes and suffixes. Warlpiri, as shown in (24), where property concepts are lexicalized as nouns, derives words denoting non-causative changes of state from the word denoting the state with a suffix. Words denoting causative changes of state are also derived from the state-denoting word, but with a different suffix.

(22) O'odham (Hale and Keyser 1998:92)

	Adjective	Non-causative COS	Causative COS	
a.	(s-) wegi	weg-i	weg-i-(ji)d	'red'
b.	(s-)moik	moik-a	moik-a-(ji)d	'soft'
c.	(s-)'oam	'oam-a	'oam-a-(ji)d	'yellow'

(23) Spanish

	Adjective	Non-causative COS	Causative COS	
a.	triste	en-triste-cer se	en-triste-cer	'sad'
b.	duro	en-dure-cer se	en-dure-cer	'hard'

(24) Warlpiri (Hale and Keyser 1998:93)

	Noun	Non-causative COS	Causative COS	
a.	wiri	wiri-jarri-	wiri-ma-	'big'
b.	maju	maju-jarri-	maju-ma-	'bad'

This situation contrasts with that observed in certain other languages, such as Tongan. In this language, as discussed above, property concepts are lexicalized as verbs and the same word is polysemous between a state and a non-causative COS denotation, as shown by the data in (25). Words denoting causative changes of state are derived from the state denoting words with a distinct morpheme, *faka-*, as shown in (25c).

(25) Tongan

- a. Ko e hala 'oku *lahi*.
 PRSTNL the road PRES wide
 'The road is wide.'

- b. Hili pe 'uluaki fo'i'akau', kuo *lahi* ia.
 after only first medicine, PRFCT big him
 'After only one pill, he became big.'
- c. Na'e *faka-lahi* e he puleanga 'a e hala.
 PAST CAUSE-wide ERG the government ABS the road
 'The government widened the road.'

Though there is no derivational morpheme signaling the difference between the state and the non-causative COS denotation in (25a,b) above, there is a difference in aspect marking—while the use of the continuous marker *'oku* correlates with an ongoing state denotation, use of the perfect marker *kuo* correlates with a non-causative COS denotation.⁷ This kind of polysemy is not unusual—it has been observed elsewhere in the literature on the typology of aspect marking that perfective marking of a stative verb often yields a change of state interpretation. Bybee *et al.* (1994:75–76) and Chung and Timberlake (1985) in particular, have made observations on a quite general scale. Chung and Timberlake's comments are particularly clear.

Applied to states, closure implies a complete change of state, specifically inception rather than cessation. Thus, languages that have a morphological category (traditionally called perfective) to specify closure for processes often use the same category to signal inception of a state. (Chung and Timberlake 1985:217)

Similar comments are found in Tatevosov (2002:340ff.), Talmy (1985:92), Smith (1997:70), and Wetzer (1996:189). Additionally, various authors have observed similar facts in numerous languages. Comrie's observations on Mandarin Chinese are representative.

In Mandarin Chinese ... a number of predicates, both adjectives and verbs, that normally refer to a state can have ingressive meaning in the Perfective, e.g. *tā gāo* 'he is tall', *tā gāo-le* (Pfv.) 'he became tall, has become tall'. (Comrie 1976:19–20)⁸

Similar observations have been made by Lefebvre and Brousseau (2002:88) for Fongbe, Prasithratsint (2000:262) for Thai, Chung and Timberlake (1985:238) for Mokilese, and Enfield (2003:262) for Lao. Representative data from some of these languages are given below.

(26) Mokilese (Chung and Timberlake 1985:238)

- a. Pahrangkije pe pwespwespwes
 iron still warm(PROG)
 'This piece of iron is still warm.'

- b. Ih lioas-ka
 he angry-PERF
 ‘He got angry.’

(27) Thai

- a. khâw kamlang ruay læy chôn s# khōng phεεng
 (s)he now rich so like buy thing expensive
 ‘(S)he is rich now. (S)he likes to buy expensive things.’
 (Prasithratsint 2000:263,(22b))
- b. khâw ruay léεw chiiwit khâw mây ma#n téε-kōn
 (s)he rich already life (s)he not like past
 ‘(S)he has become rich. His/her life is not like in the past.’
 (Prasithratsint 2000:262, (21b))

(28) Lao

- a. ?khò:j5 kamlang2 suung3
 1SG PROG tall
 ‘I am being/getting tall.’ (Enfield 2004:331, (35))
- b. khon2 suung3
 person tall
 ‘the tall person’ (Enfield 2004:336, (56))
- c. ?khò:j5 daj0 suung3
 1SG ACHV tall
 ‘I was/got to be tall.’ (Enfield 2004:331, (30))

Interestingly, all of these languages are ones in which property concepts are described as belonging to the lexical category verb. Further, it is only in such languages that I have found reports of this kind of polysemy; in languages where property concepts are lexicalized as nouns or adjectives, as in Spanish, Warlpiri, and O’odham, there is no such polysemy. Instead, the relationship between words denoting states and words denoting non-causative changes of state is signaled derivationally. These observations suggest the generalization in (29).

- (29) **Generalization 2:** When a single lexical item γ is polysemous between a state and a change of state denotation, γ belongs to the lexical category verb.

By *polysemy* in (29), I have in mind a particular kind of polysemy, called *logical polysemy* by Pustejovsky (1995: chapter 3), who makes distinctions between three different types of ambiguity. First is *contrastive ambiguity*, characterized by an “arbitrary association of multiple senses with a single word,” as with the

different senses of the lexical item *bank* (Pustejovsky 1995:29). A more systematic type of ambiguity goes by the name of *complementary polysemy*, in which the distinct senses of a word share some sort of systematic relationship to one another, often having some overlap in meaning with one another. The final type of ambiguity is known as *logical polysemy*, which Pustejovsky (1995:28) defines as "... a complementary [polysemy] where there is no change in lexical category, and the multiple senses of the word have overlapping, dependent or shared meanings." The kind of ambiguity observed for property concept words in Tongan and languages like it is this final kind, logical polysemy. States and changes of state overlap with one another in meaning, since a change into a state entails the state. This point is made clear by the event structure representations introduced below in (31). This overlap in meaning makes the Tongan situation at least one of complementary polysemy. Further, though, in Tongan and the other languages discussed above there is no change in lexical category associated with the different senses. This clearly contrasts with the situation in languages like English, where those property concept words that do not allow *-en* suffixation for phonological reasons (Jespersen 1939:53ff.) are polysemous, but only in a complementary fashion rather than a logical fashion, since there is a change in lexical category associated with the different senses. This is illustrated by the data in (30).

- (30) a. The sky is clear.
 b. The sky will clear this afternoon before the game.

Despite the fact that a word of the phonological shape *clear* appears in both (30a) and (30b), only verbs appear in the context following the auxiliary *will* (**will red, *will blue, *will large*, cf. *will redden, will become blue, will enlarge*). Further, only the COS sense of *clear* is available in (30b); it cannot have the meaning that when a particular time rolls around, the sky will (already) be in a state of clarity. Instead, (30b) must have the meaning that when the afternoon comes, the sky will undergo a change of state, from being not clear, to being clear. The upshot of this is that for cases like English *clear*, the state reading is associated only with the use of *clear* in adjectival contexts, while the COS reading is available only in verbal contexts. This is fundamentally different from the situation in languages like Tongan, where both senses belong to a single word of a single lexical category.⁹

The typological generalization, then, is that there seem to be two types of languages as far as the derivation of non-causative changes of state from property concept states is concerned, and that the type of derivation a language uses is in part correlated with how it lexicalizes property concepts. A single word is polysemous between non-causative changes of state and states only where the latter are lexicalized as verbs. What does not exist are languages of two types:

(a) languages where a single lexical item belonging to the lexical category noun has both a state and a change of state sense, and (b) languages where a single lexical item belonging to the lexical category adjective has both a state and a change of state sense. In section 4, I develop an explanation for this generalization rooted in the nature of the mapping between lexical semantics and syntactic categories, hypothesizing that verbs are the only lexical category that can denote changes of state.

4. DERIVING THE GENERALIZATIONS

In the remainder of the paper I go on to develop an analysis of the observed typological facts rooted in the nature of lexical semantic representation and how word meaning and lexical categories are related to one another. I begin by laying out my theoretical assumptions and then make use of these to derive the typological generalizations discussed above.

4.1. Theoretical assumptions

4.1.1. Event structure representations

Much work has argued in favor of two distinct components of lexical meaning—an idiosyncratic component and a more structural component meant to account for facts such as diathesis alternations and Vendlerian verb classes (Grimshaw 1993; Hale and Keyser 2002; Jackendoff 1983, 1990; Mohanan 1994; Pinker 1989; Rappaport Hovav and Levin 1998; Wierzbicka 1988). In what follows, I assume that words have meaning built up from a universal inventory of semantic operators operating on a semantic constant, the locus of idiosyncratic lexical meaning, that gives each word its unique identity (Grimshaw 1993; Rappaport Hovav and Levin 1998). The combination of these operators and these constants gives rise to various event structures, which are meant to account for similarities in meaning and behavior of different lexical items. Properties of these event structures cutting across different lexical items define lexical classes that account for the behavior of verbs with respect to their arguments (diathesis alternations; Levin 1993) and for the aspectual meanings of interest in this paper—state versus change of state.

Event structures that play a role in the remainder of the paper are those in (31), taken from Rappaport Hovav and Levin (1998:108), and might correspond loosely to the Vendlerian classes.¹⁰ These event structures encode the basic units of meaning of interest in the context of the discussion below, in particular states and changes of state. Changes of state can vary significantly in their degree of lexical semantic complexity, (31c) and (31d) both being changes of state.

- (31) Some event structure templates (Rappaport Hovav and Levin 1998:108)
- a. [x ACT_{<MANNER>}] (activity)
 - b. [x <STATE>] ((non-result) state)
 - c. [BECOME [x <STATE>]] (non-causative COS)
 - d. [[x ACT_{<MANNER>}] CAUSE (causative COS)
[BECOME [x <STATE>]]]

There are a couple of basic features of the simple ontology in (31) that should be highlighted. First, the basic building blocks are activities (31a) and states (31b), which are the primitive notions corresponding to the semantic content of all event structures. Second, more complex event structures are generated by combining operators like CAUSE and BECOME with the state and activity primitives. Indeed, this is how the event structures corresponding to non-causative (31c) and causative (31d) changes of state are arrived at. The BECOME operator is responsible for adding COS semantics, while the CAUSE operator adds causative semantics, part of which includes the introduction of a causing subevent.

4.1.2. *The principle of monotonic composition*

Following Olsen (1996) and Rappaport Hovav and Levin (1998:103), I assume that event structure representations are constructed monotonically. The idea is that while meaning, in the form of event structure operators, can be added to an event structure as a consequence of word formation processes for example, meaning may not be removed. I call this the Principle of Monotonic Composition, as stated in (32).

- (32) The Principle of Monotonic Composition (PMC)

Word meaning is constructed monotonically on the basis of event structure constants and operators.¹¹

The PMC makes predictions about how the meanings of words can expand and contract (see Rappaport Hovav and Levin (1998:103ff.) for some discussion). In the domain of investigation here, the relationship between property concept states and the non-causative COS counterparts, the PMC makes an especially interesting prediction—non-causative changes of state can be derived from states, but not vice versa, as the latter derivation would be a non-monotonic one. This can be seen clearly by considering the event structure representations laid out in (31) for states and non-causative changes of state. These are repeated in (33) and (34).

- (33) Event structure template for property concept state

[x <STATE>]

- (34) Event structure template for non-causative COS

[BECOME [x <STATE>]]

While a BECOME operator could be added to a stative event structure template in order to derive a change of state, the reverse could not happen. This would involve deletion of a BECOME operator, inconsistent with the PMC. The idea, then, is that any derivation can only be meaning adding.

A word of clarification is in order here regarding the domain of the PMC and what is meant by the word *derivation* in this context, a meaning different from how the word is used elsewhere in the paper. Up until this point, I have been concerned with the derivation of words in the context of morphological derivations; i.e., I have been concerned with the morphological shapes of words. The PMC is not about this kind of derivation, but instead about the derivation of *meanings of words*. This does, however, have an impact on what types of morphological derivations are observed, in particular in relation to the derivation of words denoting states from words denoting the corresponding changes of state, as I discuss below.

4.2. Deriving Generalization 1

Having introduced the event structure representation and the PMC, I now have enough theoretical machinery in order to derive Generalization 1, laid out in (21) above, and repeated in (35).

- (35) **Generalization 1:** If X is a property concept meaning, then the lexeme Y denoting X is morphologically basic.

Generalization 1 follows straightforwardly from the PMC, stated in (32). Put simply, there is a way to derive change of state meanings monotonically from basic state meanings, while there is no way to monotonically derive basic state meanings from change of state meanings. To see this, consider the event structure representations of state and change of state meanings in (36) and (37).

- (36) Event structure representation of a basic state

[x <STATE>]

(37) Change of state event structure representations

- a. [BECOME [x <STATE>]] (non-causative COS)
- b. [[x ACT_{<MANNER>}] CAUSE (causative COS)
[BECOME [x <STATE>]]]

While the change of state event structures in (37) can be derived from the basic state event structure in (36) via the addition of operators and constants, this would not be true for a derivation from one of the event structures in (37) to the one in (36). Such a derivation would crucially entail the deletion of bits of meaning from the event structure representation and would therefore be ruled out by the PMC. States *can* be derived from change of state event structures, but in order for this to take place in a manner consistent with the PMC, information will have to be added, leaving the change of state information as part of the decomposition. Such a decomposition would be precisely that of a result state. The prediction, then, is that while a result state could be derived from a change of state verb (via the addition of another state operator to a change of state decomposition), a property concept state could not be derived from a change of state, since this would involve the deletion of lexical semantic operators.¹²

Given the PMC, then, it cannot be the case that a property concept state could be derived from a change of state. How does this relate to the morphological shape of the words denoting these meanings, though? I take as a point a departure the idea that there are certain types of morphological operations that concomitantly effect changes on both the phonological shape and the event structure representation of words (e.g. Levin and Rappaport Hovav 1998; Stump 1998). The PMC, further, specifies that changes made to an event structure representation can only be additive. Since the PMC holds that *any* operation on event structure representations must be additive, it of course must also be true for morphological operations that effect changes on event structure representations. Despite its being vacuously true, however, it does have interesting implications for morphological typology. In the domain under investigation here, since words denoting property concepts have a proper subset of the event structure constants and operators that their associated changes of state have, it is predicted that words denoting property concepts have undergone a subset of the event structure changing morphological operations that words denoting their associated changes of state undergo. In cases where the morphological operations lead to overt morphological differences between words denoting property concepts and the associated changes of state, as with affixation, there are observable differences between the two types of words in the direction of the prediction.¹³

So far, then, given the PMC it is clear why words denoting changes of state based on property concepts should be derived from words denoting the associated property concept state. Generalization 1, however, is stronger than this—the observation was that words denoting property concepts are morphologically

basic, having undergone no meaning changing morphological operations. I believe this stronger generalization also follows from the PMC and from an idea that is at least implicit in the pioneering work of Dixon (1982)—property concept meanings are conceptually basic meanings. By this, I mean that property concept meanings are primitive bits of meaning and that the event structure representations of them are underivable from any other event structure representation. The whole enterprise of lexical decomposition rests on the idea that there are certain bits of meaning that serve as primitive lexical semantic constants. My claim is that property concept meanings are among these primitive constants.¹⁴ This claim is in need of philosophical and linguistic justification, with the linguistic justification ideally coming from a domain outside of morphology (in order to offer non-circular linguistic evidence). Such evidence will most likely need to come from the domain of lexical semantics. This is a project I leave for future research. I note, though, that if this is correct, then the strong version of Generalization 1 clearly follows. If property concepts are primitive constants, then there can be no operation on event structure that derives them from something else, since event structure operations only add meaning. Given this, it therefore follows that there exist no event structure changing morphological operations deriving words denoting property concepts; words denoting property concepts would have to be morphologically basic.

4.3. *Deriving Generalization 2*

I now return to Generalization 2 stated in (29) and repeated in (38).

- (38) **Generalization 2:** When a single lexical item γ is polysemous between a state and a change of state denotation, γ belongs to the lexical category verb.

The explanation for Generalization 2 lies in the mapping between lexical semantics and morphosyntax. Specifically, I believe that change of state semantics can only be realized by words belonging to the lexical category of verb. If true, then it should be clear why Generalization 2 would hold—this generalization is about the conditions under which one word has more than one denotation, specifically, state and change of state.

If states are lexicalized with a category other than verb, and if change of state semantics universally must be realized by verbs, then the single word denoting both of these meanings would belong to the wrong lexical category in order to denote a change of state. In such a case, then, there can be no polysemy; there must be separate (though possibly related) words denoting states and the related changes of state. This kind of explanation is reminiscent of research in

the functionalist tradition on lexical categories, which tends to appeal to the idea that the essence of the lexical categories is rooted in lexical semantics. While many have pursued the idea, due to Givón, that the lexical categories follow from some notion of time stability (Givón 1984), others have pursued something more akin to the traditional idea that nouns prototypically name places/things and verbs prototypically name actions (Croft 1991, 2001; Beck 2002). This general line of research rooting the nature of syntactic categories in lexical semantics has been met with some degree of skepticism, especially in the generative literature, in large part due to the fact that the characterization of the meaning attributed to the various lexical categories is never made very explicit, which makes the proposals somewhat difficult to rigorously evaluate. In very recent work, however, Lieber (2004) proposes an idea which opens up the possibility for capturing the best of both the more functional and more formal approaches to the study of lexical categories. Working in the context of a decompositional approach to word meaning similar in spirit to the one adopted above, Lieber (2004:37) makes a link between operators in particular positions of the event structure and syntactic category. The proposal, as stated by Lieber, is as in (39).

- (39) The outermost function of the skeleton [=event structure] determines the syntactic category (Lieber 2004:37)

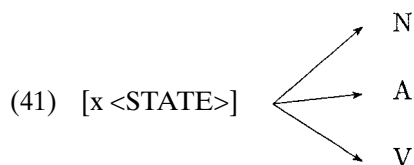
Lieber's proposal is different from prior semantically based definitions of the lexical categories in two fundamental ways. First, unlike other semantically based approaches to lexical categories, Lieber's approach admits a unidirectional interpretation. If δ is the outermost function of an event structure, it could easily be that while all δ s are mapped into verbs, there is nothing contradictory if not all verbs denote δ . This contrasts with e.g. Givón's approach, which is often criticized by generativists for reasons having to do with bidirectionality. Givón claims, for example, that nouns are the most time stable of the lexical categories, with adjectives being slightly less time stable, and verbs being the least time stable. Given this claim, then, it seems quite odd that languages should have stative verbs. Action nominalizations, such as *birth*, *conception*, etc. are often viewed as counterexamples in the other direction, since these are not at all time stable, but still belong to the nominal domain (see Baker 2003:293 for criticisms exactly along these lines).¹⁵

Given a decompositional approach to meaning and Lieber's proposal, these concerns can be overcome. Assuming Lieber's proposal in (39), the event structure assumptions laid out above, and the BECOME operator as a representation of change of state semantics (on which see Dowty 1979 for formalization), the idea that "change of state is realized only by verbs" can be made more precise, via the mapping rule in (40).

- (40) [BECOME ϕ] \rightarrow V

The idea captured by (40) is that if BECOME is the highest operator in a decomposition, then the word associated with that lexical decomposition must be assigned to the lexical category of verb.¹⁶ As discussed above, the empirical observation in (29) follows from (40)—if a word is to be polysemous between a state meaning and a change of state meaning, then given (40) such a word must belong to the lexical category verb. Otherwise, the state word would belong to a lexical category incompatible with change of state meaning, in violation of the mapping rule in (40).

The mapping rule in (40) is not the only such rule. Indeed, Lieber's proposal suggests a whole family of such mapping rules, one for each operator that can be the highest in a lexical decomposition. So, I assume that there are rules for the other operators commonly discussed in the lexical semantic literature, though more study is needed to determine exactly what the rules look like. In the case of states, the nature of the mapping rule is made clear by past research, especially that of Dixon (1982), who observes that states turn up as nouns, verbs, and adjectives across languages. The mapping rule for states, then, is the one in (41).¹⁷



The idea, then, is that there is a mapping rule for each operator. In the case of states, the mapping rule varies across languages—in some languages they are mapped to adjectives (e.g., English), while in others they are mapped to nouns (e.g., Quechua) or verbs (e.g., Tongan). The situation with the BECOME operator, however, is different. It is universal that if BECOME is the highest operator in an event structure, then the word with that meaning must be mapped to the lexical category verb. This is the idea formalized in (40).

As already discussed, the idea embodied by the mapping rule in (40) explains why Generalization 2 should hold. This rule, however, makes additional predictions that seem to be borne out by local facts from particular languages. For example, consider the contrast in (42).

- (42) a. Kim quickly believed Sandy.
 'Kim quickly came to believe Sandy.' (COS reading possible)
- b. *Quickly cool rooms are always located in the basement. (COS reading impossible)

Both *believe* and *cool* denote states, yet while the verb *believe* can also have a change of state meaning (42a), the adjective *cool* cannot (42b).¹⁸ This follows from (40), if *believe* is a verb in (42a) and if *cool* in (42b) is not.¹⁹ In this

way, then, the theory makes predictions down to the level of individual lexical items—regardless of how the language otherwise behaves, if a stative meaning is lexicalized as either a noun or an adjective, then a change of state meaning for that lexical item is unavailable (as is the case for stative meanings lexicalized as adjectives in English). In contrast, that same language can have single lexical items that are polysemous between a state and a change of state sense, as long as those lexical items belong to the lexical category verb. This is observed not only in languages like Tongan, Mokilese, and others discussed above, but in English for certain lexical items like *believe*.²⁰

The mapping rule in (40) is strong enough to explain Generalization 2 while at the same time making other falsifiable predictions, yet it is also nuanced enough that it doesn't run into the kinds of problems that other lexical semantic definitions of lexical categories run into. Consider one of the “counterexamples” to notional definitions of lexical categories highlighted by Baker above—action nouns like *destruction*, *conception*, *explosion*, etc. Nouns like these all intuitively seem to say something about a change of state. *The destruction of the city*, for example, names an event wherein a city is not destroyed at one interval and is destroyed at another, later interval. The fact that *a destruction* names an event (Zucchi 1993 and references there), though, is the crucial point—the noun *destruction* has a denotation different from that of the verb *destroy*, from which it is derived (Zucchi 1993, building on work by Zeno Vendler 1967, 1968, 1975). While a nominalization denotes an event, the verb a nominalization is derived from denotes a function from individuals to propositions. The way to handle this, on the view advocated above, then, is to say that there is an operator, EVENT, added to a decomposition when a verb undergoes nominalization. The nominalization of a change of state [BECOME ϕ], then, would look something like (43).

(43) [EVENT [BECOME ϕ]]

As a result of the nominalization operation, BECOME is no longer the highest operator in the decomposition. Though the decomposition retains its change of state semantics, the change of state semantics are embedded underneath a different operator that has its own different mapping rule. This analysis is spelled out more formally in Koontz-Garboden (2004a), who builds on the analysis of Zucchi (1993).²¹

Much more work obviously needs to be done to determine (a) what other predictions (40) makes, (b) if they are correct, and (c) what sorts of mappings to syntactic categories other event structure operators have. Assuming that (40) holds up to closer scrutiny, it gives an account of the observed impact of lexical category encoding on the derivational relationship between words denoting property concept states and words denoting their corresponding changes of state. More broadly, the general approach suggests a way of wedding notionally

oriented approaches to lexical categorization with more formal approaches to linguistic theorizing.

5. A MORE COMPLICATED TYPOLOGY?

So far as I am aware, there are no obvious counterexamples to Generalization 2. However, there do appear to be counterexamples to Generalization 1, at least on the surface. The Misumalpan languages of Nicaragua and Honduras are particularly interesting in this regard, as they appear to have words denoting property concepts and words denoting their associated changes of state that are not monomorphemic. This is illustrated in (44) for Ulwa, an endangered Misumalpan language spoken on Nicaragua's Atlantic coast.

(44) Ulwa (Hale and Keyser 2002: 122–123)

State	Non-causative COS	Causative COS	Gloss
sang-ka	sang-da	sang-pa	green/blue
yûh-ka	yûh-da	yûh-pa	long, tall
baras-ka	baras-da	baras-pa	black, dark

As the data in (44) show, Ulwa appears to have words denoting property concept states and their corresponding changes of state that are all derived from a more abstract root. Particularly interesting from the perspective of Generalization 1 is the fact that this is true even for words denoting property concept states, which appear to be composed of a root plus suffix. The obvious question that needs to be explored is what the nature of the *-ka* suffix in the first column of (44) is. If it is an inflectional affix, then it is not part of the lexeme, and therefore does not bear on Generalization 1. If it cannot be shown to be inflectional, however, then its status relative to Generalization 1 becomes more of an issue.

The available data do not at the moment offer a clear answer as to the status of the *-ka* suffix. It turns out that possessive nouns in Misumalpan languages are inflected for person and number. The paradigm is given in (45).

(45) Nominal possessive paradigm (Green 1999:78)

CNS1.SING	-ki	CNS1.PL.EXCL	-ki-na
CNS2.SING	-ma	CNS2.PL	-ma-na
CNS3.SING	-ka	CNS3.PL	-ka-na
		CNS1.PL.INCL	-ni

As illustrated in (45), the marker of the 3rd person possessive has the same phonological shape as the suffix showing up on words denoting property concept states (i.e., it is *-ka*). Some examples of *-ka* in its function as a possessive marker of nouns are given in (46).

- (46) a. *dai-ka*
 brother.in.law-CNS3
 ‘his/her brother-in-law’ (Green 1999:32)
- b. *pan-ka*
 tree-CNS3
 ‘his/her tree’ (Green 1999:38)
- c. *aidingh-ka*
 watchamacallit-CNS3
 ‘his/her watchamacallit’ (Green 1999:39)

Facts like these suggest the possibility that words denoting property concepts are nouns in these languages, and that the *-ka* marker is simply the normal nominal inflectional marker showing up on nouns.

Things are more complicated than this, however, as words denoting property concepts seem to have certain syntactic properties that *-ka* marked nouns do not have. For example, the semantic headedness of a phrase syntactically headed by a *-ka* marked word denoting a property concept is different from the semantic headedness of a phrase syntactically headed by a *-ka* marked word with a more referential meaning.

- (47) a. *Yang kuh pan-ka buk-payang.*
 PRN1 firewood stick-CNS3 split-PRES1
 ‘I am splitting a stick of firewood.’ (Green 1999:70)
- b. *yang û-ki*
 PRN1 house-CNS1
 ‘my house’ (Green 1999:82)
- (48) a. *muih yam-ka*
 person good
 ‘good person’ (Green 1999:136)
- b. *damaska sik-ka*
 jungle great
 ‘deep forest’ (Green 1999:135)
- c. *pukka bara-ka*
 night dark
 ‘dark night’ (Green 1999:139)

While the semantic head of the highlighted phrases in (47) is the possessive marked noun, the semantic head of the phrases in (48) is the word preceding the *-ka* marked word. Study of available materials suggests that there are further asymmetries in the behavior of *-ka* marked words that bear on the questions discussed here. For example, while some property concept words require *-ka* affixation, according to Green (1999:134), it is optional with others. Further, preliminary fieldwork on the language suggests that there are still other property concept words which do not take *-ka* at all. These facts are complicated and not yet well understood. They require further research focusing on noun phrase syntax, adjectival semantics, and the nature of the lexical category system of the language before their implications for Generalization 1 can be determined.

6. CONCLUSION

Though the research I have reported is still in its preliminary stages, several important empirical generalizations have already emerged. First, I have shown, following Dixon (1982) that there are (at least) two different types of states from a lexical semantic perspective—property concepts, which entail no prior event, and result states, which do entail a prior event giving rise to a result state. This lexical semantic distinction was found to have an impact on how words with these types of meanings are lexicalized. While property concepts are lexicalized as morphologically basic forms, this is not always the case for result states. I have argued that these facts follow from the Principle of Monotonic Composition, a principle that holds that the construction of meaning is carried out monotonically. Secondly, I find that some languages have words that are polysemous between a state and a change of state denotation. Due to a constraint on the mapping between lexical semantic and syntactic categories that only verbs can denote changes of state, polysemy arises only in languages where property concepts are lexicalized as verbs.

From a theoretical perspective, these observations suggest that theories of event structure that give homogeneous representations to all COS predicates (e.g. Baker 2003) need to be revisited. There seems to be a contrast in the behavior of words denoting property concept states and result states that seems in part attributable to the fact that these are fundamentally different kinds of states, with different types of meanings. Further, I also find that there are differences in how non-causative changes of state are derived from property concepts, depending on the lexical category of property concepts, which varies crosslinguistically. These are typological differences that theories of event structure should seek to capture, regardless of the nature of the theory (i.e., syntactically or semantically based). I have developed analyses for these observations within a decompositional approach to lexical meaning. Future work should focus on further investigation of

the empirical observations made here and the predictions of the theory. As well, further research is needed exploring how competing theories of event structure might capture the same set of facts.

NOTES

* I would like to thank audience members at both the 2004 meeting of the LSA and the Fourth Mediterranean Morphology Meeting for their questions and comments. I am also grateful to Sisilia Lutui for help with Tongan and to Martina Faller for help with Quechua. I have received extremely useful feedback on this paper from three anonymous *Yearbook of Morphology* reviewers as well as John Beavers, Paul Kiparsky, and Beth Levin. I also acknowledge the following people for comments and data related to this work: Elena Benedicto, Lev Blumenfeld, Melissa Bowerman, Joan Bresnan, Luis Casillas Martínez, Eve Clark, Mike Dukes, Hana Filip, Ivan García-Álvarez, Martin Haspelmath, Shelly Lieber, Jimmy Lin, D. Gary Miller, Yuko Otsuka, Sergei Tatevosov, and Judith Tonhauser. This paper builds on initial joint work with Beth Levin (Koontz-Garboden and Levin 2004) and was supported in part by Graduate Research Funds from the Department of Linguistics at Stanford as well as by an NSF Small Grant for Exploratory Research (NSF Grant BCS-0004437) to Beth Levin.

¹ Here and throughout the paper I intend the phrase “morphologically basic” to mean that a word stripped of its inflectional morphology (i.e., a lexeme) and of any semantically irrelevant derivational morphology (e.g. syntactically conditioned derivational morphology) has undergone no morphological rules (ablaut, reduplication, affixation, etc.). In contrast, I use the phrase “morphologically derived” when a lexeme has undergone such a rule. For my purposes, the determining factor for whether a lexeme is considered morphologically derived or basic has to do with whether the lexeme has undergone any morphological rules effecting changes in the lexical semantic representation (i.e., lexical conceptual structure) of the lexeme. For further discussion see Stump’s (1998: 13ff.) characterization of the notion *lexeme* and of morphological rules deriving new lexemes from lexemes, effecting a change in the lexical semantics of the original lexeme.

² See also Dixon (1982:22) for discussion of the conditions. According to him, *-en* can be attached to adjectives ending in *p, t, k, f, s, ʃ, θ*, and *d*.

³ Megerdooomian does not discuss how result state meanings are expressed in Eastern Armenian. The point of her discussion was simply that there is no morphologically basic form for states associated with an externally caused change of state. Further data collection needs to be undertaken to determine how these meanings are expressed. The implication of her discussion is that if they are expressible at all, it is via some sort of more complex form, as observed for the other languages discussed in this paper.

⁴ See Koontz-Garboden (2004b) for further discussion and analysis of the Tongan facts.

⁵ An anonymous reviewer notes that words denoting property concepts could be morphologically complex for reasons independent of lexical semantics. For example, there may be adjective specific inflectional morphology, as in, e.g. Latin. Things are similar in Mohawk, as another reviewer points out, where, according to the reviewer “... the predicate meaning ‘white’ is a verb made up of two morphemes, the bound root ‘rak’ and the “stative” affix *-v*. Many other state-denoting words are similar.” It is for this reason that Generalization 1 in (21) is formulated in terms of the notion of a lexeme. In cases like these, though the word denoting

the property concept may be morphologically complex, the complexity is due to the addition of morphology that does not effect any changes at the level of the lexical semantics; the lexeme denoting the property concept state is morphologically basic. As the second reviewer points out, an independent definition of what a lexeme is, is needed. I agree, though providing such a characterization is beyond the scope of this paper (though see Stump 1998 for discussion).

⁶ The lexicalization of result states and the COS predicates related to them requires further research, as some languages such as Lakhota (Boas and Deloria 1939; Foley and Van Valin 1984) and Tagalog (Foley and Van Valin 1984) seem to lexicalize result states as morphologically basic forms, with words denoting the non-causative and causative changes of state built on top of them. For this reason, Generalization 1 is stated as a unidirectional conditional, rather than as a bidirectional conditional; further research is required to determine how strongly the converse of Generalization 1 holds crosslinguistically (that if X is a result state meaning, then the lexeme Y denoting X is morphologically derived).

What is noteworthy, though, is that in all languages I have examined, the paradigms involving result states are morphosyntactically distinct from those involving property concepts. For example, only roots with property concept meanings can be used without additional affixes in Lakhota, while roots with result state meanings must combine with the affix *-hā* to give rise to a stative meaning. This contrast is illustrated by the data in (i) and (ii).

- (i) Property concepts (Foley and Van Valin 1984:41)
 - a. *spaya* ‘be wet, get wet’
 - b. *čāze* ‘be angry, get angry’
 - c. *yazā* ‘feel pain, be sick’
- (ii) Result States (Foley and Van Valin 1984:42)
 - a. *-blaza* ‘be ripped open’
blaza-hā ‘it is rent, torn open’
 - b. *-blečha* ‘be shattered (said of brittle material)’
blečha-hā ‘it is shattered’
 - c. *-wega* ‘be fractured (said of a long round object)’
wega-hā ‘it (e.g. a pole) is broken’
 - d. *-khīča* ‘be scraped’
-khīča-hā ‘the outside is scraped off’

Data like these and those discussed above support the idea that property concepts and result states are two fundamentally different types of states, often down to the level of morphological encoding.

⁷ Here I am actually simplifying significantly due to space considerations. A COS meaning can arise with *‘oku* marked states in the presence of an adverb requiring such a meaning, though the default interpretation of *‘oku* constructions is a stative one. This suggests that what determines whether a property concept word has a state or a COS reading goes beyond grammatical aspect marking. Which reading arises depends on the sentential context, which can lead to the coercion of one meaning or another (Zucchi 1998). These issues are discussed extensively in Koontz-Garboden (2004b).

⁸ Comrie cites Jaxontov (1957:116) as making a similar observation.

⁹ See Koontz-Garboden (2004b) for further discussion of the Tongan facts.

¹⁰ Rappaport Hovav and Levin (2002) show that change of state verbs vary in telicity, and argue based on this that "... lexical aspectual class alone does not determine argument expression" (Rappaport Hovav and Levin 2002:4). These findings are evidence that event structure needs information beyond that which is relevant for Vendlerian classification.

¹¹ One reviewer asks whether the PMC forms part of an innate UG. My view is that the PMC is probably consistent with many different types of theories, and I would not think that one would have to accept the existence of UG in order to accept the PMC. It might be that word meaning is constructed monotonically for functional reasons that have nothing to do with UG. I think this is an open question, one worthy of further research.

¹² I am sympathetic with an anonymous reviewer who says it would be helpful to see the lexical semantic representation of a result state. I hesitate to give such a representation because it is not clear to me whether these tend crosslinguistically to be derived from non-causative or from causative changes of state, or whether languages vary in this regard. Additionally, further research is needed to work out the right kind of formal interpretation of the lexical semantic operator responsible for giving rise to the result state interpretation. (By this, I have in mind the kind of careful formal research carried out by Dowty 1979 for other decompositional operators.) I leave this typological and formal research for the future, noting that these are areas to expand on the research presented in this paper.

¹³ Two anonymous reviewers raise the issue of the relationship between inchoatives and causatives, an issue not directly discussed in this paper, but which the PMC makes predictions about. Both observe that the PMC seemingly leads to the prediction that causatives should only be derived from inchoatives. Despite this, many languages show for many predicates what is often called an 'anticausative' alternation, whereby the word denoting the non-causative COS is derived from the word denoting the causative COS. This type of alternation is illustrated for Spanish by the data in (i).

- (i) a. Juan cocin-ó el arroz.
 Juan cook-3SING.PAST the rice
 'Juan cooked the rice.'
- b. El arroz se cocin-ó.
 the rice SE cook-3SING.PAST
 'The rice cooked.'

If the event structure representations of non-causative and causative changes of state are as I have suggested in (37), then data like those in (i) would be a clear counterexample to the PMC. There is, however, a relatively well-established tradition cutting across formal (Levin and Rappaport Hovav 1995: chapter 3) and functional (Croft 1990:65ff.) studies arguing that in cases such as the one illustrated above, the non-causative COS is indeed derived from the causative COS, consistent with the PMC. The thrust of the evidence, following Croft's formulation, is that for pairs with an anticausative alternation, even though an agent is not overtly encoded for the non-causative alternant, it is still entailed that an agent gave rise to the event. Levin and Rappaport Hovav (1995: chapter 3) state things slightly differently from Croft, appealing to the notion of external causation, but the idea is the same—that although syntactically one of the arguments the causative COS has is missing in the non-causative COS alternant, some semblance of this argument is still semantically entailed. Levin and Rappaport Hovav (1995: chapter 3) formalize this idea by representing the event structure of the non-causative COS in

pairs such as (i) above as having an existentially bound argument. In this way, the non-causative COS event structure is derived from the causative COS event structure monotonically.

What this means, then, is that the representation of the non-causative COS I have given in (37) is, for some types of predicates, an oversimplification. Since my focus is more on the state/change of state alternation, I gloss over this here. Still, I believe that the facts discussed in this note are not inconsistent with a broad interpretation of my theoretical proposals.

¹⁴ Incidentally, this claim is implicit in the event structure representation I have given to property concept states in (36), for example. Claims that certain bits of meaning, e.g. stative meaning, are primitives are implicitly made throughout the lexical decomposition literature, though they are rarely made explicit or their consequences discussed. The literature tends to focus instead on what Rappaport Hovav and Levin (1998) call event structure templates. Much less well-studied are what Rappaport Hovav and Levin (1998) call the *constants* of event structure, namely the primitive units of meaning that event structure operators take as arguments. Bowerman (2004) is one of the only studies I am aware of that looks in any detail at the issue of the primitive constants of event structure, finding crosslinguistic differences, suggesting that languages may differ in interesting ways in the bits of meaning that are taken as primitive constants. Bowerman's preliminary results suggest that languages may in part select bits of meaning to take as primitive constants from a universal semantic space. In this way, primitive constants may be in part language specific, rather than completely universal. I do not believe that such a finding would have a significant impact on the results presented here, though once further research has been carried out, the results should certainly be considered in the context of my findings and theoretical claims.

¹⁵ An anonymous reviewer makes the observation that in Lieber (2004) both adjectives and stative verbs are characterized by the feature [-dynamic]. The reviewer claims that this poses a problem for Generalization 2. In fact, it does not. First, though I adopt Lieber's idea that the outermost function in a lexical decomposition determines syntactic category, I do not necessarily adopt her features (indeed, I make use of BECOME, which Lieber does not use). Further, even if I did make use of Lieber's feature set, this observation would still not be a problem, since the claim is unidirectional, not bidirectional. So, having a single function (or feature) in a decomposition correspond to more than one syntactic category is perfectly consistent with my (and Lieber's) claims. Indeed, the mapping rule in (41) is precisely such a rule.

¹⁶ I agree with two reviewers who remark that the mapping rule in (40) likely has a deeper explanation. At the moment, however, I remain agnostic as to what that might be, though there are certainly proposals in the lexical category literature (cited above) that would help to explain why (40) might hold. Outside of the lexical category literature another, not unrelated possibility, is that (40) and rules like it, should be thought of as bootstrapping rules that aid the process of language acquisition. If (40) were part of an innate UG, then the child trying to acquire language would know to assign words with change of state meaning to a particular lexical category. I have no particular stake at the moment in how (40) is thought about, though I do believe that it is an important outstanding issue, and unraveling how it fits into the larger picture will entail the development of a comprehensive theory of grammar and language and language acquisition in general. These issues obviously go beyond the scope of this paper, but are areas for exciting future research.

¹⁷ This is a very coarse-grained rule that can probably be made more specific based on the semantics of the particular state being mapped to a syntactic category. For example, it is known

from the work of Stassen (1997) that property concepts form an implicational hierarchy in terms of their likelihood to be lexicalized as verbs (at least on Stassen's definition of verbhood).

(i) Stassen's (1997:169) Adjective Hierarchy

HUM. PROP. < PHYS. < DIMENSION/COLOR < VALUE/AGE/Form < MATERIAL/GENDER

The stative mapping rule ultimately should be sensitive to a more fine-grained analysis of categories of states, such as the implicational hierarchy discovered by Stassen. Further work should not only focus on making this more explicit, but also investigating the morphosyntactic and semantic consequences of crosslinguistic variation in the state to lexical category mapping rule.

¹⁸ An analysis based on the notion of coercion (Zucchi 1998; de Swart 1998) is developed to account for these facts and for the Tongan pattern more generally in Koontz-Garboden (2004b).

¹⁹ In addition to sounding odd to me and to other native speakers of English I have consulted, *Google* searches like "the quickly cool" and "the quickly red" also fail to turn up plausible examples.

²⁰ It should also be pointed out, as an anonymous reviewer reminds me, that the theory allows for the existence of languages where property concepts are lexicalized as verbs, yet there is no single polysemous word with both property concept state and change of state senses. According to the reviewer, Mohawk is such a language, where there is an inchoative affix deriving change of state verbs from verbs denoting property concept states. The same is true for Central Alaskan Yup'ik (Jacobson 1984, 1995).

²¹ In the very brief sketch of the analysis above, I have couched the discussion in terms of events. It should be noted, though, that regardless of whether one adopts a Neo-Davidsonian semantics or not, there are semantic differences between verbs and their nominalizations that any formal theory must capture. These differences are made especially clear by Zucchi (1993). Koontz-Garboden (2004a) also gives some additional empirical arguments for a distinction. These empirical differences show that there must be a semantic difference between verbs and their associated nominalizations, regardless of how one chooses to formalize it. This distinction alone is sufficient to support my proposal that there is an additional operator in the decomposition of nominalizations when compared to the verbs they derive from. So, one could likely reject a Neo-Davidsonian semantic analysis and still accept my conclusions.

REFERENCES

- Baker, Mark C. (2003). *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.
- Bauer, Winifred, William Parker, and Te Kareongawi Evans (1993). *Maori. Descriptive Grammars*. London: Routledge.
- Beck, David (2002). *The Typology of Parts of Speech Systems: The Markedness of Adjectives*. Outstanding Dissertations in Linguistics. New York: Routledge.
- Boas, Franz, and Ella Deloria (1939). *Dakota Grammar*. *Memoirs of the National Academy of Sciences*, volume 23, number 2.

- Bowerman, Melissa (2004). Event categorization in adult and child language: The case of 'cutting' and 'breaking'. Paper presented at the Stanford University Construction of Meaning Workshop. April 15, 2004.
- Bybee, Joan, Revere Perkins, and William Pagliuca (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Chung, Sandra, and Alan Timberlake (1985). Tense, aspect, and mood. In Timothy Shopen (ed.) *Language Typology and Syntactic Description*. vol. 3. Cambridge: Cambridge University Press, 202–258.
- Comrie, Bernard (1976). *Aspect*. Cambridge: Cambridge University Press.
- Croft, William (1990). Possible verbs and the structure of events. In S. L. Tsonatzidis (ed.), *Meanings and Prototypes*. London: Routledge, 48–73.
- Croft, William (1991). *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press.
- Croft, William (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Cusihuaman, Antonio (1976). *Gramática Quechua: Cuzco-Collao*. Lima, Perú: Ministerio de Educación, Instituto de Estudios Peruanos.
- Dixon, Robert M. W. (1982). *Where Have All the Adjectives Gone?: and Other Essays in Semantics and Syntax*. The Hague: Mouton.
- Dowty, David (1979). *Word Meaning and Montague Grammar*. Dordrecht: Reidel.
- Enfield, Nick J. (2004). Adjectives in Lao. In Robert M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *Adjective Classes: A Cross-linguistic Typology*. Oxford: Oxford University Press.
- Foley, William A., and Robert D. Van Valin (1984). *Functional Syntax and Universal grammar*. Cambridge, UK: Cambridge University Press.
- Givón, Talmy (1984). *Syntax: A Functional-Typological Introduction*. Amsterdam: John Benjamins.
- Green, Thomas Michael (1999). *A lexicographic study of Ulwa*. Ph.D. dissertation, Cambridge, MA: Massachusetts Institute of Technology.
- Grimshaw, Jane (1993). Semantic structure and semantic content in lexical representation. MS, Rutgers University.
- Hale, Kenneth L., and Samuel Jay Keyser (1998). The basic elements of argument structure. *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect*, vol. 32. MITWPL, 73–118.
- Hale, Kenneth L., and Samuel Jay Keyser (2002). *Prolegomenon to a Theory of Argument Structure*. Cambridge, MA: MIT Press.
- Haspelmath, Martin (1993). More on the typology of inchoative/causative verb alternations. In Bernard Comrie and Maria Polinsky (eds.), *Causatives and Transitivity*. Amsterdam: John Benjamins, 87–120.
- Jackendoff, Ray (1983). *Semantics and Cognition*. Cambridge, MA: MIT Press.
- Jackendoff, Ray (1990). *Semantic Structures*. Cambridge, MA: MIT Press.
- Jacobson, Steven A. (1984). *Yup'ik Eskimo Dictionary*. Fairbanks, AK: Alaska Native Language Center.
- Jacobson, Steven A. (1995). *A Practical Grammar of the Central Alaskan Yup'ik Eskimo Language*. Fairbanks, AK: Alaska Native Language Center.

- Jaxontov, S. E. (1957). *Kategorija glagola v kitajskom jazyke*. Leningrad: Izd-vo Leningrad-skogo universiteta.
- Jespersen, Otto (1939). The history of a suffix. *International Journal of Structural Linguistics* 1, 48–56.
- Koontz-Garboden, Andrew (2004a). Nominalization and the lexical semantics of syntactic categories. MS, Stanford University.
- Koontz-Garboden, Andrew (2004b). Aspectual coercion and the typology of change of state predicates. MS, Stanford University.
- Koontz-Garboden, Andrew, and Beth Levin (2004). The morphological typology of change of state event encoding. Proceedings of the 4th Mediterranean Morphology Meeting, Catania, Sicily.
- Lefebvre, Claire, and Anne-Marie Brousseau (2002). *A Grammar of Fongbe*. Berlin: Mouton de Gruyter.
- Lehmann, Christian (1990). Towards lexical typology. In William Croft, Keith Denning, and Suzanne Kemmer (eds.), *Studies in Typology and Diachrony*. Amsterdam: John Benjamins, 161–185.
- Levin, Beth (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
- Levin, Beth, and Malka Rappaport Hovav (1995). *Unaccusativity: At the Syntax-Lexical Semantics Interface*. Cambridge, MA: MIT Press.
- Levin, Beth and Malka Rappaport Hovav (1998). Morphology and lexical semantics. In Arnold Zwicky and Andrew Spencer (eds.), *Handbook of Morphology*. Oxford: Blackwell, 248–271.
- Lieber, Rochelle (2004). *Morphology and Lexical Semantics*. Cambridge: Cambridge University Press. References to pre-publication manuscript.
- Megerdumian, Karine (2002). *Beyond words and phrases: A unified theory of predicate composition*. Ph.D. dissertation, Los Angeles: University of Southern California. Available at <http://members.cox.net/karinem/dissert.html>.
- Mohanan, Tara (1994). *Argument Structure in Hindi*. Stanford, CA: CSLI.
- Nedjalkov, V. P., and G.G. Silnitsky (1973). The typology of morphological and lexical causatives. In Ferenc Kiefer (ed.), *Trends in Soviet Theoretical Linguistics*. Dordrecht: D. Reidel, 1–32.
- Olsen, Mari Broman (1996). *A semantic and pragmatic model of lexical and grammatical aspect*. Ph.D. dissertation, Northwestern University. Published 1997 by Garland.
- Pinker, Steven (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.
- Prasithrathsint, Amara (2000). Adjectives as verbs in Thai. *Linguistic Typology* 4, 251–271.
- Pustejovsky, James (1995). *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Rappaport Hovav, Malka, and Beth Levin (1998). Building verb meanings. In Miriam Butt and Wilhelm Geuder (eds.) *The Projection of Arguments: Lexical and Compositional Factors*. Stanford, CA: CSLI, 97–134.
- Rappaport Hovav, Malka, and Beth Levin (2002). Change of state verbs: Implications for theories of argument projection. In *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA, 269–280.
- Smith, Carlota (1997). *The Parameter of Aspect*. Dordrecht: Kluwer (2nd ed.).
- Stassen, Leon (1997). *Intransitive Predication*. Oxford: Oxford University Press.

- Stump, Gregory T. (1998). Inflection. In Arnold Zwicky and Andrew Spencer (eds.), *Handbook of Morphology*. Oxford: Blackwell, 13–43.
- de Swart, Henriëtte (1998). Aspect shift and coercion. *Natural Language and Linguistic Theory* 16, 347–385.
- Talmy, Leonard (1985). Lexicalization patterns. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol. 3. Cambridge: Cambridge University Press, 57–149.
- Tatevosov, Sergej (2002). The parameter of actionality. *Linguistic Typology* 6, 317–401.
- Vázquez Soto, Verónica (2001). Some constraints on Cora causative constructions. In Masayoshi Shibatani (ed.), *The Grammar of Causation and Interpersonal Manipulation*. Amsterdam: John Benjamins, 197–244.
- Vendler, Zeno (1967). *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press.
- Vendler, Zeno (1968). *Adjectives and Nominalizations*. The Hague: Mouton.
- Vendler, Zeno (1975). Causal relations. In D. Davidson and G. Harman (ed.), *The Logic of Grammar*. Encino, CA: Dickenson, 255–261.
- Weber, David (1989). *A Grammar of Huallaga (Huánuco) Quechua*. Berkeley: University of California Press.
- Wetzer, Harrie (1996). *The Typology of Adjectival Predication*. Berlin: Mouton de Gruyter.
- Wierzbicka, Anna (1988). *The Semantics of Grammar*. Amsterdam: John Benjamins.
- Zucchi, Alessandro (1993). *The Language of Propositions and Events: Issues in the Syntax and the Semantics of Nominalization*. Dordrecht: Kluwer.
- Zucchi, Alessandro (1998). Aspect shift. In Susan Rothstein (ed.), *Events and Grammar*, Dordrecht: Kluwer, 349–370.

Department of Linguistics
Stanford University
Stanford, CA 94305-2150, USA

andrewkg@csl.stanford.edu

Pleonasm and hypercharacterisation

CHRISTIAN LEHMANN*

1. INTRODUCTION

Hypercharacterization¹ (German *Übercharakterisierung*) may be introduced per ostensionem: it is visible in expressions such as those of the second column of Table 1.

Table 1. Stock examples of hypercharacterisation

Language	Hypercharacterised	Basic	Surplus element
German	<i>der einzigste</i> 'the most only'	<i>der einzige</i> 'the only'	superlative suffix <i>-st</i>
Old English	<i>children, brethren</i>	<i>childer, brether</i>	plural suffix <i>-en</i>

While it is easy, with the help of such examples, to understand the term and get a feeling for the concept 'hypercharacterization', a precise definition is not so easy. The concept has, in fact, never been formally defined. Most of the time it has been taken for granted, and often it has been explicitly equated with neighbouring concepts. The concepts against which it must be delimited include pleonasm, tautology, redundancy, reinforcement and hypercorrection. Some of these are well established in certain scientific disciplines, others are no clearer than hypercharacterisation itself. I will therefore

1. start by defining pleonasm and delimiting it against neighbouring concepts;
2. articulate the concept by reviewing a set of suggestive cases;
3. define hypercharacterisation as a specific kind of pleonasm;
4. describe a set of cases of hypercharacterization within the framework outlined so far;
5. draw some conclusions which are of relevance for linguistic theory.

Since this procedure is not entirely deductive, but instead both based on an intuitive understanding of the concept of hypercharacterization and inspired by a variety of data, the definition resulting from it will be open to discussion and further refinement.

Pleonasm and hypercharacterisation are absolutely pervasive at different levels of style and at all the levels of the linguistic system, from discourse down to inflectional morphology and even to phonology.² Moreover, pleonasm has obvious rhetorical and poetic functions which would deserve a study of its own. In this paper, the approach is purely linguistic: the structure and linguistic (communicative, semantic, grammatical) function of hypercharacterised expressions in syntax and morphology will be studied.

From among the concepts akin to hypercharacterization in its semantic field, ‘hypercorrection’ must be separated out. **Hypercorrection** is the use of an expression X, in an attempt to speak correctly, in a context C where the norm forbids it, the background being that X does not occur at all in unconstrained colloquial speech, but is required by the norm in certain contexts other than C. Hypercorrection is frequent in situations where the speaker feels it would be important to conform to the norm, for instance in language acquisition. A typical example is *Whom shall I say was calling?* Hypercorrection has nothing to do with the topic under study here.

2. PLEONASM

2.1. Definition

The most general concept in our domain is **redundancy**. A message is redundant iff it contains such elements which contribute nothing to the information not already conveyed by the rest of the message. Repeating an utterance is redundant, and much of grammatical agreement, as in German *eine alte Eule* (INDEF:F.SG old:F.SG OWL.F.SG) as compared to English *an old owl*, is redundant.

However, a simple information-theoretical conception of redundancy does not lead us very far in the analysis of linguistic structure. In particular, a simple-minded conception of redundancy where ‘redundant’ implies ‘superfluous’ and therefore ‘useless’ would be inadequate. Redundancy fulfills functions at all levels of communication and grammatical structure. At the highest level (which is well recognised in information theory, too), redundancy ensures understanding even under difficult communication conditions. At the level of communicative intentions, it may be employed to overwhelm or impress the receiver. Redundancy may have poetic functions in the sense of Jakobson’s (1960) projection of paradigmatic relations onto the syntagmatic axis. And last but not the least, the combination of partly or wholly synonymous elements may fulfill various grammatical functions, as we shall see in section 3.1. Thus, ‘redundant’ does not by any means entail ‘functionless’.

The concepts of **pleonasm** and **tautology** have been current in rhetoric, linguistics and philosophy since antiquity. At the beginning, we can exclude the logical approach and with it the meaning of the term *tautology* in propositional logic, where it refers to a proposition that is always true independently of the truth values of its constituents, as e.g. *It will rain or it will not rain*. In rhetoric and linguistics, the two terms have been treated as interchangeable and been variously delimited against each other with about equal frequency. The following properties have usually played a criterial role:

- Tautology and pleonasm are kinds of redundancy.
- Both terms are used as *nomina acti*, referring to linguistic acts (specifically to rhetorical figures),³ and as *nomina patientis*, referring to expressions resulting from such acts.
- Both of them refer to complex expressions some of whose constituents bear some semantic similarity.
- Traditionally, the constituents in question are words or phrases.
- In the typical case, the similarity in question obtains between just two constituents.
- In the clearest and extreme case, one constituent is synonymous with the other. That case is called tautology.
- In less extreme cases, the meaning of one constituent entails the meaning of the other without being identical to it. ‘Pleonasm’ may either be restricted to this relationship or be used as a cover term for both kinds of semantic relation.

Tautology may be illustrated by the examples in E1:

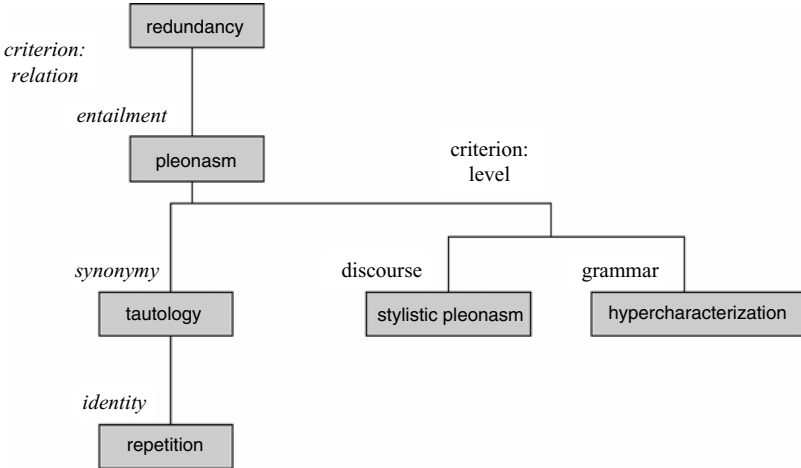
- E1. a. each and every, null and void, useless and unnecessary
 b. German *plötzlich und unerwartet* ‘sudden and unexpected’
 c. business is business, enough is enough

Apart from their semantic properties, many collocations like those of E1 also have poetic qualities, to be seen in such features as alliteration and meter. Such expressions may in fact even be motivated to some extent by the analogical model of binomials like *kith and kin*.

Pleonasm will here be used to include tautology as a special case. In general, a pleonastic expression contains constituents—typically two—one of which implies—technically: entails—the other. Thus, the meaning of the latter constituent is part of the meaning of the former. For instance, the meaning of *return* is roughly ‘go back’. *Return back* is pleonastic because the meaning of *back* is included in, or implied by, the meaning of *return*. This shared semantic component in pleonastic expressions will be called the **focal component**.

We will drop the traditional implicit presupposition that the expressions in question are words or phrases in order to be able to apply these concepts at the morphological level. **Hypercharacterization** will be defined in section 2.2.3 as pleonasm at the level of grammar.⁴ **Repetition** may be regarded as a special kind of tautology where the relation between the elements involved in the process is (type-)identity. We will have occasion to come back to it in section 2.5 At this point, the relation between the concepts introduced so far may be visualised as in S1:

S1. Conceptual field of pleonasm



We may now refine the definition of pleonasm:

An expression $E_1 + E_2 \dots E_n$, is **pleonastic** iff it contains a meaning component F that is included in the meaning of more than one E_i . Typically, F is the intension of one E_i and properly included in the intension of $E_{j \neq i}$; and $E_1 + E_2 \dots E_n$, reduces to a binary construction $E_1 + E_2$.

The notion of a component Q being properly included in a meaning P can be explicated as semantic entailment: $P(x)$ entails $Q(x)$. This formula would directly fit such examples as *Essential* (x) \rightarrow *wesentlich* (x), but would not apply in a straightforward way to others such as *return* (x) \rightarrow *back* (x). We will therefore assume that the pleonastic character of an expression $E_1 + E_2$ is tested by an implication $E_1 \rightarrow E_1 + E_2$. For instance, *return* (x) \rightarrow *return back* (x), and *example* (x) \rightarrow *specific example* (x).

Now for any construction $E_1 + E_2$, the entailment $E_1 \rightarrow E_1 + E_2$ is unusual and defines its pleonastic character, while the reverse entailment $E_1 + E_2 \rightarrow E_1$ is always valid provided the construction $E_1 + E_2$ is at all semantically compositional. Now $((p \rightarrow q) \& (q \rightarrow p)) \leftrightarrow (p \leftrightarrow q)$. In other words, what we have is **synonymy** of a pleonastic construction with one of its members. In this light, the difference between a pleonastic and a tautological construction consists in the fact that in a pleonastic construction, one member is synonymous with the construction, while in a tautological construction, each member is synonymous with the construction.

In ancient rhetoric, the hyperonym for pleonasm is *adiectio*, i.e. the addition of linguistic material. Its opposite is *detractio*, the suppression of linguistic

material,⁵ which we may translate by **pregnancy** (conciseness). The publicity slogan in E2 is a relatively recent example:

E2. Deutschlands meiste Kreditkarte
GERM ‘Germany’s most [common] credit card’

As we shall see below, pleonasm and hypercharacterisation are of interest to the grammarian because they may be at the origin of new grammatical structure. Pregnancy is the opposite in this respect, too: in order to be viable, it must rely on established linguistic structure and exploit it to the utmost. Pregnancy will not occupy us further here.

Finally, a methodological remark must be made. In section 2, many examples of pleonastic expressions are adduced without individual analysis. Some of them have both pleonastic and plain uses. For instance, *repeat* means ‘do something for the n th time, with $n > 1$ ’. Therefore, the literal meaning of *repeat again* is ‘do something for the n th time, with $n > 2$ ’. If the expression is used in this sense, it is not pleonastic. Purists will restrict their use of it to this sense. However, it suffices to observe actual speech (or to do an internet search) in order to become aware that most uses of *repeat again* actually mean ‘do something for the second time’, being thus included in the meaning of *repeat*. Consequently, while it may be observed that several of the examples below are not necessarily pleonastic, this does not invalidate the point that they do have pleonastic uses; and that is all that is necessary for the argument.

2.2. Structural types of pleonastic constructions

Since pleonasm is a purely semantic (or stylistic/rhetorical) concept, it implies very little about the structure of pleonastic expressions. These are therefore structurally quite heterogeneous. At least the following criteria are useful in their classification.

2.2.1. Grammatical level of the pleonastic construction

The principal distinction here is between

- a syntactic construction, as in *resulting effect*,
- and a word (form), as in German *bestmöglichst* ‘best (most) possible’.

Of course, different syntactic levels may be distinguished, if necessary; for instance, *see with one’s eyes* is a verbal, *resulting effect* is a nominal. There could, in principle, be pleonastic sentences, too, like *this whale is a mammal*; but they probably occur chiefly as examples of analytic sentences in logic books.

2.2.2. Nature of the elements expressing the focal component

The criterion of section 2.2.1 may be applied again to the focal component of a pleonastic expression itself. Stepping down the hierarchy of grammatical levels, it may be expressed by

- a syntagma, as in *fly through the air*,
- a word (form), as in *return back*,
- a stem, as in German *Eichbaum* ‘oak tree’,
- a derivational morpheme, as in German *Reformierung* ‘reform’,
- an inflectional morpheme, as in *spaghettis*.

In the prototypical pleonastic construction, the focal component is expressed twice, once by a **dedicated unit** (underlined in the above examples) whose meaning is exhausted by the focal component, once as part of the meaning of another unit. It is, however, not excluded that the focal component is represented by a dedicated unit more than once. For instance, in OE *children*, plural is expressed by each of the suffixes *-(e)r* and *-en*. Consequently, this parameter may be applied separately to each of the occurrences of the focal component, leading by itself to a cross-classification of pleonastic constructions. At the morphological level, naturally tautologies of derivational and of inflectional morphemes are of special interest.

2.2.3. Relation between elements containing the focal component

Given that the dedicated unit and the unit including the focal component are members of a construction, they are in some structural relation. At the higher levels of grammar, this will be one of the generic syntactic relations of

- sociation, as in German *mit Fug und Recht* ‘with full right’,
- government, as in *dream a dream*,
- modification, as in *return back*.

In a **sociative** pleonastic construction, the two related elements are generally synonymous. The construction is then a tautology, as in E1. In a **governing** construction, the dependent is by definition selected by the head. The meaning of the latter then includes a selection restriction that embodies a hyperonym of the dependent. For instance, the meaning of *mow* includes as a selection restriction a component that represents (an area covered by) a uniform collection of plants of a certain shape, which is a hyperonym to such nouns as *lawn*. This is also true for such cognate object constructions as *dream a horrible dream*, *sleep a restful sleep*. Governing constructions, cognate or otherwise, are generally not regarded

as pleonastic,⁶ probably because the verbal selection restriction is unavoidable. There is, however, an extreme variant of the cognate object construction where the object is not further specified: In Korean, verbal concepts such as ‘sleep’ and ‘dream’ are obligatorily rendered by a cognate object construction of the form ‘sleep a sleep’, ‘dream a dream’ (S.-R. Ryu p.c.). As far as the semantic relation between the two units is concerned, it seems to be a matter of definition whether it should be regarded as pleonastic or tautological.

The core of pleonastic constructions is constituted by **modificative** constructions. Typical examples have one of the following syntactic structures⁷:

- a nominal consisting of a head noun containing, and an adjective attribute expressing, the focal component, as in *original source*, *free gift*; German *die wesentlichen Essentials* ‘the substantial essentials’;
- a verbal consisting of a verb containing, and an adverbial expressing, the focal component, as in *fly through the air*, *return back*;
- an adjectival consisting of an adjective containing, and an adverbial expressing, the focal component, as in *potentially capable*, *more than unique*.

These constructions have a modifier in common that is syntactically optional and semantically redundant. It is, however, the modifier, not the head, that codes the focal component more explicitly.

The preceding classification is restricted to the syntactic level, i.e. it is a subclassification of the first class of section 2.2.1. For present purposes, it does not seem necessary to take up the issue of grammatical relations at the word level; the categorical distinctions introduced in section 2.2.2 will suffice.

This discussion amounts to a recognition that in the prototypical pleonastic construction, the dedicated unit modifies the unit that properly includes the focal component. We will treat this as an empirical generalization over many examples from English, German and a couple of other European languages, based on the semantic definition of pleonasm given in section 2.1. Although pleonastic constructions are typically modificative, it is probably wise not to elevate this to the status of a definitory criterion, because then the concept would consist of purely semantic and purely structural criteria which seem to be essentially independent.⁸

Intuitively, **hypercharacterization** is pleonasm at the level of grammar. We can now refer this to the structural distinctions introduced in section 2.2.2. Hypercharacterization may then be defined as that kind of pleonasm where the focal component is expressed by an inflectional or derivational morpheme. This is taken as criterial no matter whether this morpheme can be identified as the surplus element in the construction and whether the other occurrence of the focal component in the construction takes the form of a dedicated unit, too.

The methodological upshot of section 2.2 is, then

- The classification of 2.2.1 is a prerequisite for the classification of section 2.2.3.
- The classification of 2.2.2 is presupposed for the delimitation of hypercharacterisation.
- The classification of section 2.2.3 yields an empirical generalisation over pleonastic constructions.

As a consequence of this, hypercharacterised constructions will be just as heterogeneous structurally as pleonastic constructions in general. This should be kept in mind for section 3.

2.3. *Asymmetry in pleonasm*

Given a tautological expression $E_1 + E_2 \dots E_n$, $E_1 \dots E_n$ each make an equal contribution to the overall tautological character of the expression. If $E_1 \dots E_n$ are linked by a sociative relation, we can choose any one of them at random, omit the others and still have the same total meaning. For instance, we can easily reduce *useless and unnecessary* to either *useless* or *unnecessary*. In this sense, binary tautologies are symmetric.

Now the question arises whether non-tautological pleonastic expressions are semantically symmetric in the sense that the focal component can be omitted either in the head or in the modifier, or whether they are asymmetric in the sense of having a legitimate core and a superfluous periphery. Since non-tautological pleonastic expressions generally have a dependency structure, they are structurally asymmetric, so that one can leave out the modifier, but one cannot simply leave out the head. We will therefore assume that the methodological counterpart to leaving out the dependent in an expression such as *wesentliche Essentials* is to replace the head by a hyperonym that does not contain the focal component, e.g. *wesentliche Punkte* ‘essential points’. Semantically, then, pleonasm might be symmetric in the sense explained.

An examination of a large set of data—some of which are adduced in section 2.4—shows that the procedure of replacing the head by an appropriate hyperonym is not viable in many cases because there is no such hyperonym. *Exactly the same, potentially capable, original source* illustrate this point. On the other hand, omission of the modifier is always possible both syntactically and semantically. It is also the simpler procedure. I will therefore assume that non-tautological pleonastic expressions are asymmetric not only structurally—by virtue of their dependency structure—but also semantically in the sense that the syntactic head is the semantic core and the syntactic modifier is the surplus element that renders the expression pleonastic. In other words, the implication

used to operationalise the definition of pleonasm will be directed in this way: ‘head → dependent’ or else ‘head → head + dependent’. We will see in section 2.4.6 that this assumption is not entirely unproblematic.

2.4. The motivation for pleonasm

Everywhere in linguistic structure, a movement descending the levels of structure from discourse down to the morpheme correlates with a decrease in the freedom of selection and combination of the units of those levels. At the highest level, these operations are motivated by semantic, stylistic, pragmatic, etc. considerations, i.e. by considerations concerning the cognitive and communicative aims the speaker is pursuing. At the lowest level, such motivations no longer exert any influence, because it is the linguistic system that dictates them. If hypercharacterization differs from other kinds of pleonasm only by the lower level at which it plays, it is foreseeable that there will be a variety of extra-structural motivations for pleonasm in general, which will be relevant only in a diluted and weakened form for hypercharacterization.

2.4.1. Intensity

Given a predicate that may be true of its argument to different degrees, there may be a default value for that predicate for that class of arguments, and there may be particular individuals that the predicate is true of to a higher degree or even to the highest conceivable degree. To express such a situation, ascription of that predicate to that argument may be intensified. E3 contains some relevant examples.

- E3. a. bitterly cold, boiling hot
 b. German *bärenstark* (bear-strong) ‘husky’, *strohdumm* (straw-dumb) ‘empty-headed’

The expressions chosen for intensification are often based on exaggeration. Probably, somebody calling a person *bärenstark* is not committed to a bet that that person could stand a test against a bear. But the concept of intensification is indifferent to the validity of such literal interpretations. It suffices that *bärenstark* is not synonymous with *stark*, but assigns its argument a position on the relevant scale that is above the default. This kind of intensification will be called **polar extreme enforcement**.

This analysis implies that polar extreme enforcement is not a kind of pleonasm in the sense defined in section 2.1. It is nevertheless necessary to start our treatment of motivations for pleonasm with intensification, because the conditions for intensification are often loosened. That is, intensification often treats

predicates as gradable that are inherently absolute. We are coming to this in the next section.

2.4.2. *Emphasis*

Like most kinds of redundancy, pleonasm is often regarded as bad style. Ancient rhetoric did, in fact, classify it as a kind of solecism (Lausberg 1990, §502). It is also true that unwitting pleonasm violates the Gricean maxim of quantity and may insofar be irritating. On the other hand, many pleonastic expressions are evidently no unwitting slips, but are meant to lend emphasis to the message. Examples of such **emphatic pleonasm** are given in E4; moreover, all of the examples of tautology given in E1 are motivated by emphasis.

- E4. a. completely deaf, perfectly legitimate, surrounded on all sides, diametrically opposed, coal-pitch-black, totally unnecessary
- b. exactly the same, exact replica, completely empty, more than unique/extremely unique, I have been there myself, with these very eyes I saw it

The emphatic character of the expressions in E4 is verifiable by a test: in all of them, the modifier may receive emphatic stress.

Analysing the examples in E4, we see at once that in most of them the modifier is an intensifier. Emphatic pleonasm may be subdivided as follows:

1. **Default confirmation:** In E4(a), the head may be interpreted more or less liberally. However, what the modifier says is the default interpretation of the head, anyway, and insofar the expression is pleonastic. The intensifier confirms this default interpretation, forestalling a possible moderate interpretation of the head.
2. **Insistence on focal component:** In E4(b), the meaning of the head is absolute in the sense that it applies to something in a yes-or-no fashion rather than to some extent. Consequently, the intensifier cannot do more than underline the significance of what the head implies.

Many heads in emphatic pleonasm admit of a less-than-perfect reading and, correspondingly, of an attenuative modifier. Thus, expressions like *inexact replica*, *almost the same*, *surrounded on almost all sides* are unobjectionable. They presuppose the possibility of cancelling the perfect interpretation of the head. In this perspective, intensification has a purely semantic justification in the cases of section 2.4.1 (E3); it is semantically motivated to some extent in

‘default confirmation’ (E4(a)) and only stylistically motivated in ‘insistence on focal component’ (E4(b)). In other words, the three varieties of intensive and emphatic pleonasm seen so far may be ordered on a scale from purely semantic to stylistic motivation as in S2:

S2. *Motivation of intensive and emphatic pleonasm*

polar extreme enforcement > default confirmation > insistence on focal component

Insistence on the focal component is legitimated by analogy to default confirmation, and default confirmation is legitimated by analogy to polar extreme enforcement. Emphatic pleonasm sails under the flag of polar extreme enforcement. The latter is just a kind of intensification, which, in itself, is not (yet) pleonastic. Thus, S2 symbolises the emergence of pleonasm; its central position may be taken to mark the pole of incipient, unobtrusive pleonasm.

2.4.3. *Rhematicity*

Functional sentence perspective is gradual in many ways. One of these is the fact that the difference between thematic and rhematic material is greater at higher levels of syntactic complexity and shrinks down to the lowest level, viz. the level of the word form. Now if I have a sentence in which the focal component is to be rhematic, this will not be sufficiently represented by the word of whose meaning it is but a component. The modifier codes the focal component separately so that it can receive rhematic status in the utterance. This is the typical motivation for expressions such as those of E5.

E5. specific examples, sudden impulse, little baby, original source, free gift, pre-planning/forward planning, potentially capable, may possibly

Here again, the focal component may, in some cases, be a defeasible implication of the meaning of the head noun. Thus, the concepts of a *big baby*, of an *intermediate source* or of a *Danaans’ gift* are not self-contradictory. As in the default confirmation variety of emphatic pleonasm, the modifier here makes explicit a component that is part of the default interpretation of the head.

On the other hand, non-pleonastic uses of some of the phrases in E5 are possible. One might construct a text that meaningfully opposes *potentially capable* to *actually capable*. The point here is that these phrases are generally used in a pleonastic fashion where *potentially capable* is not opposed to *actually capable*, but just means *capable*.⁹

2.4.4. Safety

The set of examples in E6 shows another motivation for pleonasm:

- E6. bound affix, handwritten manuscript, joint cooperation, collaborate together, circulate around, postponed until later, vacillating back and forth

If you are not sure whether the head actually possesses the focal component, you play it safe by expressing the component separately in a modifier. We will call this **safety pleonasm**.¹⁰ There are several fields in which safety pleonasm appears to be commonly operative. An especially important one is loanwords, as in Table 2.

Table 2. Safety pleonasm in loan words

Language	Expression	Comment
English	Rio Grande river	Spanish <i>rio</i> ‘river’
English	Sahara desert	Arabic <i>sahara</i> ‘desert’
English	Mount Fujiyama	Japanese <i>yama</i> ‘mountain’
Italian	Mongibello	Sicilian <i>mon</i> = Arabic <i>gebel</i> ‘mountain’
German	die La-Ola-Welle	Spanish <i>la ola</i> = German <i>die Welle</i> ‘the wave’

In a speech community, there is variation with respect to command of the donor language of loans. Those that borrow an expression may be assumed to have some knowledge of the meaning and even structure of the loan. To other speakers of the recipient language, the structure of such foreign names is either unknown or irrelevant. For these, *rio* is not another word for ‘river’, but part of the proper name *Rio Grande*. To this extent, such formations are not really pleonastic in the recipient language.

In general, safety pleonasm is a symptom of instability of variation, at the level of the individual or of the speech community. For some speakers, *handwritten manuscript* is clearly pleonastic, while for others it is not, but just means ‘handwritten paper’. Safety pleonasm therefore indicates that at least part of the speech community does not feel that the base of the expression is (already) characterised for the focal component.

2.4.5. Verbosity

Yet other examples evince a desire to equip a naked noun, verb or adjective with a companion so that it need not stand alone. The word alone seems too weak. E7 contains a couple of relevant examples.

- E7. past experience, resulting effect, unexpected surprise, return back, sink down, fall down, repeat again, fly through the air

In some cases, this horror vacui may be motivated purely phonologically, by reasons of rhythmic euphony. Observe also that several of the heads are monosyllabic.¹¹ This variety may be called **phatic pleonasm**. The modifier is not stressed and in most cases cannot even be stressed because there is no possible contrast.

2.4.6. Concord

A pleonastic combination may become usual to the extent that it is less marked than its non-pleonastic counterpart which lacks the modifier. The expansion of *repeat/wiederholen* to *repeat again/noch einmal wiederholen* is almost an automatism. To the extent that there is a rule that requires that modifier and head agree in the focal component, we have a kind of semantic concord at the syntactic level. That some such mechanism must be operative becomes more plausible if this rule manifests itself at the morphological level. This may be seen in the following two sets of examples.

A variant of the pleonastic nominal appears in diminutive expressions of the kind illustrated in E8–E10:

E8. Gyricons SmartPaper besteht im Wesentlichen aus kleinen zweifarbigen Kügelchen, die in einer dünnen flexiblen Plastikschiicht eingebettet sind. Die Kügelchen drehen ... (c't 1/2004:22)

GERM 'Gyricon's SmartPaper essentially consists in small dichromic mini-balls embedded in a thin flexible plastic layer. These mini-balls turn ...'

E9. Dieser [Chip] ist mit seinen 1024 × 576 kleinen Spiegelchen auf die hierzulande übliche PAL-Norm ... abgestimmt. (c't 1/2004:22)

GERM 'With its 1024 × 576 small mini-mirrors, this [chip] is attuned to the PAL norm which is standard in this country.'

E10. kleine vorgelagerte Inselchen (MDR Kultur, 31 October 2004)

GERM 'small islets situated in front'

Such examples share with the foregoing types the fact that the focal component is expressed more explicitly by the syntactic modifier than by the head. However, something similar to agreement appears to be operative in such a combination, in that once we have chosen the adjective *klein* as a modifier, diminution of the head noun is almost an automatism.¹² We therefore call this variant **concord pleonasm**. In this and the following case, concord pleonasm manifests itself at the morphological level: The focal component is not just a semantic feature of the lexical meaning of the noun, but expressed separately by the diminutive morpheme. Because of this, either the syntactic or the morphological modifier is freely omissible, with little difference in meaning.

The same phenomenon may be illustrated with female sex marking in German. E11 illustrates the range of phenomena relevant here.

- E11. a. Im Jahr 1884 wird Sofja Kovalevskaja in Stockholm die erste weibliche Professorin Europas. (<http://ruprecht.fsk.uni-heidelberg.de/ausgaben/58/ru05.htm>, 3 February 1999)
- GERM 'In 1884, Sofja Kovalevska in Stockholm becomes the first female she-professor of Europe.'
- b. Liselotte Welskopf-Henrich...1960 erster weiblicher Professor an der Berliner Alma Mater (<http://home.t-online.de/home/ametas/welskopf.htm>, 16 June 1999)
- 'Luise Welskopf-Henrich... first female professor at the Alma Mater of Berlin in 1960'

In E11(a), the female sex of the referent is expressed twice, by the adjective attribute and by the female derivational suffix. In E11 (b), it is only expressed by the attribute. Moreover, the NP has feminine gender in E11(a) and masculine gender in E11(b). A search on the web (Google, 12 February 2004) turns up 57 examples of *weibliche Professorin* and 4 examples of *weiblicher Professor*.

The motivation of this kind of pleonasm is intricate. In a diachronic perspective, one can be sure that the numerical ratio would have been inverse if this web search had been executed 50 years ago. At that time, *weibliche Professorin* would either have seemed unnecessarily redundant or else it would have meant 'feminine [i.e. womanly] professor'. In contemporary German, female human beings are mostly designated by nouns of feminine gender and, if possible, derived with the female suffix. In an NP containing the adjective *weiblich* as a modifier, this rule is almost obligatory, as the numerical ratio shows.¹³ Thus, the use of the adjective *weiblich* in sentences like E11 is contrastive, while the use of the female suffix *-in* is due to concord pleonasm.

In the varieties of pleonasm analysed before, the focal component is expressed separately by the modifier, but is just a semantic component of the head. Pleonasm in such constructions is thus a purely semantic, not a morphological phenomenon. In morphological concord pleonasm, the focal component receives separate expression by a bound morpheme on the head, fulfilling thus the condition for hypercharacterization. In other words, the focal component is expressed twice separately, so that its two occurrences are near-synonymous. Given that one of them is a word, while the other is a derivational morpheme, they can hardly be totally synonymous. However, morphological concord pleonasm as illustrated here constitutes a transition from a purely semantic to a morphological phenomenon.

Both the syntactic modifier and the bound morpheme are optional, but to the extent that the focal component is more explicitly coded at the syntactic

level, its morphological coding is perceived as a contingent phenomenon. From a processing perspective, the asymmetry postulated for pleonastic constructions in section 2.3 may be reversed here: In the constructions of sections 2.4.2–2.4.5, it appears that the speaker first selects the head and then expands it into a pleonastic construction, succumbing to one of the motivations discussed there. Contrariwise in morphological concord pleonasm, it appears that the speaker first selects the syntactic modifier together with an unmarked version of the head and only then pleonastically marks the latter for the focal component. The morphological marking of a feature of one word on another member of its construction is like agreement. However, concord pleonasm differs from agreement not only in being largely optional, but also in its direction: inside the noun phrase, agreement works from the head towards the modifier,¹⁴ while in the noun phrases of the present section, concord works from the modifier towards the head.

2.4.7. *Summary of motivations*

The kinds of motivation for pleonasm that we have distinguished differ in their strength and may accordingly be arranged on a scale, as in S3 (which embodies S2 at its start):

S3. *Strength of motivation of pleonasm*

(intensive >) emphatic > rhematic > safety > phatic > concord pleonasm

There are several criteria for identifying the motivation of a given pleonastic construction:

- **Entailment:** At the end of S3, the meaning of the modifier is entailed by the meaning of the head. At the start of the scale, the former merely pins down a possibility provided by the latter.
- **Usualness:** At the start of S3, pleonasm is marked; at the end, it becomes the normal way of expression.
- **Contrast:** Contrastive stress on the modifier is normal at the start of S3 and then recedes down the scale until it becomes outright impossible at the end.

Although the phenomena analysed so far abide at the lexical-syntactic and derivational levels, it may be seen that these three criteria are reminiscent of the criteria that define grammaticalisation (cf. Lehmann 2002, chapter 4). That is, increasing entailment is an early stage of desemanticisation, usualness is a milder form of obligatoriness, and loss of the ability to contrast is the prerequisite for cliticisation. It is as if S3 were a pre-stage of a grammaticalisation scale. And, of

course, a given expression or construction may move down S3 from left to right. We will come back to this in section 3.1.1.

As we have seen in section 2.4.1, there is a kind of intensification which we called polar extreme enforcement, illustrated by expressions like *boiling hot*, which insist that the extreme pole of a scale is being referred to and which are not yet pleonastic. At the opposite pole, pleonasm becomes similar to syntactic agreement. At the beginning of S3, semantic and pragmatic considerations condition the choices; at the end, usage and grammar start to dictate them. The scalar representation of pleonasm as in S3 makes us see how it ties in with related phenomena.

The classification of the five kinds of pleonasm is, of course, not clear-cut. *Little baby*, for instance, need not be rhematic in every case, often it will be merely a phatic pleonasm. Again, the various motivations do not exclude each other. *Fall down* and *repeat again*, which I classified as phatic pleonasm, are almost obligatory collocations, a feature that they share with concord pleonasm.

Finally, many pleonastic expressions share with the tautologies reviewed in E1 the property of being phraseologisms. In such cases, pleonasm is not a collateral result of a constellation at the syntactic or the discourse level, but something built into the lexicon of a language.

2.5. Repetition

As anticipated in section 2.1, one can conceive of repetition as a particular kind of tautology where the synonymous elements are identical. And repetition does have some of the functions of pleonastic expressions that we saw in section 2.4. It may have the same function as tautology—insistence on the focal component, as illustrated by the German adverbials in E12.

- E12. a. immer und ewig ‘for ever and ever’
 b. immer und immer (ditto)

E12(a) is synonymous with E12(b) (although their use is slightly different). E12(a) is an example of tautology like the phrases of E1. E12(b) differs from those cases only by the formal identity of the synonymous items.

Repetition may have a purely phatic function, as in E13.

- E13. That is totally impossible—totally impossible.

Thus, the scale of S3 may be applied to repetition as to non-identical redundancy. As has been indicated above and as will be argued further in section 3.1.1, pleonasm has a grammaticalised manifestation, which is hypercharacterization. In the same way, reduplication may be seen as grammaticalised repetition (cf. Marantz and Wiltshire 2000:558). We will therefore consider repetition as a

limiting case of the redundancy phenomena analysed here. A couple of examples involving repetition will come up; however, reduplication and iteration will not be treated per se.

A peculiar kind of repetition may be seen in abbreviation elaboration, which is, at the same time, a kind of hyponym compounding and therefore treated in section 3.4.

3. HYPERCHARACTERIZATION

In hypercharacterization, the focal component is expressed by an inflectional or derivational morpheme (cf. section 2.2.2). It should be born in mind that the concept of hypercharacterization imposes no conditions on the expression of the second occurrence of the focal component constitutive of any pleonasm. Thus, English *more easier* is hypercharacterised by the adverb *more* combining with a morphological comparative form; but so is German (*der*) *einzigste* ‘the most only’, where the superlative suffix repeats the idea of singling out one individual fulfilling a relevant condition, which is also part of the concept of *einzig*. We already saw some relevant cases of hypercharacterization in morphological concord pleonasm (section 2.4.6). Other typical examples include, in the domain of inflection, the English *children* and *brethren* adduced in Table 1, and in the domain of word-formation, the German examples given in Table 3.

Table 3. Hypercharacterization in German word-formation

Hypercharacterised	Basic	Meaning
aufoktroyieren	oktroyieren	impose, force upon
wegeskamotieren	eskamotieren	retract, play down

Given that hypercharacterization is a kind of pleonasm, it may be motivated in the same ways seen before. The German preverbs are added to their bases in order to make explicit a meaning component commonly expressed by these preverbs, as in the near synonyms *aufzwingen* and *wegschaffen*, respectively. Analogy is clearly at work here. Since the bases are French loans of whose meaning one cannot be entirely sure, the motivation of these formations combines rhematicity with safety.

At the grammatical level, pleonasm concerns linguistic theory in a much more vital way. All of the expressions analysed in section 2 are syntactically and semantically well-formed, so that they do not constitute a problem for either syntax or formal semantics.¹⁵ Their peculiarity may thus safely be relegated to stylistics. At the level of morphology, however, we deal with specific operators combining specific structural features with their operands in a rule-governed

way. Now how can OE *brether* and *childer* take a plural affix if they are already marked for plural? Any theory of grammar that constructs complex forms in a compositional fashion by combining an operand of a certain category with an operator that transforms it into a resulting expression of another category has a serious problem here.¹⁶ We shall come back to these problems below and first review a couple of examples of hypercharacterization in order to familiarise ourselves with the phenomenon.

3.1. Hypercharacterization in syntax

In doing this, we can take up where we left off in section 2.4.6, viz. at the level of syntactic concord.

3.1.1. Personal agreement

In Latin just as in the written norm of several Romance languages, the personal ending contained in the finite verb form is sufficient reference to the subject; thus neither grammar nor semantics require an overt subject. In several spoken varieties, and in French even in the written standard, the subject pronoun is obligatory. Table 4 visualises the situation in two Romance languages in a simplified way.

Table 4. Pronominal subjects in Romance ‘we live’

Language	Pronominal subject			
	Without		With	
	Example	Value	Example	Value
Italian	<i>viviamo</i>	normal	<i>noi viviamo</i>	emphatic
French	<i>vivons</i>	ungrammatical	<i>nous vivons</i>	normal

In the right-hand column of Table 4, we have hypercharacterization of the subject reference. In Italian, its motivation is emphasis, while in French, it is sheer concord. As is well known, the French construction evolves by grammaticalization of a Proto-Romance construction that is reflected in Italian. This shows that once we concentrate on grammatical pleonasm, the scale S3 becomes a manifestation of a general grammaticalization scale.

At the start of the development, the verbal clause is expanded by an emphatic subject pronoun. The Italian line of Table 4, read from left to right, illustrates this process. This kind of pleonasm comes under the concept of **reinforcement** (cf. Lehmann 2002, chapter 2.5). At this stage, the subject pronoun is clearly

the surplus element of the pleonastic construction. In the further course, the reinforcement of the subject reference no longer works at the communicative, but merely at the syntactic level, i.e. the pronoun is needed to ensure the subject reference in the first place. This shows that the concept of pro-drop occasionally used to describe the Italian situation is misconceived with regard both to the diachrony and to the function of the construction. It is only from an anglocentric perspective that Italian drops some element that should be there. On the contrary Italian optionally and French obligatorily add a subject pronoun.

As concord hypercharacterization is grammaticalised to mere agreement, redundancy seems to be introduced into the grammar. However, in this course it loses its pleonastic function at the communicative level and gets a new function at the structural level, in the marking of syntactic relations.

3.1.2. *Spatial relations*

Another area where hypercharacterisation is very frequent in the languages of the world is spatial deixis and spatial relations. E14 illustrates four variants of a sentence containing the deictic *da* 'there' in Standard German (a), Northern Colloquial German (b), Bavarian (c) and Alemannic German (d).

- E14. a. Davon weiß ich nichts.
 GERM b. Da weiß ich nichts von.
 c. Dadavon weiß ich nichts.
 d. Da weiß ich nichts davon.
 'I know nothing of it.'

All of these variants are attempts to solve the problem of the topicalisation of the pronominal complement of the preposition. Both Standard German pied piping and Northern Colloquial German preposition stranding solve it without redundancy. Southern dialects avoid preposition stranding while feeling that mere pied piping is communicatively insufficient and the topical pronoun must be present separately.

E15 illustrates three different uses of the spatial relator *ex* 'out of', all of which are from Classical Latin. In E15(a), the relator appears only as a preposition; in (b), it appears only as a preverb; in (c), it is used pleonastically both as a preposition and a preverb.

- E15. a. ex urbe fugere
 LATIN out.of town:ABL.SG flee:INF
 b. urbe effugere

- c. ex urbe effugere
 ‘to flee out of town’

German constructions of the kind illustrated in E16 appear to be structurally similar to E15(c).

- E16. a. an etwas anschließen
 GERM ‘to adjoin to something’
 b. auf jemanden aufpassen
 ‘to watch over somebody’

They differ from the Latin construction in several respects. One that we can forgo relates to the fact that the German compound verb is separable. What is of more importance is that the compound verb governs the preposition of its complement.¹⁷ This pleonasm is therefore completely grammaticalised or lexicalised.

Both redundant demonstrative topicalisation and preverb–preposition concord are cases of concord hypercharacterization; but they are special in that they involve repetition of the same element. Since hypercharacterization plays at the level of grammar, the choice of synonymous morphs decreases, so that pleonasm often takes the form of identical repetition [sic!].

3.1.3. Other cases

In German, subjunctive II, which like a Romance conditional marks unreal propositions, is obligatory after certain modal adverbs such as *beinahe* ‘almost’, as in E17.

- E17. beinahe wäre ich gefallen
 GERM ‘I almost fell’

The adverb is syntactically optional, but if it is omitted, the meaning changes. The unreal subjunctive here is redundant, because the adverb by itself says that the situation was not realised. Other languages, e.g. English and Latin, have the indicative in such sentences. Since the subjunctive here is predictable, we deal with a case of concord pleonasm.

Another kind of construction, known from Latin, involves what has been called the ‘pleonastic reflexive’ for a long time. Since it is analysed at length in Cennamo 1999, an example may suffice here:

- E18. Quid igitur sibi volt pater?
 LATIN what(ACC) then RFL:DAT wants father(NOM.SG)
 ‘What then does my father want?’ (Ter. *Andr.* 375 ap. Cennamo 1999:117)

The reflexive pronoun in E18 is omissible with no change in meaning. It is pleonastic insofar as it underlines the subjective component inherent in the notion of volition. It also combines with other verbs of inactive meaning, focusing on the fact that the process abides in the sphere of the subject. In Vulgar and Late Latin, the construction loses its marked character and evinces some symptoms of grammaticalization.

3.2. Hypercharacterization in inflection

Hypercharacterization in inflection has been a topic in linguistics at least since Paul (1920:162f), where it is treated as ‘pleonasm of formative elements’. This term allows for the possibility that an inflectional category may be hypercharacterized by different morphological processes. Since the phenomenon is well documented (see also Haspelmath 1993, section 5f, and Dressler 2004), we can limit ourselves here to a couple of examples.

In Middle High German, the suppletive comparative of the adjective *guot* ‘good’ was *bass* ‘better’, as in E19 (from ~1200)¹⁸:

E19. von Veldeke der wîse man! der kunde se baz gelobet hân. (Parz. 8, 404, 29f.)

MHG ‘von Veldeke, the wise man! He could have praised her better.’

Secondarily, the form gets the comparative suffix *-er*, which triggers metaphony, so that the modern form *besser* results.

Redundant comparative and superlative marking is common in Indo-European languages. In colloquial English, we find *more easier*, in French and Spanish, we find E20f:

E20. le plus meilleur pays au monde (<http://www.frapru.qc.ca/Comm/Comm044.html>, 29 June 2000)

FR ‘the best country in the world’

E21. KEV ...el mas lindo, el mas mejor!!!! (<http://www.fotoslocas.com/usuarios/k/kevinstone.htm>, 11 February 2004)

SPAN ‘KEV ... the most handsome, the best one!’

The examples from the three languages have it in common that the surplus element is an analytic marker attached to a synthetic form of grading. They differ in that the synthetic comparative has a morpheme of its own in the English example, while E21f evince a suppletive superlative. A pleonastic superlative is, of course, motivated by emphasis. In addition, it may be relevant that the pleonastic comparative and superlative in the Romance languages is restricted

to adjectives with suppletive grading. Insofar, it is safety pleonasm. On the other hand, no emphasis and no safety is discernible in *more easier*; this is just phatic pleonasm.

The examples of E22 are similar both functionally and structurally:

- E22. a. der einzigste/extremste/optimalste
 GERM 'the most only/extreme/optimal'
- b. in keinster Weise 'in no way'

The underlying Latin forms *extremus* 'outermost' and *optimus* 'best' have the position of superlatives in their paradigm and are even marked as such by an—admittedly irregular—superlative allomorph. Naturally, this does not matter for German grammar.¹⁹ Here, the examples in E22(a) are on a par: The focal component—the function of the superlative suffix—is something like 'the relevant domain (identified by the argument of the adjective) is restricted to that subset (or individual) that occupies the positive pole of the scale designated by its host (the adjective stem)'. It is represented by a dedicated inflectional or derivational morpheme (which assigns these cases to hypercharacterization), but otherwise just entailed by the meaning of the latter's host. The application of the superlative suffix to *kein* (E22(b)) works similarly insofar as it pretends *kein* to mean something like 'occupying the positive pole on a scale of scarcity'.

A related phenomenon occurs in German adjectival compounds whose determinans is a superlative form of some adjective and whose determinatum is another adjective or participle, as those in E23(a).

- E23. a. bestmöglich/kürzestmöglich/meist verkauft
 GERM 'best possible/shortest possible/most sold'
- b. das bestmöglichste Ergebnis/der kürzestmöglichste Weg/das meist-verkaufteste Buch
 'the best possible result/the shortest possible way/the best selling book'

Hypercharacterised forms as those in E23(b) are very frequent. In this case, we clearly have concord pleonasm. In addition, pleonasm is here motivated by the principle that inflection should be at the word margin.²⁰

The Old English forms *children* and *brethren* illustrate hypercharacterised nominal plural. Other examples of this kind are Dutch *kinderen* 'children' and German *Jungens* 'boys', all with two different allomorphs of the plural morpheme. Pleonastic plural marking is particularly common in loans. Thus the Italian plural form *spaghetti* ends up as *spaghetts*, with a plural *-s*, in English, Spanish and optionally in German. In contemporary German, the plural *-s* is sufficiently productive to yield such hypercharacterised forms as

Praktikas = *Praktika* ‘practical courses’, *Visas* = *Visa* ‘visas’, *Lexikas* = *Lexika* ‘lexicons’.²¹

All of these examples clearly involve analogical transfer of a marker from a context in which it is the only operator to fulfill the function in question to a context where it pleonastically duplicates an operator already applied. We may generalize that hypercharacterization in morphology is based on **analogy**.²² Moreover, in a diachronic perspective, the two concurrent markers are not on the same level. There is an inner marker which for some reason does not quite do the job, and an outer marker which is currently productive and which speakers feel should appropriately appear on such a word form (cf. Dressler and Dziubalska-Kolaczyk 2001, section 5, Dressler 2004). A more precise formulation of the analogical account might therefore say that hypercharacterization is a kind of **adaptation** of a stem or word form based on paradigmatic pressure (Koefoed and Marle 2004:1581).

A special case of inflectional hypercharacterization may be seen in **word-internal agreement**. E24 provides an example from Lithuanian.

E24.	balt-os-i-os		nakt-ys		mane
LITH	white-NOM.PL.F-DEF-NOM.PL.F		night-NOM.PL.F		I:ACC
	veik-ia	kaippaslapt-ing-i		ker-ai	
	seem-PRS.3	like	mysterious-NOM.PL.M		witchcraft-NOM.PL.M
	‘the white nights seem like mysterious witchcraft to me’ (Stolz 2004:17)				

Synchronically, the Lithuanian definite adjective consists of the adjective stem, inflected for case, number and gender, and a suffixal definiteness morpheme that is again marked for the same categories, often with the same declensional allomorph (Stolz 2004). Such cases arise by grammaticalization, where an erstwhile syntagma consisting of two words showing syntactic agreement is unverbated. At the level of syntax, agreement, although pleonastic, fulfills a function in marking syntactic relations (cf. section 3.1.1). At the level of morphology, it loses any kind of motivation.

3.3. Hypercharacterization in derivation

3.3.1. German action nouns

Consider the derivational relationship between noun and verb. Since we have both deverbal nouns and denominal verbs, this relationship is not per se directional. From the root of the German verb *konzipieren* ‘conceive’, we form the action noun *Konzeption* ‘act of conceiving’, and on the basis of the noun *Analyse* we form the verb *analysieren* ‘make an analysis’. In both cases, an

iconic interpretation of the derivational process would make one believe that the derived stem is semantically more complex than the base; but since the two processes are mirror images of each other, this would lead into a contradiction. We have to conclude that a stem does not, in general, become semantically more complex by mere derivational transferral into a different category. As a matter of fact, we simply get the same concept in two different syntactic categories.

Deverbal nouns in *-ion* (with its allomorphs) such as *condition* and *relation* have been polyfunctional since Latin times. They are primarily action nouns (**nomina actionis**), as *relation* originally signifies the act of referring. Secondly, they are act nouns (**nomina acti**), as *relation* signifies the result of referring something to something else. Moreover, such a verbal noun from a transitive base may develop a **nomen patientis** reading, as in *derivation* (= *derivatum*) ‘derived word’, which shares its non-dynamic character with the *nomen acti*. Once the noun has acquired the secondary meaning, it may seem too weak to serve as an action noun; it may seem to lack in ‘dynamic force’. A clear example is the English noun *position*, which no longer signifies the act of putting, but only its result. The act must now be expressed by *positioning*, which itself is on the way of losing its dynamic character.

The semantic passage of *nomina actionis* into *nomina acti* and *nomina patientis* and the corresponding functional shift in the derivational morpheme forming such deverbal nouns is probably widespread. The German derivational suffix *-ung* is subject to the same process. Thus *Glättung* ‘smoothing’ is a *nomen actionis*, *Bewerbung* ‘(job) application’ is a *nomen actionis* and *acti*, *Spannung* ‘tension, voltage’ is only a *nomen acti*, *Packung* ‘package’ only a *nomen patientis*. There seems to be a drift towards stativisation and reification.

Sometimes the speaker wants to make sure that an action is being designated. His problem is then to signal that whatever nominalising process is applied to the verbal base is not subject to the semantic shift just observed. Abstract nouns that are not overtly derived, like those of the first column of Table 5, are the first to become suspect of stativity. To ensure their dynamic character, they are first verbalized by the suffix *-ier-* (second column), which forms verb stems chiefly from non-German bases. In a second step, these verb stems are nominalised by *-ung*, which, one hopes, conserves the action meaning (third column).

Table 5. Action noun renewal in German

Nominal base	Denominal verb	Action noun	Meaning
Reform	reformieren	Reformierung	reform(ation)
Typologie	typologisieren	Typologisierung	typologisation
Metapher	metaphorisieren	Metaphorisierung	(application of) metaphor
Hypostase	hypostasieren	Hypostasierung	hypostasis

Although the nouns of the first column do conserve a *nomen actionis* reading besides the frequent *nomen acti* use, they are now mostly replaced—chiefly in the former function—by the nouns of the third column.

Nouns derived in *-tion* are not exempt from this remodelling. There are two variants of applying to them the combination of operations observed in Table 5. The first is illustrated by Table 6. Here, the base underlying the derivation in *-tion* is derived, by means of *-ier-*, into a verb. The latter is then nominalised by *-ung*. Thus we find, instead of the age-old action/act nouns in the left column of Table 6, alternate action nouns newly derived in *-ierung*, as in the middle column.²³

Table 6. Alternate action noun derivation in German

Latinate	Germanised	Action reading
Integration	Integrierung	integrating
Qualifikation	Qualifizierung	qualifying (oneself)
Klassifikation	Klassifizierung	classifying
Konversion	Konvertierung	converting
Konzeption	Konzipierung	conceiving/planning
Revision	Revidierung	revising
Tradition	Tradierung	transmitting

The second solution to the expression problem—this one involving hypercharacterisation—is to derive a verb from the act noun itself and nominalise this again. For instance, *Konzeption* ‘conception’, both an action and an act noun, can be verbalised by the suffix *-ier*, yielding *konzeptionieren*, and this can be nominalised again by the suffix *-ung*, yielding *Konzeptionierung*.²⁴ This is visualised in S4, together with the parallel *Revisionierung* ‘revision’.

S4. Recursive nominal and verbal derivation in German

operation	[[X]-tion] _N	[[X]-ier] _V	[[X] _V -ung] _N
product	konzip(ieren) revid(ieren)	Konzeption Revision	konzeptionier(en) reversionier(en) Konzeptionierung Revisionierung

Konzeptionieren is the same as *konzipieren*, and *Konzeptionierung* is the same as *Konzeption* (or *Konzipierung*, for that matter). As the examples show, the processes of action noun derivation and denominal verb derivation may be applied recursively, either one undoing the result of the application of the other. Hypercharacterization here requires the execution of two derivational operations

in tandem, since if I am not content with *Konzeption*, I must first verbalise and then nominalise it again in order to arrive at *Konzeptionierung*.

The motivation behind this trend is not easy to pin down. *-ung* by itself displays the same polysemy as *-tion*, which makes it hard to believe that speakers trust in its dynamicity. One might hypothesize that it is the component *-ier* in *-ierung* which guarantees the action noun reading because *-ier* reflects the underlying verbality. (For speakers' motivations, it would not matter that no theoretically sound argument could be made along these lines, since verbs derived in *-ier* are also at the basis of categorically stative deverbal nouns, e.g. nomina agentis in *-ierer* like *Kopierer* 'copying machine'.) We would then be faced with a new suffix *-ierung*, exclusively dedicated to the formation of nomina actionis. Some documented cases do presuppose that if there is a pair of nouns one of which is derived in *-tion* and the other in *-ierung*, then the second is dynamic. Clear witness of such reasoning is a publication title such as E25:

E25. Konzeptualisierung von Motivation und Motivierung im Kontext situ-
ierten Lernens

GERM 'conceptualisation of motivation and motivating in the context of situated learning'²⁵

Here *Motivation* refers to the pupils' disposition, while *Motivierung* refers to the teachers' action. *Klassifikation* vs. *Klassifizierung* is a stock example of the distinction intended here.²⁶ However, a web search turns up a host of examples like *Deutsche Hotelklassifizierung* 'German hotel classification',²⁷ which clearly refer to the result of the action. Equally in E26f, the nouns derived in *-ierung* are clearly nomina acti.

E26. Der räumlich-zeitliche Bereich ist auf die Positionierung des Referenten
im Raum sowie deren Beibehaltung im Zeitverlauf bezogen. (linguistic
term paper 2004).

GERM 'The spatio-temporal area concerns positioning of the referent in
space and maintenance of the latter [i.e. the position] in the course of
time.'

E27. "Ostasien" older "East Asia"—eine deutsche Konzeptualisierung ([http://
www.lvk-info.org/nr17/lvk-17polap.htm](http://www.lvk-info.org/nr17/lvk-17polap.htm) 1 November 2004)

GERM 'Ostasien' or 'East Asia'—a German conceptualisation

An alternative hypothesis is that—apart from a couple of specific cases like E25—no semantic issue is involved here, and what matters instead is only the replacement or reinforcement of an unproductive derivation mechanism (*-tion*)

by a productive one (*-ierung*). It is true that many formations in *-ierung* are just now replacing older formations in *-tion*. Examples include *Demonstrierung* ‘demonstration’ and *Variierung* ‘variation’, each found several hundred times on the web (1 November 2004), but absent from *Duden Wörterbuch*, 2001 edition. On the other hand, there are also recent well-established neologisms in *-tion* like *Animation*, *Emission*, *Präsentation* (and even more scientific terms such as *Extrapolation*, *Kollokation*, *Partizipation*) which show no tendency to get ousted by counterparts in *-ierung*.

Thus, every attempt to come up with a specific semantic motivation of the redundant application of the derivation in *-ierung* fails. It seems that the general motivation of phatic pleonasm must suffice: the sheer desire to make words sound more impressive. The hypercharacterization resulting from this in cases such as S4 is not specifically intended, but does not bother most speakers either.

The analysis shows that hypercharacterization in derivational morphology must be seen in the context of the renewal of inherited derivata by productive means.²⁸ This renewal itself is not hypercharacterization; but sometimes the renewal does not go back to the roots, but simply works on some available base, which may or may not already be marked for the category in question.

3.3.2. Other cases of derivational hypercharacterization

Derivational processes which come under intensification in the broadest sense, including diminution, augmentation, iteration, etc., are particularly prone to hypercharacterization. Diminution provides some well-known examples. Sometimes different allomorphs of the diminutive marker are stacked, as in Italian *Bertinetto* ‘little Bertie’, *librettino* ‘little booklet’.²⁹ Sometimes the most productive diminutive suffix can be iterated, as in Spanish *chiquitito* ‘tiny little’. There is also a derivational counterpart to the pleonastic superlative in such Italian forms as *ultrabellissimo* ‘most hyperbeautiful’, typical of the language of publicity. The Latin intensive-iterative suffix *-(i)t-* is reapplied in verbs such as *dic-t-it-o* (say-INTS-INTS-1.SG) = *dic-t-o* ‘say repeatedly’, *iac-t-it-o* (throw-INTS-INTS-1.SG) = *iac-t-o* ‘throw repeatedly’.

However, derivational hypercharacterization occurs in other functional contexts as well. The German suffix *-lich* may derive adjectives like *freundlich* ‘kind’, but also adverbs like *schwerlich* ‘hardly’, *gröblich* ‘in a gross way’, *fälschlich* ‘wrongly’. In the latter function, the suffix is barely productive today. Since most adjectives can be used in adverbial function without morphological change, there are many words derived in *-lich* that function both as adjectives and as adverbs, like *wissentlich* ‘knowing(ly)’. Furthermore, there is a more recent and productive adverbialising suffix *-weise*, which has an analogous diachronic origin as Romance *-mente*, viz. its basis are circumlocutions such as *in freundlicher Weise*

‘in a friendly way’, which get univerbated to derived adverbs like *freundlicher-weise* ‘kindly’. Now this suffix is also sometimes added to adverbs derived in *-lich*. Thus we find E28.

E28. ...ein Turmalin, der fälschlicherweise lange für einen Rubin gehalten wurde. (MDR Kultur, 22 February 2004)

GERM ‘...a turmaline that was long regarded as a ruby in a wrongly way.’

Besides such hypercharacterisations, there are also constructions like E29 which go back to the periphrasis, but use an adverb as the attribute to *Weise*³⁰:

E29. Ein Mitglied kann durch den Vorstand ausgeschlossen werden, wenn es in gröblicherWeise gegen die Vereinsinteressen verstoßen hat. (Förderverein Bilzingsleben 31 August 1996)

GERM ‘A member may be excluded by the executive board if he has infringed the association’s interests in a seriously way’.

It seems that the authors of E28 and E29 regarded *fälschlich* as synonymous with *falsch*, and *gröblich* as synonymous with *grob*; i.e. they did not feel that *fälschlich* and *gröblich* are characterised as adverbs. This is then safety pleonasm.

Transitivity of transitive verbs also belongs here. The German applicative prefix *be-* generally transitivity verbs, as in *singen* ‘sing’, *besingen* ‘sing to the honour of’. It applies redundantly in examples like *befüllen* = *füllen* ‘fill’, *bejagen* = *jagen* ‘hunt’,³¹ *befüttern* = *füttern* ‘feed’. Such examples appear to be due to phatic pleonasm.

If, however, a loan is provided with a marker specifying its grammatical class irrespective of the fact that, in the donor language, it already belongs to that class, it is rather a case of safety pleonasm. Thus Spanish *alcanzar* ‘reach’ is a transitive verb, but as a loan in Yucatec Maya, it is provided with the transitivity suffix *-t-*, as shown in E30 and a myriad of similar examples.

E30. k-u alcanzar-t-ik
 YM IMPFV-SBJ.3 achieve-TRR-INCMPL³²
 ‘he achieves it’

3.4. Hypercharacterization in compounding

In the endocentric nominal compound, the determinans forms a more specific concept on the basis of the determinatum. There are at least two pleonastic varieties of this compounding type. The first is illustrated in Table 7.

Table 7. German hyponym compounds

Expression	Composition	Meaning
Sturmwind	storm:wind	storm
Enkelkind	grandchild:child	grandchild
Eichbaum	oak:tree	oak
Grammatikalisierungsprozeß	grammaticalization:process	grammaticalization

Here the determinans is a hyponym of the determinatum; the construction may therefore be called **hyponym compounding**.³³ Although these compounds satisfy the semantic characterization of pleonasm given in section 2.1, they differ in their structure from the phenomena considered so far because it is the modifier alone, not the structural head, that is synonymous to the complex. Unlike all the other cases of pleonasm, it is, thus, the head of the construction that is redundant. It is known that the strategy of hyponym compounding can give rise, through grammaticalization, to a system of nominal classification. In this, hyponym compounding is functionally similar to concord pleonasm, which, as we saw, may evolve into agreement.

A subvariety of hyponym compounding may be seen in the left-hand column of Table 8:

Table 8. Abbreviation elaboration

Elaborated abbreviation	Resolution of abbreviation
ABS system	Anti-lock Braking System
HIV virus	Human Immunodeficiency Virus
LCD display	Liquid-Crystal Display
PIN number	Private Identification Number
ABM-Maßnahme	Arbeitsbeschaffungsmaßnahme 'labour provision measure'

This construction may be called **abbreviation elaboration**. It is very common both in English and in German.³⁴ As in Table 7, the added noun repeats a semantic component already contained in the abbreviation. It also has the same kind of motivation as other pleonasms: In abbreviation elaboration, safety pleonasm concurs with phatic pleonasm. The peculiar feature of abbreviation elaboration is that the focal component is identically present in the base. It is not a matter of synonymy, but of repetition, although the component is not spelt out in the base.

The other main variety of pleonastic compounding is **synonym compounding**, as in E31f:

E31. German *schlußendlich* 'end-finally', *letztendlich* 'last-finally'

E32. *duo-yu*

CHIN extra-remaining

'excessive, extra' (Chao 1968:374f)

While this type does not appear to constitute a productive pattern in German, it has been very important in Mandarin Chinese, apparently as a form of safety pleonasm to disambiguate homonymous bases.

4. THEORETICAL IMPLICATIONS

There is pleonasm at all the grammatical levels from the sentence down to the stem. As in any movement downwards the grammatical levels, paradigmaticity increases. At the highest level, pleonasm starts out as a strategy that lends emphasis to an expression. At the lowest level, it is a strategy that fits an expression in a paradigm or in a structural class. This has become clearest in the two sections (3.2f) devoted to hypercharacterising inflectional and derivational morphemes. Discourse motivations here give way to system-internal pressure.

Theoretically, the pleonastic comparative of English could involve a repetition of the operator *more*, just as the pleonastic action noun in German could involve a repetition of the nominaliser *-tion*. However, this is not what happens. As we have seen, repetition is only a limiting case of redundancy. Hypercharacterization is therefore not just a process of copying. It shares properties with contamination (blending), where the speaker cannot make a choice between two synonymous expressions.³⁵

Hypercharacterization provides some important lessons for linguistic theory. In the fields of syntax, inflection and word-formation, formal descriptive models describe the formation of complex units of a certain category by the combination of an operator with an operand of a certain category. An extreme form of this approach has been known as the 'unitary base hypothesis'. As we have seen (and as has been shown repeatedly in the literature), this purely analytic description runs into problems. There is one category in syntactic and morphological constructions that is of prime importance, and that is the category of the resulting syntagma. The speaker uses an operator to create an expression with certain grammatical properties. Operators are often sensitive to the properties of the operand. But these are of secondary importance, and often the speaker simply does not care.³⁶ This is clearest in the treatment of loans. Here one might want to argue that the grammatical properties of the donor language cannot possibly play a role in the recipient language. However, things are not so simple. Borrowing an item presupposes some degree of knowledge of the donor language, and the item is borrowed precisely for its properties. The most one can say is that the

speaker wants to make sure, with the means of the code he is currently using, that the item has the properties needed in the discourse. The most transparent way of guaranteeing this is the application of a productive operation whose operator confers just the desired property. This speaks in favour of a goal-directed theory of language and of a **holistic approach** to grammatical description, to complement the otherwise needed analytic approach.

Grammaticalization has often been described as a transition from universal iconic discourse strategies to language-specific system-dependent grammatical rules. As far as that goes, grammaticalization involves a loss in motivation. As we have seen, pleonastic phenomena can be arranged on a scale that starts from full motivation in terms of intensification and emphasis and where motivation weakens gradually. As soon as we get to a stage where elements in a sentence start being in concord with each other, we enter the domain of rules of syntax. From here on, the phenomenon goes by the name of hypercharacterization. Further grammaticalization leads to grammatical agreement, first at the syntactic level, finally even inside a word-form.

Some have proposed a principle of **derivational blocking** which says that an otherwise productive derivational process is blocked for a particular base if there is already a derivatum—formed from this base by another process—that occupies the target position. The facts adduced in section 3.3 falsify a simple general version of such a principle. Sometimes the opposite principle seems to be active: concepts that continue to be needed deserve to be expressed by currently productive means, which may lead to a renewal of their expression. This phenomenon is well known from grammaticalization research. It suffices to mention a stock example like the renewal of various verbal categories of Ancient Indo-European languages in their modern descendants (see Lehmann 2002, chapter 2.4, for details). The marking of inflectional categories is thus far from being constrained by a blocking principle. Quite on the contrary, if the system of grammatical meanings includes a certain category, then that category will be marked by such structural means which correspond to the type the language is currently following; and at the same time, their marking by means that belonged to a previous type will fossilise. Research into hypercharacterization may shed new light onto the corresponding issue in derivational morphology. Blocking is counteracted by renewal there, too.

Safety pleonasm evinces a basic insecurity in the control of the code. Since none of us is the master of the norm, we do not have full certainty of the meaning of a word and the service it can do in our speech. Therefore we prefer to play it safe and to combine it with another sign which should also contribute the desired meaning and of which we may feel a little more sure.

Pleonasm and hypercharacterization thus provide evidence of a peculiar kind that language is not a stable system. Older textbooks teach that language changes because we have to adapt it to new needs. Younger textbooks teach that it changes because the language acquisition device comes up with an original

analysis of the input. Pleonasm and hypercharacterization confirm what Coseriu (1958, chapter III) said long ago (cf. also Booij 2005b): Language changes because we create it every day. We have to do so to the extent there is no ready-made language that we could rely upon.

LANGUAGE ABBREVIATIONS

Chin(ese), Fr(ench), Germ(an), Lith(uanian), M(iddle) H(igh) G(erman), Span(ish), Y(ucatec), M(aya)

NOTES

* I thank Giorgio Banti, Dagmar Haumann, Johannes Helmbrecht, Nils Jahn, Yoko Nishina, Su-Rin Ryu and two anonymous reviewers of the *Yearbook of Morphology* for helpful comments on the first draft and for a couple of examples.

¹ The word *overcharacterisation* ‘exaggeration, caricature’ is not a technical term of linguistics.

² In phonemics, a phoneme is hypercharacterised if it differs from the closest less marked members of its subsystem by more than one feature value, as e.g. in Ancient Greek the high round vowel opposed to /o/ was not /u/, but /y/. On the syntagmatic dimension, Sherer (1994) applies the concept of hypercharacterization to syllable structure.

³ Cf. Lausberg (1990, §§ 502, 604).

⁴ Given this, the title of this paper is slightly ill-formed; it should read ‘Pleonasm and, in particular, hypercharacterisation’.

⁵ One of whose manifestations is ellipsis.

⁶ However, Lausberg (1990:328) cites a Roman author who does subsume this construction (in Latin) under pleonasm.

⁷ Most of the English examples in section 2 are from www.wordexplorations.com/pleonasm.html as of 10 February 2004. Some of them may already be found, in their Latin or Greek version, in ancient treatments of rhetoric.

⁸ The modificative nature of pleonastic constructions is, in fact, methodologically ambivalent: It is here treated as an empirical generalization over a phenomenon whose concept does not entail it. However, as Lausberg (1990, §502) shows, already in ancient rhetoric some authors defined pleonasm with respect to modificative constructions.

⁹ A Google search (6 May 2005) for *potentially capable* yields 40,400 pages, 20 of which oppose it to *actually capable* and one to *capable*.

¹⁰ It follows the (German proverbial) maxim *doppelt genäht hält besser* ‘double-stitched lasts longer’.

¹¹ Cf. Malkiel (1957f:79, 98f) on rhythmic aspects of hypercharacterised Spanish *pieses* ‘feet’, Löfstedt (1933) on the idea that a word may be perceived as too short for what it signifies, and Haiman (1985) for theoretical aspects of quantitative isomorphism.

¹² How strong this automatism is may depend on the particular language and a variety of other factors. A text count of combinations of *malen’kij* ‘small’ with diminutive nouns in Russian reported on in Rusakova (2004) finds the following numbers of tokens: (a) no such adjective +

diminutive noun: 200; (b) *malen'kij* + non-diminutive noun: 58; (c) *malen'kij* + diminutive noun: 14. In that corpus, the concord tendency is thus relatively weak. Moreover, in contrast with E8–E10, there are non-pleonastic combinations of ‘small + diminutive’, viz. whenever the diminutive does not mean ‘small’.

¹³ Another example, just to show that professors enjoy no preferential treatment: *daß er eine Beziehung zu einer weiblichen Managerin des Konzerns unterhielt* ‘that he entertained a relationship with a female she-manager of the company’ (*Der Standard* 6 March 2005, p. 3).

¹⁴ With some simplification; cf. Lehmann (1982) for more accuracy.

¹⁵ There may be exceptions to this. Taken literally, a predicate like *more than unique* is self-contradictory. However, a literal interpretation is out of place, because then *more* would have to be the syntactic head of the phrase, while in fact *more than* is a modifier to *unique*.

¹⁶ Ortman (1999) quotes a number of morphological theories—and defends one himself—that exclude hypercharacterization because of its redundancy. Such theories declare the non-existence of facts like those adduced here and in the literature.

¹⁷ It could, in principle, govern any preposition. For instance, in *mit x aufhören* ‘to stop doing x’, the verb particle and the preposition governed are distinct.

¹⁸ I will assume without further discussion that comparison is an inflectional category in the languages at hand.

¹⁹ It may be comforting for German speakers that already the Romans did not shun *extrem-issimus*.

²⁰ See Haspelmath (1993), section 2.4 for discussion of cases of this kind and Dressler (2004) for the sequence of hypercharacterised and hypercharacterising affix.

²¹ Cf. Booij (2005a:259) for similar examples in Dutch.

²² In view of Haspelmath’s (1993, section 5.2) objections against an analogical account, it should be stressed that an analogical model need not be perfect in motivating each and every feature of the transformed item; it suffices that it share some features with the latter.

²³ Some of the nouns in the left-hand column have actually lost their action meaning. Thus: *Unser aller Pflicht ist die Tradierung/² Tradition von Werten*. ‘Everybody among us has as his duty the tradition of values.’

²⁴ *Konzeptionierung* is absolutely fashionable; a Google search (2 November 2004) turns up 29,200 examples. One can also hear *Konzeptionalisierung* (Google: 2,280 examples). Many of the examples of both nouns exhibit a stative sense.

²⁵ Stark, Robin and Mandl, Heinz (1998), *Konzeptualisierung von Motivation und Motivierung im Kontext situierter Lernens* (Forschungsbericht Nr. 091). München: Ludwig-Maximilians-Universität, Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie, Internet.

²⁶ It is adduced, e.g., in Fleischer (1971:156f). In other cases, an investigation into the differential function of rival forms of derivation does yield positive results; cf., e.g. Kaunisto (1999) on the English suffixes *-ic* and *-ical*.

²⁷ It is adduced, e.g., in Fleischer 1971:156f and Knobloch 2002:336. The main title of the website <http://www.hotelsterne.de/>.

²⁸ The derivational suffix German *-ität* = English *-ity* forms abstract nouns on the basis of adjectives chiefly of Greco-Latin origin, as in *Publizität* ‘publicity’. This is currently one of the most productive means towards this end. Other suffixes like *-ie* ‘-y’, as in *Monotonie*

'monotony', are losing ground. For some time, old derivata in *-ie/-y* have been replaced by more modern (and longer!) ones in *-ity*. Thus, *Anonymie/anonymy* have been all but ousted by *Anonymität/anonymity*; *Synonymie/synonymy* and *Homonymie/homonymy* still go strong, but *Synonymität/synonymity* and *Homonymität/homonymity* are on the advance.

²⁹ German allows this to a much more limited extent, as in *Kinderleinchen* 'children-DIM-DIM' or *Schatzleinchen* 'darling-DIM-DIM-DIM'.

³⁰ The typo in E29 is telling: the text is evidently an emendation of an earlier version that contained *gröblicherweise* 'seriouslywise'.

³¹ *Bejagen* also means 'hunt in (a hunting-ground)' and then is a regular, non-pleonastic applicative derivation.

³² IMPFV imperfective, SBJ subject, TRR transitiviser, INCMPL incomplete.

³³ Fleischer (1971:93f) speaks of 'clarifying compounds'. See Bloomer (1996) for a detailed study.

³⁴ According to anecdotal evidence provided by G. Banti (p.c.), abbreviation elaboration occurs in Italian, too, although there it has a different structure, e.g. *virus HIV*.

³⁵ This point is stressed in Dressler et al. (2001). Cf. Haspelmath (1993, section 6.2) for some discussion.

³⁶ Plag (2005) argues emphatically that categorical properties of the bases of word-formation processes are in general irrelevant. Earlier proponents of a holistic, semantic rather than structural approach to word-formation include Plank (1981).

REFERENCES

- Bloomer, Robert K. (1996). Die pleonastischen Zusammensetzungen der deutschen Gegenwartssprache. *American Journal of Germanic Linguistics and Literatures* 8, 69–90.
- Booij, Geert (2005a). *The Grammar of Words. An Introduction to Linguistic Morphology*. Oxford: Oxford University Press.
- Booij, Geert (2005b). Paradigmatic morphology. In Bernard Tranel and Georgette Dal (eds.), *Hommage a Danielle Corbin*. Amsterdam: Benjamins.
- Cennamo, Michela (1999). Late Latin pleonastic reflexives and the unaccusative hypothesis. *Transactions of the Philological Society* 97(1), 103–150.
- Chao, Yuen Ren (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Coseriu, Eugenio (1958). *Sincronía, diacronía e historia. El problema del cambio lingüístico*. Montevideo: Universidad de la República de Uruguay, Facultad de Humanidades y Ciencias (Repr.: Madrid: Gredos, 1973.)
- Dressler, Wolfgang U. (2004). Hypercharacterisation and productivity in inflectional morphology. In Thomas Krisch, Thomas Lindner and Ulrich Müller (eds.), *Analecta homini universali dicata. Arbeiten zur Indogermanistik, Linguistik, Philologie, Politik, Musik und Dichtung. Festschrift für Oswald Panagl zum 65. Geburtstag*. 2 Bde. Stuttgart: H.-D. Heinz (Stuttgarter Arbeiten zur Germanistik, 421); 2, 515–524.
- Dressler, Wolfgang U., Katarzyna Dziubalska-Kolaczyk, and Rossella Spina (2001). Sources of markedness in language structures. *Folia Linguistica Historica* 22, 103–135.

- Fleischer, Wolfgang (1971). *Wortbildung der deutschen Gegenwartssprache*. Tübingen: M. Niemeyer (2. unveränd. Aufl.; 1.: Leipzig: VEB Verlag Enzyklopädie, 1969).
- Haiman, John (1985). *Natural Syntax. Iconicity and Erosion*. Cambridge: Cambridge University Press.
- Haspelmath, Martin (1993). The diachronic externalization of inflection. *Linguistics* 31, 279–309.
- Jakobson, Roman (1960). Closing statement: Linguistics and poetics. In Thomas A. Sebeok (ed.), *Style in Language*. Cambridge, MA: MIT Press; New York: Wiley, 350–377.
- Kaunisto, Mark (1999). *Electric/electrical and classic/classical*: Variation between the suffixes *-ic* and *-ical*. *English Studies* 80, 343–370.
- Knobloch, Clemens (2002). Zwischen Satz-Nominalisierung und Nennerivation: *-unq*-Nomina in Deutschen. *Sprachwissenschaft* 27, 333–362.
- Koefoed, Geert and Jaap van Marle (2004). Fundamental concepts [of morphological change]. In Geert Booij, Christian Lehmann, Joachim Mugdan and Stavros Skopeteas (eds.), *Morphologie. Ein internationales Handbuch zur Flexion und Wortbildung. 2. Halbband*. Berlin: W. de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft, 17.2); 1574–1589.
- Lausberg, Heinrich (1990). *Handbuch der literarischen Rhetorik. Eine Grundlegung der Literaturwissenschaft*. Stuttgart: F. Steiner (3rd edn. = repr. of the 2nd., enl. edn. 1973).
- Lehmann, Christian (1982). Universal and typological aspects of agreement. In Hansjakob Seiler and Franz Josef Stachowiak (eds.), *Apprehension. Das sprachliche Erfassen von Gegenständen. Teil II: Die Techniken und ihr Zusammenhang in den Einzelsprachen*. Tübingen: G. Narr (LUS, 1, II); 201–267.
- Lehmann, Christian (2002). *Thoughts on Grammaticalization*. Erfurt: Seminar für Sprachwissenschaft der Universität (ASSidUE, 9) (2. revised edition).
- Löfstedt, Einar (1933). Wortform, Wortumfang und Verwandtes. In Einar Löfstedt, *Syntactica. Studien und Beiträge zur historischen Syntax des Lateins. Zweiter Teil: Syntaktisch-stilistische Gesichtspunkte und Probleme*. Lund etc.: Gleerup (Acta Reg. Societatis Humaniorum Litterarum Lundensis, X:2); 35–62.
- Malkiel, Yakov (1957f). Diachronic hypercharacterization in Romance. *Archivum linguisticum* 9, 79–113; 10, 1–36.
- Marantz, Alec and Caroline Wiltshire (2000). Reduplication. In Geert Booij, Christian Lehmann and Joachim Mugdan (eds.), *Morphologie. Ein internationales Handbuch zur Flexion und Wortbildung. 1. Halbband*. Berlin: W. de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft, 17.1); 557–567.
- Ortmann, Albert (1999). Affix repetition and non-redundancy in inflectional morphology. *Zeitschrift für Sprachwissenschaft* 18, 76–120.
- Paul, Hermann (1920). *Prinzipien der Sprachgeschichte*. Tübingen: M. Niemeyer (5th edn.).
- Plag, Ingo (2005). Syntactic category information and the semantics of derivational morphological rules. *Folia Linguistica* 38, 195–225.
- Plank, Frans (1981). *Morphologische (Ir-)Regularitäten. Aspekte der Wortstrukturtheorie*. Tübingen: G. Narr (Studien zur deutschen Grammatik, 13).
- Rusakova, Marina V. (2004). Russian diminutive/hypocoristic adjectives: Between inflection and derivation. *11th International Morphology Meeting*, Vienna, 14–17 February 2004.

Sherer, Tim D. (1994). *Prosodic phonotactics*, PhD dissertation. Amherst: University of Massachusetts.

Stolz, Thomas (2004). Pleonastic morphology dies hard. Evidence from Lithuanian and sundry languages. *11th International Morphology Meeting*, Vienna, 14–17 February 2004.

*Seminar für Sprachwissenschaft
Universität Erfurt
Deutschland*

christian.lehmann@uni-erfurt.de

Pulaar verbal extensions and phonologically driven affix order

MARY PASTER*

1. INTRODUCTION

Pulaar, a West Atlantic language spoken across a wide area of West Africa, has a number of verb suffixes that can occur in combination, offering the linguist an opportunity to examine their relative ordering and the principles governing it. Arnott (1970:333,366) reported that in the Gombe Fula dialect, the order of affixes is largely fixed. In particular, according to Arnott, the first four suffixes to come after the verbal stem are consonantal suffixes ordered according to the formula 'TDNR': all of the /-t/ suffixes precede the /-d/ suffixes, which precede the /-n/ suffix, which in turn precedes the /-r/ suffixes (1970:366). As discussed in this paper, many of the verb suffixes, including several of the 'TDNR' suffixes that are the focus of this paper, enter into semantic scope relations with each other. Therefore, if it is true that their order is fixed, then the behaviour of these suffixes contradicts the claim of Rice (2000) that affixes are ordered according to their relative semantic scope and that templatic (fixed) affix order results only when the affixes in question do not have a scope relationship. In this paper, I present new data from a speaker of a related dialect of Pulaar showing that scope relations do play a crucial role in the ordering of these suffixes, and I then show that such an explanation is also consistent with Arnott's (1970) data and in fact accounts for a larger set of Arnott's examples than did his own claim of fixed ordering. I also discuss implications of this reanalysis of Pulaar affix order for Rice's (2000) claim as well as for the morphological model advanced by McCarthy and Prince (1993a,b).

The structure of the paper is as follows. First, in the remainder of section 1, I discuss Rice's (2000) Scope Hypothesis and other proposals relating the order of affixes to their scope (Baker 1985, Bybee 1985, Condoravdi and Kiparsky 1998), and then provide background on the Pulaar language. In section 2, I present Arnott's (1970) affix order data from Gombe Fula and discuss Arnott's claim that the order of affixes is fixed. In section 3, I present new data from a speaker of Fuuta Tooro Pulaar and an analysis of these data in terms of scope. I then present in section 4 a reanalysis of Arnott's (1970) Gombe Fula data similar to the one proposed for the Fuuta Tooro dialect discussed in the preceding section. In section 5, I discuss some theoretical implications of this new analysis of Pulaar affix order. section 6 concludes and summarizes the paper.

1.1. Scope-based affix order and emergent templates

Several researchers have claimed that affix order is related to syntactic/semantic principles (see, for example, Baker 1985, Bybee 1985, Condoravdi and Kiparsky 1998, Rice 2000). Specifically, Rice (2000) claims that the relative order of affixes corresponds to their scope (defined in terms of semantic compositionality), such that if morpheme X has scope over morpheme Y, then morpheme X occurs further from the root than morpheme Y. I will refer to this as the Scope Hypothesis, and as we will see, this principle is useful in analysing affix order in Pulaar.

One alternative approach to affix order is the use of fixed or ‘templatic’ order. Analyses in these terms often propose ‘slots’ on either side of the root in which specific affixes invariably occur. Templatic analyses are proposed when affix order appears to be arbitrary and does not relate straightforwardly to any external principle (see Bloomfield 1962, Zwicky 1985, Anderson 1986, Simpson and Withgott 1986, Speas 1990, Stump 1992, Inkelas 1993, Hyman and Inkelas 1999, and Good 2003 for examples of template-based analyses of affix order).

Rice (2000) characterizes templates as ‘emergent’ rather than as playing a fundamental role in morphological systems. Specifically, Rice argues (2000:396) that templatic order can arise only when there are no scope relations among the affixes in question. As I discuss in section 5, this claim may need to be weakened in light of the facts of Pulaar affix order, which reveal that templatic orderings can and do arise even when the affixes involved have clear scope relations. In general, though, the findings reported in this paper will be demonstrated to be consistent with Rice’s (2000) Scope Hypothesis.

1.2. The Pulaar language

The Pulaar language is spoken in a wide area of West Africa and comprises a number of dialects that are mutually intelligible to varying degrees. The name ‘Fula’ is sometimes used as a cover term for all of the Pulaar dialects plus other languages known by names such as Fulfulde, Fulani, and Fulbe. However, the Fula name usually does not include Pulaar, so there is no single good cover term for all of these languages, even though they are very closely related and seem to form a continuum of mutual intelligibility. Since the primary focus of this paper is on a dialect of Pulaar, I use ‘Pulaar’ to refer to the entire language group including Pulaar proper as well as the Fula languages.

There is a large literature on the Pulaar languages. Of particular relevance to this paper are two works that discuss verbal extensions in different Pulaar dialects. The first is Arnott’s (1970) description of the Gombe Fula dialect spoken in northern Nigeria. Arnott (1970) provides the first in-depth description of the verbal extensions and explicitly discusses their relative order. In a series of articles, de Wolf (1985, 1986, 1987, 1991) discusses the verbal extensions as they

are used in Noŋaare Fulani, a dialect spoken in western Niger. de Wolf gives examples of extensions in combination, showing their relative order, though he does not provide an analysis of the order.

The Fuuta Tooro dialect, which I describe in this paper in section 3, is spoken in the Fuuta Tooro region along the border between Senegal and Mauritania. The consultant for the present study is a 42 year-old speaker who moved to the US from a town near Matam, which is in Senegal in the eastern part of the Fuuta Tooro region.

2. GOMBE FULA

In this section I present Gombe Fula data from Arnott (1970) showing the meaning and usage of verb suffixes, focusing on the consonantal suffixes which are the focus of this paper. I give examples from Arnott (1970) in which two or more of these suffixes are combined, then discuss Arnott's interpretation of the relative ordering of the suffixes.

2.1. *The verb suffixes of Gombe Fula*

Arnott (1970:334) lists nineteen verb 'extensions' in Gombe Fula (1) (examples pp. 340–364)¹. In each example, the relevant extension appears in bold text.

(1) Shape	Label	Example
-d	Denominative (DEN)	fur- d -a 'be grey'
-t	Reversive (REV)	taar- t -a 'untie'
-t	Repetitive (REP)	soor- t -o 'sell again'
-t	Reflexive (REF)	ndaar- t -o 'look at oneself'
-t	Retaliative (RET)	jal- t -o 'laugh at... in turn'
-t	Intensive (INT)	yan- t -a 'fall heavily'
-d	Associative (ASS)	nast- id -a 'enter together'
-d	Comprehensive (COM)	janng- id -a 'read, learn all...'
-n	Causative (CAU)	woy- n -a 'cause to cry'
-r	Modal (MOD)	ɓe mah- ir -i dɪ 'they built them with'
-r	Locative (LOC)	'o 'yiw- r -ii 'he came from'
-an	Dative	'o wolw- an -ii 'he spoke to'
-indir	Reciprocal	ɓe koomn- indir -ii 'they greeted e.o.'
-ootir	Reciprocal	ɓe tokk- ootir -i 'they followed e.o.'
-kin	Simulative	'o wum- kin -o 'pretend to be blind'
-law	Celerative	'o maɓɓ- ilaw -ii 'he shut... quickly'
-oy	Distantive	yahu wi'- oy ɓe 'go and tell them'
RED-n	Iterative	'o wari- war-in -ii 'he kept coming'
RED-tir	Iterative-Reciprocal	ɓedon pii- pii-tir -a 'they keep hitting e.o.'

Arnott lists both -C and -VC forms of the single consonant extensions (where V is usually [i]), but I give only the -C form here. The -VC form of each consonantal suffix occurs in a phonological environment that is consistent but difficult to capture using phonological features: it occurs after consonant clusters and geminates, and after non-implosive obstruents, 'y, *sh*, and *h* (Arnott 1970:335). The -VC form also occurs variably after some single consonants not included in the set given above; though Arnott does not mention morphological conditioning, it appears that the -VC form occurs when preceded by another extension.

According to Arnott (1970), the suffixes generally occur in the order in which they are presented above. Arnott makes a more specific claim (1970:366) that the order of the consonantal suffixes (the first eleven suffixes listed above) is fixed: 'As far as [the 'TDNR'] extensions are concerned (the purely verbal extensions consisting basically of a single consonant), [the] normal order can be summarized by the formula T-D-N-R'². There are a few minor discrepancies between the order that Arnott lists on page 334 and the order that is revealed in the examples he provides throughout the grammar, so I have adjusted the ordering of items in the list to reflect the ordering found in the examples (none of the changes affects the consonantal suffixes that are of particular interest in this paper). Each of the consonantal suffixes is discussed in more detail below.

2.1.1. *The Denominative -d suffix*

The Denominative -*d* suffix generally attaches not to verb roots but to adjectival roots. The result is a verb stem to which any number of verbal suffixes may attach. The -*d* suffix invariably occurs immediately after the root, which is unsurprising since the other suffixes to be discussed below attach only to verb stems; the -*d* suffix must therefore attach first to an adjectival root, 'converting' it into a verb stem suitable to host verbal extensions. Some examples of the -*d* suffix are provided below (Arnott 1970:363).

(2)	fur-	'grey'	fur-d-a	'be grey'
	yam-	'healthy'	yam-d-a	'be healthy'
	'ool-	'yellow'	'ool-d-a	'be yellow'
	barka	'blessing'	bark-id-a	'be blessed'
	semmb	'strength'	semmb-id-a	'be strong'
	meere	'in vain'	meer-id-a	'be worthless'

de Wolf (1987) also discusses this suffix as used in Noŋaare Fulani, referring to it as the Verbaliser since it can attach to adjectives, nouns, and adverbs, converting each to a verb.

2.1.2. The *-t* suffixes

Arnott lists five different verbal extensions whose basic shape is */-t/*. One issue that arises when sets of homophonous extensions are considered is how to determine what constitutes a separate morpheme. Each of the *-t* suffixes has at least a slightly different meaning, though some of their meanings overlap. Arnott's basis for distinguishing these suffixes is not only semantic, but also phonological: three of the *-t* suffixes (Reversive, Repetitive, and Intensive) are reported to have *-ut* allomorphs in addition to *-t* and *-it*, while the Reflexive and Retaliative have only *-t* and *-it*. The same extensions are discussed by de Wolf (1985) for Nongaare Fulani, where these extensions are also homophonous except for a *-c* allomorph in the Reflexive and Repetitive. Based on the phonological evidence in combination with some generalizations about the type of verb that each suffix will attach to, de Wolf follows Arnott in assuming that there are five distinct *-t* suffixes. In sections 2.1.2.1–2.1.2.5, I describe each of the five *-t* suffixes distinguished by Arnott (1970) for Gombe Fula. For each suffix, examples are provided showing bare verbs and the corresponding extended verbs. Arnott gives these examples in citation form without sentential context or morpheme-by-morpheme glosses; their purpose is simply to illustrate the basic meaning change that applies to the stem when each suffix is attached.

2.1.2.1. The Reversive *-t* suffix

According to Arnott, the Reversive suffix causes the extended stem to have a meaning 'opposite' that of the root, as shown in the examples below (Arnott 1970:340).

(3)	fi6a	'tie'	fi6-t-a	'untie'
	taara	'wind'	taar-t-a	'unwind'
	6ila	'hang up'	6il-t-a	'take down'
	soma	'become tired'	som-t-a	'lose one's tiredness'
	sad'a	'be difficult'	sad-t-a	'be easier'
	ja66o	'welcome'	ja66-it-o	'take leave of'

The Reversive allomorphs are *-t*, *-it*, and *-ut*.

2.1.2.2. The Repetitive *-t* suffix

This suffix denotes repetition of an action. Examples are shown below (Arnott 1970: 341).

(4)	'yama	'ask'	'yam-t-o	'ask again'
	rema	'hoe'	rem-t-o	'do a second hoeing'
	soora	'sell'	soor-t-a	'sell again'

wi'a	'say'	wii-t-o	'say again, repeat'
fiya	'hit'	fii-t-o	'hit again'
fud'fa	'begin'	fud'f-it-a	'start again'

The Repetitive allomorphs are *-t*, *-it*, and *-ut*. Note that these are the exact same allomorphs as listed for the Reversive above. Furthermore, there is semantic overlap between the two suffixes, so that there are several instances where the identity of the suffix is ambiguous even when the translation of the utterance is known. For example, Arnott cites the following as an example of a Reversive: *feew-t-a* 'cool down (after being hot)', formed from the verb *feewa* 'be cold'. This could also be interpreted as a Repetitive form with the literal meaning 'be cold again,' which would imply that one was hot in between the two instances of being cold.

2.1.2.3. *The Reflexive -t suffix*

The Reflexive *-t* is an argument structure-changing suffix that reduces the total number of arguments of the verb by one, such that the subject performs the action on him/herself or for his/her own benefit (Arnott 1970:342). Examples are shown below (p. 342).

(5)	ndaara	'look at'	ndaar-t-o	'look at oneself'
	wara	'kill'	war-t-o	'commit suicide'
	ta'ya	'cut'	ta'y-it-o	'cut oneself'
	nana	'hear'	nan-it-o	'hear oneself'
	jala	'laugh'	jal-it-o	'laugh at oneself'
	yima	'sing'	yim-t-o	'sing to oneself'

The Reflexive has the allomorphs *-t* and *-it*.

2.1.2.4. *The Retaliative -t suffix*

When added to a verb stem, this suffix indicates that an action is done to someone else in retaliation, as seen in the examples below (Arnott 1970:342–343).

(6)	ndaara	'look at'	ndaar-t-o	'look at ... in turn'
	jala	'laugh at'	jal-t-o	'laugh at ... in turn'
	food'a	'pull'	food-t-o	'pull ... in turn'
	lata	'kick'	lat-it-o	'kick back'
	hud'a	'abuse'	hud-t-o	'abuse in turn'
	fiya	'hit'	fii-t-o	'hit ... back'

The Retaliative has the allomorphs *-t* and *-it*.

A comparison of these examples with those found in the preceding sections indicates that in some cases, the same verb root may have more than one of the *-t* suffixes affixed to it, in some cases resulting in a single phonetic form with multiple possible meanings. For instance, the form *ndaar-t-o* may have the Reflexive meaning ‘look at oneself’ or the Retaliative meaning ‘look at ... in turn’. Similarly, *fii-t-o* may have the Repetitive meaning ‘hit again’ or the Retaliative meaning ‘hit ... back’.

2.1.2.5. *The Intensive -t suffix*

This suffix indicates ‘completeness, severity, intensity, etc.’ as shown in the examples below (Arnott 1970:343).

(7)	fooda	‘pull’	food-t-a	‘pull tight’
	sada	‘be hard, difficult’	sad-t-a	‘be very hard, difficult’
	yana	‘fall’	yan-t-a	‘fall heavily’
	majja	‘get lost’	majj-it-a	‘get completely lost’
	daro	‘stand’	dar-t-o	‘stand firm’
	’yama	‘ask’	’yam-t-a	‘interrogate’

The Intensive has the allomorphs *-t*, *-it*, and *-ut*.

Again, a comparison with previous examples shows that at least some roots are compatible with other *-t* suffixes in addition to the Intensive: *fooda* ‘pull’ also takes the Retaliative suffix, while *’yama* ‘ask’ takes the Repetitive.

2.1.3. *The -d suffixes*

Arnott lists another set of homophonous suffixes, the Associative and Comprehensive *-d* suffixes. Since both suffixes have the same set of allomorphs (*-d*, *-id*, *-ud*, *-od*), they cannot be distinguished phonologically. Both suffixes can attach to any verb in Gomba Fula. Arnott distinguishes the Comprehensive from the Associative based on their syntactic context: the Comprehensive occurs with prepositional phrases introduced by *’e* ‘with’, while the Associative occurs with *fuu* ‘all’ added to the subject or object (1970:346). Arnott acknowledges that this distinction allows for a significant amount of ambiguity between the two different suffixes, since neither is required to occur in the syntactic environment that distinguishes it. In his description of the same suffixes in Noŋaare Fulani, de Wolf (1991) distinguishes the Associative and Comprehensive (which are also homophonous in that dialect) based on the restriction that the Associative cannot attach to active verbs. The weak evidence for a distinction between these suffixes, in combination with their semantic similarity, suggests that these may really be a single suffix, which I will claim is the case in Fuuta Tooro. Leaving

this issue aside for the moment, below I present descriptions and examples for the Associative and Comprehensive suffixes in Gombe Fula.

2.1.3.1. *The Associative -d suffix*

According to Arnott, the Associative suffix denotes either ‘joint action’ or ‘action in association with some person or thing’ (1970:344). The effect on the arguments of the verb is to require either a plural subject or else any subject plus a second actor introduced by a preposition. Examples are shown below (Arnott 1970:345).

(8)	wara	‘come’	war-d-a	‘come in company’
	joodo	‘sit down’	jood-d-o	‘sit, settle together’
	yaha	‘go’	yaa-d-a	‘go together’
	wada	‘do’	waa-d-a	‘do together’
	wolwa	‘speak’	wol-d-a	‘speak with’
	nasta	‘enter’	nasd-id-a	‘enter together’

2.1.3.2. *The Comprehensive -d suffix*

The Comprehensive suffix indicates ‘totality or completeness’ of the subject or object (Arnott 1970:345). Examples are shown below (Arnott 1970:346).

(9)	nyaama	‘eat’	nyaam-d-a	‘eat up completely’
	habba	‘tie’	habb-id-a	‘tie up all...’
	winnda	‘write’	winnd-id-a	‘write all...’
	yara	‘drink’	yar-d-a	‘drink up (completely)’
	mabba	‘close’	ma66-id-a	‘[close] all...’

2.1.4. *The Causative -n suffix*

The causative suffix adds an object to the verb and contributes the meaning ‘cause to,’ ‘arrange for,’ or ‘make’ (Arnott 1970:346–347), as shown in the examples below (p. 347). de Wolf (1986) discusses the use of the same extension in Nongaare Fulani, where its behaviour and shape seem to be identical to those of the Gombe Fula Causative.

(10)	hula	‘fear’	hul-n-a	‘frighten’
	jala	‘laugh’	jal-n-a	‘amuse’
	woya	‘cry’	woy-n-a	‘cause to cry’
	nyaama	‘eat’	nyaam-n-a	‘give to eat, feed’
	hoya	‘be easy’	hoy-n-a	‘make easy’
	wooja	‘be red’	wooj-in-a	‘redden’
	lugga	‘be deep’	lugg-in-a	‘deepen’

The Causative has two allomorphs, *-n* and *-in*, distributed according to the principle discussed above for the distribution of *-C* vs. *-VC* allomorphs for all of the consonantal extensions.

2.1.5. The *-r* suffixes

Arnott distinguishes two *-r* suffixes, Modal and Locative. The Modal introduces a noun that is either an instrument or a manner in which an action is done. The Locative introduces a noun that is a location in or near which an action is done. Both suffixes have the same allomorphs (*-r*, *-d*, *-ir*, and *-or*), and both can attach to any verb. Arnott's basis for distinguishing the two is their meaning difference, but it is not clear why the modal and instrumental meanings were deemed similar enough to be represented by a single Modal suffix while the Locative was distinguished. It seems plausible to assume that there is only one *-r* suffix in Gombe Fula, which functions as a modal, instrumental, and locative marker. de Wolf (1991) follows Arnott in proposing two separate *-r* extensions for Noŋaare Fulani, but in that dialect, the suffixes can be distinguished phonologically and by the more restricted distribution of the Locative suffix. Below I provide descriptions and examples of the Modal and Locative in Gombe Fula.

2.1.5.1. The Modal *-r* suffix

The Modal suffix indicates either the manner in which an action is done, or else an instrument with which an action is done. In each case, the addition of this suffix changes the argument structure of the verb such that it supports an additional object (Arnott 1970:348). Examples of the Modal suffix are shown below (Arnott 1970:348–349).

- (11) 'o ha66-ir-ii gujjo boggol
 3sg tie-MOD-past thief rope
 'he tied up the thief with rope'
- 'o ma66-ir-ii yolnde semmbe
 3sg close-MOD-past door force
 'he shut the door with force'
- baŋŋaaro ta'y-ir-i kusel lafi
 butcher cut-MOD-past meat knife
 'the butcher cut the meat with a knife'

2.1.5.2. The Locative *-r* suffix

The Locative suffix indicates a location in or near which an action takes place. Examples are shown below (Arnott 1970:352).

- (12) 'yiw-r-ii fuuna
 come-LOC-past east
 'came from the east'
- to ɓe njood-or-ii
 where 3pl sit.down-LOC-past
 'where did they settle?'
- 'o hoot-ir-ii ta ɗatal ngala
 3sg return-LOC-past by road that
 'he returned by that road'

Now that the individual consonantal extensions have been introduced, we will proceed in the following section to examine their relative ordering when two or more of them are used in a single verb.

2.2. Order of the consonantal suffixes

Arnott gives examples showing most of the possible combinations of 'TDNR' suffixes. An exhaustive list of Arnott's examples with two or more (non-homophonous) 'TDNR' suffixes is given in (13)³.

- (13) T-N-R
 'o yam-ɗ-it-in-ir-ii mo lekki gokki kesi
 3sg_i healthy-DEN-[REP]-CAU-MOD-past 3sg_j medicine other new
 'He_i cured him_j with some new medicine' (p. 368)
- T-D-R
 'o jaɓ-t-id-ir-an-ii yam depte 'e semmbe
 3sg take-INT-COM-MOD-dative-past 1sg books with force
 'He snatched all my books from me by brute force' (p. 367)
- T-D
 'o maɓɓ-it-id-ii jold'e fuu
 3sg close-REV-COM-past doors all
 'He opened all the doors' (p. 367)
- T-R
 'o maɓɓ-it-ir-ii yolnde hakkiilo
 3sg close-REV-MOD-past door slowly
 'He opened the door slowly' (p. 367)
- T-R
 war-t-ir-
 come-REV-MOD-
 'bring back' (p. 367)

D-R
 no njood-od-or-too maβbe 'e mi
 how sit/live-ASS-MOD-relative.future 3pl with 1sg
 'How shall I sit/live with them?' (p. 367)

D-R
 to njood-od-or-too maβbe 'e mi
 where sit/live-ASS-LOC-relative.future 3pl with 1sg
 'Where should I sit with them?' (p. 367)

Arnott also cites some forms where it appears that the ordering of affixes violates his 'TDNR' generalization. All of these 'exceptional' forms are given in (14).

(14) D-T
 mi wol-d-it-at-aa 'e maβbe
 1sg speak-COM-REP-future-negative with 3pl
 'I won't speak with them again' (p. 368)

N-D
 'o nyaam-n-id-ii di
 3sg eat-CAU-COM-past 3pl
 'He fed them all' (p. 368)

R-D
 mi yaa-r-id-ii di
 1sg take-MOD-COM-past 3pl
 'I took them all' (p. 368)

T-R-D
 mi war-t-ir-id-an-te di
 1sg come-REV-MOD-COM-DAT-future 3pl
 'I'll bring them all back to you' (p. 368)

N-T
 mi hul-n-it-oo mo
 1sg fear-CAU-RET-subjunctive 3sg
 '(If he frightens me,) I'll frighten him in turn' (p. 368)

Arnott explains away the exceptional orderings as cases of lexicalised stem-extension combinations: 'Variation from the usual order seems to be confined to cases where the basic radical and first extension . . . frequently occur together as an extended radical . . .' (p. 367). One diagnostic for identifying lexicalised forms is to determine whether the meaning of the form is compositional or idiomatic. Lexicalised forms are more likely to have idiomatic meanings (where the meaning of the root-affix combination is not predictable from the meaning of the root and affix taken separately), yet all of the forms in (14) have compositional meanings; that is, in each case, the meaning of the putative lexicalised

root-extension combination is straightforwardly predictable from the meaning of the root and the extension. For example, Arnott assumes that in the example *mi wol-d-it-at-aa 'e mabbe* 'I won't speak with them again' cited above, which contains a Comprehensive and a Repetitive (both shown in bold), the first extension is actually part of a separately listed verb stem, *wold* 'speak with', so that the actual morphological structure of the verb is *wold-it-at-aa*. This assumption saves the 'TDNR' generalization, since otherwise the *-d-t* order in this form would constitute a counterexample. However, since the meaning of the putative stem *wold* 'speak with' follows straightforwardly from the semantics of the root meaning 'speak' combined with the Comprehensive, which contributes a 'joint action' meaning, the only evidence for this root-suffix combination being lexicalised is Arnott's observation that it is a frequently occurring combination. This evidence is thus somewhat weak and warrants further examination. As will be discussed in section 4, once the 'TDNR' generalization is abandoned in favour of a scope-based analysis, the non-TDNR orderings can be explained without the assumption that they involve lexicalised root-extension combinations.

The Gombe Fula data that have been presented above are all consistent with Rice's (2000) Scope Hypothesis. In no example do we find that adherence to a fixed 'TDNR' ordering schema causes the order of affixes not to correspond to their scope. Furthermore, as was shown, there are some exceptions to the 'TDNR' order, though Arnott claims that all of these exceptions involve lexicalised stems. Thus, the evidence presented by Arnott (1970) is consistent with both a scope-based and a templatic analysis. More data are needed to determine what happens when the Scope Hypothesis and the template make conflicting predictions for the relative order of specific combinations of suffixes. In the following section, I present new data from a Senegalese dialect of Pulaar, and I show how these data support a scope-based analysis of Pulaar affix order.

3. FUUTA TOORO PULAAR

A new study of suffix order in Pulaar was carried out in consultation with a speaker from north-eastern Senegal. His dialect is known as Fuuta Tooro, since the region where it is spoken (northern Senegal and southern Mauritania) was formerly the Fuuta Tooro state (*Ethnologue* 14). Although this area is relatively far from the area where Gombe Fula is spoken, the two dialects are likely to be mutually intelligible, as evidenced by the fact that the Fuuta Tooro speaker was able to understand all of the Gombe Fula examples from Arnott (1970) that were presented to him. In this section, I present the consonantal affixes found in Fuuta Tooro and examine in pairwise fashion each of the possible combinations of the consonantal affixes to determine their relative ordering. I then present an analysis of the order of these suffixes based on semantic scope and a partial templatic ordering.

3.1. The consonantal extensions of Fuuta Tooro Pulaar

The consonantal suffixes of Fuuta Tooro are given in (15).

(15)	Shape	Label	Example
	-d'	Denominative (DEN)	mi dom-d'-ii ⁴ 'I became thirsty'
	-t	Separative (SEP)	mi udd-it-ii baafal ŋgal 'I opened the door'
	-t	Repetitive (REP)	o haal-t-ii 'he spoke again'
	-d	Comprehensive (COM)	mi udd-id-ii baafe d'e 'I closed all the doors'
	-n	Causative (CAU)	mi jaŋŋg-in-ii 'I taught'
	-r	Modal (MOD)	mi dog-r-ii paɗe 'I ran with shoes'

As in Gombe Fula, the consonantal suffixes of Fuuta Tooro each have -C and -VC allomorphs. The -VC allomorph occurs after geminates, consonant clusters, and variably after some single consonants; the -C allomorph occurs elsewhere. I do not find evidence for any more than two allomorphs for each (-C and -iC) in Fuuta Tooro, since alternations in the vowel can be accounted for via harmony rules.

Note that I propose only six consonantal suffixes for Fuuta Tooro, in comparison to Arnott's eleven for Gombe Fula. One reason for this was hinted at in my discussion of the Gombe Fula suffixes in section 2.1: in some cases where Arnott distinguished separate suffixes, it seems more appropriate to propose a single suffix. For example, the Comprehensive and Associative -d suffixes can reasonably be reduced to a single suffix (I label this 'Comprehensive' above, not to suggest that the Comprehensive meaning is the more basic, but simply as a shorthand). Similarly, Arnott's Modal and Locative -r suffixes correspond to a single Modal suffix here, since the shape, distribution, and function of Modal and Locative are basically identical. Another reason for the smaller number of suffixes shown here is that some of the extensions distinguished by Arnott are simply not used very productively in Fuuta Tooro. For instance, in the speech of this particular consultant, I did not observe instances of Reflexives, Intensives, or Retaliatives formed with -t.

3.1.1. The -d' suffix

The use of the Denominative suffix does not appear to be very robust in Fuuta Tooro, as evidenced by the fact that the consultant volunteered forms without the Denominative when given English sentences meant to elicit Denominative forms (for example, 'He has become well again,' 'We all became poor together,' and 'He got well with medicine'). In one accepted Denominative example, the Denominative is combined with the Causative -n, and the Denominative occurs in first position (both extensions appear in bold): *mi dom-**d'-in-ii** mo* 'I made him thirsty'. This is as expected, since the Causative attaches to verb stems, so the root would have to be converted to a verb by the Denominative suffix before

accepting the Causative suffix. I will not discuss the Denominative or its ordering properties in further detail since the Denominative suffix is not common and it is not a verbal suffix, strictly speaking.

3.1.2. *The -t suffixes*

The *-t* suffixes have two primary meanings: Separative and Repetitive. I treat these as two separate suffixes because, as will be discussed, they have different distributions and different ordering properties when combined with other suffixes. The Fuuta Tooro consultant does not volunteer forms with a *-t* suffix to give the Reflexive, Retaliative, or Intensive meanings as found in Gombe Fula.

3.1.2.1. *The Separative -t suffix*

The Separative suffix corresponds roughly in meaning to Arnott's (1970:340) Reversive suffix. I have relabelled this suffix as Separative because in Fuuta Tooro, the Separative appears only to occur with verbs that involve putting objects together in some way, so that the extended verb has a meaning relating to the separation of objects⁵. Examples are shown below.

- | | | | | |
|------|-----|------------------------|--------|------|
| (16) | mi | udd-it-ii | baafal | ɲgal |
| | 1sg | close-SEPAR-past | door | det. |
| | | 'I opened the door' | | |
| | mi | soom-t-ii | gawri | |
| | 1sg | bundle-SEPAR-past | millet | |
| | | 'I un-bundled millet' | | |
| | o | sok-t-ii | baafal | ɲgal |
| | 3sg | lock-SEPAR-past | door | det. |
| | | 'he unlocked the door' | | |
| | mi | ha66-it-ii | foggol | ɲgol |
| | 1sg | tie-SEPAR-past | rope | det. |
| | | 'I untied the rope' | | |

3.1.2.2. *The Repetitive -t suffix*

The Repetitive suffix is homophonous with the Separative suffix, but is less restricted in its distribution than the Separative suffix. It appears that virtually any verb can have a Repetitive form. Those verb roots that can take the Separative suffix have homophonous forms with Repetitive meanings; for example, *mi ha66-it-ii foggol ɲgol* can mean either 'I untied the rope' or 'I tied the rope again'.

- (17)
- | | | | |
|-----|----------------------------------|--------------|------|
| mi | yaa-t-ii | | |
| 1sg | go-REPET-past | | |
| | ‘I went again’ | | |
| o | haal-t-ii | | |
| 3sg | speak-REPET-past | | |
| | ‘he spoke again’ | | |
| o | def-t-ii | faataata | |
| 3sg | cook-REPET-past | sweet potato | |
| | ‘he cooked a sweet potato again’ | | |
| min | cok-t-ii | baafal | ɲgal |
| 1pl | lock-REPET-past | door | det. |
| | ‘we locked the door again’ | | |
| mi | udd-it-ii | baafal | ɲgal |
| 1sg | close-REPET-past | door | det. |
| | ‘I closed the door again’ | | |

3.1.3. The Comprehensive/Associative *-d* suffix

The *-d* suffix in Fuuta Tooro seems to have the same functions as in Gombe Fula. As was discussed earlier, the Comprehensive and Associative *-d* suffixes distinguished by Arnott (1970:346) on the basis of their syntactic distribution are probably better analyzed as a single suffix with a pluralizing and/or comprehensive meaning. The same is true in Fuuta Tooro, since there is no phonological, morphological, or semantic distinction between Comprehensive and Associative, and since the meanings are similar. Some uses of this suffix are shown in the examples below.

- (18)
- | | | | |
|-----|-------------------------------------|--------|-------|
| o | haal-d-ii | e | am |
| 3sg | speak-ASSOC-past | with | 1sg |
| | ‘he spoke with me’ | | |
| mi | yaa-d-ii | e | makko |
| 1sg | go-ASSOC-past | with | 3sg |
| | ‘I went with her’ | | |
| 6e | ɲgudd-id-ii | baafal | ɲgal |
| 3pl | close-COMP-past | door | det. |
| | ‘they all closed the door together’ | | |
| min | cok-d-ii | baafal | ɲgal |
| 1pl | lock-COMP-past | door | det. |
| | ‘we all locked the door together’ | | |

mi	ha66-id-ii	6oggi	di
1sg	tie-COMP-past	ropes	det.

‘I tied all the ropes’

The English translation of the example *mi ha66-id-ii 6oggi di* ‘I tied all the ropes’ is consistent with at least two interpretations: one in which all of the ropes are tied in sequence, and one in which all are tied simultaneously. However, only the latter reading occurs in Pulaar. In every example where the Comprehensive applies to plural objects of an action, the action is understood to occur all at once rather than iteratively.

3.1.4. The *-n* Causative suffix

The *-n* Causative suffix appears to behave identically to the Gombe Fula Causative. Although the Fuuta Tooro consultant often first volunteers a periphrastic construction rather than using *-n* when a Causative is elicited, he uses *-n* productively with a wide range of verb roots when prompted to give a ‘shorter’ form. The consultant reports that the *-n* suffix is used more commonly in other dialects than in his own, but he nonetheless judges Causative forms using *-n* to be correct and natural. Some examples are shown below.

- (19)
- | | | | | |
|-----|---------------|-----|--------|------|
| o | ha66-in-ii | kam | 6oggol | ngol |
| 3sg | tie-CAUS-past | 1sg | rope | det. |
- ‘he made me tie the rope’
- | | | |
|-----|-----------------|-----|
| o | jaɲɲg-in-ii | kam |
| 3sg | learn-CAUS-past | 1sg |
- ‘he taught me’
- | | | |
|-----|---------------|-----|
| min | ñaam-n-ii | mo |
| 1pl | eat-CAUS-past | 3sg |
- ‘we fed her’
- | | | |
|-----|-----------------|-----|
| 6e | njaal-n-ii | mo |
| 3sg | laugh-CAUS-past | 3sg |
- ‘they made him laugh’
- | | | |
|-----|---------------|-----|
| mi | dog-n-ii | 6e |
| 1sg | run-CAUS-past | 3pl |
- ‘I made them run’
- | | | | | |
|-----|----------------|-----|------|------|
| o | irt-in-ii | kam | supu | o |
| 3sg | stir-CAUS-past | 1sg | soup | det. |
- ‘he made me stir the soup’

o	ñoot-in-ii	kam	simis	o
3sg	sew-CAUS-past	1sg	shirt	det.
'she made me sew the shirt'				

3.1.5. The Modal -r suffix

As discussed with respect to the Gombe Fula *-r* suffixes, the Modal/Instrumental and Locative *-r* suffixes may be best analyzed as a single suffix in Gombe Fula, and the same is true of these suffixes in Fuuta Tooro. Both suffixes introduce a noun phrase, and both indicate something about the way in which an action is done—in what way, using what tool, or in what location. I therefore treat these suffixes as a single Modal suffix. In this dialect, the most common usage of *-r* is the Instrumental. Examples of its use are provided below.

- (20) mi udd-ir-ii baafal ŋgal sawru
 1sg close-MODAL-past door det. stick
 'I closed the door with a stick'
- mi haɓɓ-ir-ii ɓoggol ŋgol juude am
 1sg tie-MODAL-past rope det. hands 1sg
 'I tied the rope with my hands'
- ɓe talɗ-ir-ii ɓoggol ŋgol lafi
 3pl cut-MODAL-past rope det. knife
 'they cut the rope with a knife'
- mi irt-ir-ii supu o kuddu
 1sg stir-MODAL-past soup det. spoon
 'I stirred the soup with a spoon'
- o sok-r-ii baafal ŋgal coktirgal
 3sg lock-MODAL-past door det. key
 'she locked the door with a key'
- min ñoot-ir-ii simis o meselal
 1sg sew-MODAL-past shirt det. needle
 'I sewed the shirt with a needle'
- o dog-r-ii paɗe
 3sg run-MODAL-past shoes
 'he ran with shoes'

Now that the consonantal suffixes have been introduced, in the following section I present data showing the relative order of these suffixes when they occur in combination.

3.2. Order of the consonantal suffixes

This study of suffix ordering was undertaken to test the extent to which the fixed ‘TDNR’ order proposed by Arnott (1970) is upheld, particularly in cases where adherence to this ordering principle is in conflict with the order predicted by the Scope Hypothesis (Rice 2000) and other proposals relating order to scope (Baker 1985, Bybee 1985, Condoravdi and Kiparsky 1998), since as discussed earlier, Arnott (1970) does not provide any examples where the order of suffixes does not correspond to their scope. As I will demonstrate in this section, the order of consonantal suffixes in Fuuta Tooro corresponds closely with their scope, and the ‘TDNR’ generalization plays no role. In this section, I go through each possible pairwise combination of the ‘TDNR’ suffixes, giving examples of each and discussing the extent to which the ordering of each pair corresponds to their scope⁶. After looking at the Fuuta Tooro data and a theoretical account of affix order in this dialect, we will return in section 4 to Arnott’s (1970) Gombe Fula data to determine whether scope may also be a better predictor of affix order in that dialect as well. Note that in the discussion in this section, when I refer to the ‘Scope Hypothesis’, this is meant to include not only the specific hypothesis by that name advanced by Rice (2000), but also the previous proposals relating affix order to scope (Baker 1985, Bybee 1985, Condoravdi and Kiparsky 1998).

The relative ordering of the Separative *-t* with the Comprehensive *-d* exhibits an alternation that correlates directly with scope. When the Separative has scope over the Comprehensive, the Comprehensive *-d* is ordered before the Separative *-t* (21). We know that Separative has scope over Comprehensive in these examples because the action is described iteratively. Recall from section 3.1.3 that when the Comprehensive applies to multiple objects, the action is understood to take place on all objects simultaneously. When Separative applies to the result of a Comprehensive action, however, the ‘undoing’ action does not necessarily take place simultaneously, since Separative does not contribute a simultaneity meaning. Thus, the iterative reading in (21) follows directly from the fact that Separative has scope over Comprehensive, as reflected in the ordering of the Separative outside the Comprehensive.

- (21) D-T
 mi udd-id-it-ii baafe d’e fof
 1sg close-COM-SEP-past door det. all
 ‘I opened all the doors (in sequence)’
- D-T
 mi haββ-id-it-ii boggi d’i fof
 1sg tie-COM-SEP-past ropes det. all
 ‘I untied all the ropes (in sequence)’

D-T

o	sok-d-it-ii	baafe	dɛ	fof
3sg	lock-COM-SEP-past	doors	det.	all

'he unlocked all the doors (in sequence)'

As predicted by the Scope Hypothesis, when the Comprehensive has scope over the Separative, the Separative *-t* is ordered before the Comprehensive *-d* (22). We know that Comprehensive has scope over Separative in these examples because the 'undoing' action in each example occurs all at once. This meaning of simultaneity is contributed by the Comprehensive, and the only way that the simultaneity can apply to the 'undoing' action (and not just the original action) is if the Comprehensive applies to a stem that already includes the Separative meaning. Thus, the Separative attaches first, and the Comprehensive attaches to the output of Separative affixation, as reflected in the ordering of the Comprehensive outside the Separative.

(22) T-D

mi	udd-it-id-ii	baafe	dɛ	fof
1sg	close-SEP-COM-past	door	det.	all

'I opened all the doors (at once)'

T-D

mi	haɓɓ-it-id-ii	ɓoggi	di	fof
1sg	tie-SEP-COM-past	ropes	det.	all

'I untied all the ropes (at once)'

T-D

o	sok-t-id-ii	baafe	dɛ	fof
3sg	lock-SEP-COM-past	doors	det.	all

'he unlocked all the doors (at once)'

The ordering of the Repetitive with the Comprehensive is also consistent with scope. The Repetitive *-t* is ordered after the Comprehensive *-d* when the Repetitive has scope over the Comprehensive (23)⁷. The fact that Repetitive has scope over Comprehensive is evidenced by the fact that the repetitive meaning in each case applies not only to the verb, but also to the same participants referred to by the Comprehensive. For example, *o haal-d-it-ii e am* 'he spoke with me again' means not only that he spoke once before and spoke again, this time with me. It means, more specifically, that he spoke once before with me and spoke again with me. We can understand this if we assume that the Comprehensive applies first to the verb, and then the Repetitive applies to the verb with the Comprehensive already affixed to it. This is reflected in the ordering of the Repetitive *-t* after the Comprehensive *-d*.

(23) D-T
 o haal-d-it-ii e am
 3sg speak-COM-REP-past with 1sg
 ‘he spoke with me again’ (he spoke with me before)

D-T
 mi yaa-d-it-ii e makko
 1sg go-COM-REP-past with 3sg
 ‘I went with her again’ (I went with her before)

D-T
 o def-d-it-ii e makko faataata
 3sg cook-COM-REP-past with 3sg sweet potato
 ‘he cooked a sweet potato with her again’ (he cooked a sweet potato with her before)

When the Comprehensive has scope over the Repetitive (24), the Repetitive *-t* is always ordered first. These examples cannot be produced with the Comprehensive *-d* ordered first, because this produces readings as in (23) above where the same participants were involved in both the original and repeated action. In the examples shown below, the evidence for Comprehensive having scope over Repetitive is that the same participants are not necessarily involved in both the original and repeated actions. The Repetitive applies only to the verb, and then the Comprehensive applies to the output of Repetitive affixation, which is a repeated action. Thus, the subjects/objects referred to by the Comprehensive participate in the repeated action but not necessarily in the original action. The ordering of the Comprehensive *-d* outside the Repetitive *-t* is therefore consistent with the scope of the suffixes.

(24) T-D
 6e ngudd-it-id-ii baafal ηgal
 3pl close-REP-COM-past door det.
 ‘they all closed the door again together’ (someone else closed it before)

T-D
 6e kaββ-it-id-ii βoggol ηgol
 3pl tie-REP-COM-past rope det.
 ‘they all tied the rope again together’ (someone else tied it before)

T-D
 min cok-t-id-ii baafal ηgal
 1pl lock-REP-COM-past door det.
 ‘we all locked the door again together’ (someone else locked it before)

Although as mentioned in section 3.1.4 the Causative *-n* is apparently not as commonly used in this dialect as are the other consonantal extensions, the consultant nonetheless has clear intuitions regarding its use and compatibility with other verbal extensions. As shown in (25), the Causative *-n* is ordered after the Separative *-t*. This is consistent with the scope of the suffixes, since the Causative refers not to the original action, but to the ‘undoing’. Thus, the Causative applies to a verb that already has the separative meaning, which is consistent with the ordering of the Causative *-n* outside the Separative *-t*.

- (25) T-N
 o udd-it-in-ii kam baafal ŋgal (*o udd-in-it-ii kam)
 3sg close-SEP-CAU-past 1sg door det.
 ‘he made me open the door’
- T-N
 o haββ-it-in-ii kam βoggol ŋgol (*o haββ-in-it-ii kam)
 3sg tie-SEP-CAU-past 1sg rope det.
 ‘he made me untie the rope’

If order is scope-based, we predict that the opposite ordering of these affixes should correspond to the opposite scope relation between the two, as was seen in the examples shown above where Separative-Comprehensive and Repetitive-Comprehensive were combined. In the case of Causative-Separative, however, it is impossible to find an ordering alternation corresponding to a meaning change because it is apparently impossible for Separative to have scope over Causative. This can be explained by the fact that Separative generally applies to a verb whose semantics involve putting things together. Thus, in order for Separative to apply to a Causative, the entire Causative verb would have to have a ‘putting together’ meaning. There are apparently no verbs corresponding to ‘make be together’ that use the Causative suffix, such that a Separative would be expected to attach to the Causativised stem. Even if the Separative were found to have the less restricted meaning of the Reversive in Gomba Fula and were therefore not limited to verbs with the ‘putting together’ meaning, we would not necessarily expect to find forms where Reversive had scope over Causative. This is because of the pragmatically marked nature of ‘uncausing’. Perhaps the nature of ‘causing’ in Fuuta Tooro (as in English) is such that it generally cannot be reversed, except perhaps in some very specific and/or uncommon contexts. Given the apparent impossibility of ‘uncausing’, the fixed order of the Separative before the Comprehensive is predicted by the Scope Hypothesis.

When the Repetitive *-t* combines with the Causative *-n*, both orderings are acceptable, corresponding to scope. When the Repetitive has scope over the Causative, the Causative *-n* precedes the Repetitive *-t* (26). The scope relation

is made clear by the fact that in these examples, it is required that the same agent caused both the original action and the repetition of the action. Thus, the repetitive meaning must apply to the Causativised verb, corresponding to the ordering of the Repetitive suffix outside the Causative suffix. Beneath the English glosses in the examples below, I give a bracketed version of the gloss to show how the wide scope of the repetitive corresponds to the ‘same agent’ meaning implied in each example sentence.

(26) N-T

o jaŋŋg-in-it-ii kam
 3sg learn-CAU-REP-past 1sg
 ‘he taught me again’ (he taught me before)
 [[he taught me] again]

N-T

min ñaam-n-it-ii mo
 1pl eat-CAU-REP-past 3sg
 ‘we fed her again’ (we fed her before)
 [[we fed her] again]

N-T

o d’aan-in-it-ii kam
 3sg sleep-CAU-REP-past 1sg
 ‘she put me to sleep again’ (she put me to sleep before)
 [[she put me to sleep] again]

N-T

o sood-in-it-ii een deftere nde
 3sg buy-CAU-REP-past 1pl book det.
 ‘she made us buy the book again’ (she made us buy the book before)
 [[she made us buy the book] again]

As predicted by the Scope Hypothesis, the opposite order of the Causative and Repetitive suffixes corresponds to the opposite scope relation from that seen in the examples given above. When the Causative has scope over the Repetitive (27), the Repetitive *-t* suffix precedes the Causative *-n*. The scope relation is evidenced by the fact that in each of these sentences, the original action is understood to have been done voluntarily rather than being caused by the same agent who causes the repeated action. Thus, the Repetitive applies to the bare verb, and the Causative applies to the Repetitive verb, meaning that the causation applies to the repeated action (and not necessarily to the original action). As in the examples above, I provide bracketed glosses under the English glosses in each example below to illustrate the narrow scope of the Repetitive.

(27) T-N

o jaŋŋg-it-in-ii kam
 3sg learn-REP-CAU-past 1sg
 ‘he made me learn again’ (I learned before voluntarily)
 [he made me [learn again]]

T-N

min ñaam-t-in-ii mo
 1pl eat-REP-CAU-past 3sg
 ‘we made her eat again’ (she ate before voluntarily)
 [we made her [eat again]]

T-N

o ñaam-t-in-ii kam
 3pl eat-REP-CAU-past 1sg
 ‘he made me eat it again’ (I ate it before voluntarily)
 [he made me [eat it again]]

T-N

o sood-it-in-ii een deftere nde
 3sg buy-REP-CAU-past 1pl book det.
 ‘she made us buy the book again’ (we bought the book before voluntarily)
 [she made us [buy the book again]]

The relative order of the Separative *-t* and Modal *-r* corresponds to their scope. In the examples below in (28), the Modal has scope over the Separative, as indicated by the fact that the instrument in each example is used to undo the action and not necessarily to do the original action. For example, in ‘I opened the door with a stick,’ it is understood that the stick is used to open the door, not that the stick is used to close the door, as would be the case if the Separative had scope over the Modal. Thus, the scope of the two suffixes in these examples corresponds to the ordering of the Modal *-r* outside the Separative *-t*.

(28) T-R

mi udd-it-ir-ii baafal ŋgal sawru (*mi udd-ir-it-ii)
 1sg close-SEP-MOD-past door det. stick
 ‘I opened the door with a stick’

T-R

o ha66-it-ir-ii boggol ŋgol juŋŋgo makko (*o ha66-ir-it-ii)
 3sg tie-SEP-MOD-past rope det. hands 3sg
 ‘he untied the rope with his hands’

T-R

a sok-t-ir-ii baafal ŋgal coktirgal (*a sok-r-it-ii)
 2sg lock-SEP-MOD-past door det. key
 ‘you (sg.) unlocked the door with a key’

It is apparently impossible to produce a single verb form where Separative has scope over Modal. When asked to produce such a form corresponding to, e.g., ‘we un-sewed the shirts with a needle,’ ([we un-[sewed the shirts with a needle]]) where the needle was used to do the sewing but not the unsewing, the speaker is unable to express this in Pulaar with a single verb. The forms **min ñoot-ir-it-ii* and **min ñoot-it-ir-ii* are unequivocally rejected. Therefore, we are unable to test the prediction of the Scope Hypothesis that the Separative suffix should occur after the Modal suffix when Separative has scope over Modal.

When Modal has scope over Repetitive, as in the examples shown in (29) below, the Modal *-r* suffix is ordered after the Repetitive *-t* suffix. This is as predicted by the Scope Hypothesis. It is clear in these examples that Modal has scope over Repetitive because in each example, it is specified that a different instrument is used to do the original vs. the repeated action. This reading follows if the Repetitive applies to the verb first, and then the Modal applies to the Repetitive stem, such that the use of the instrument introduced by the Modal applies to the repeated action, but not necessarily to the original action.

- (29) T-R
 o udd-it-ir-ii baafal ñgal sawru wofñdu (*o udd-ir-it-ii)
 3sg close-REP-MOD-past door det. stick different
 ‘he closed the door again with a different stick’
- T-R
 mi irt-it-ir-ii supu o kuddu godf’o (*mi irt-ir-it-ii)
 1sg stir-REP-MOD-past soup det. spoon different
 ‘I stirred the soup again with a different spoon’
- T-R
 o sok-t-ir-ii baafal ñgal coktirgal godñgal (*o sok-r-it-ii)
 3sg lock-REP-MOD-past door det. key different
 ‘she locked the door again with a different key’
- R
 mi udd-ir-ii baafal ñgal juufe am
 1sg close-MOD-past door det. hands 1sg
 ‘I closed the door with my hands . . .
- T-R
 mi udd-it-ir-ii baafal ñgal sawru (*mi udd-ir-it-ii)
 1sg close-REP-MOD-past door det. stick
 then I closed the door again with a stick’

When Repetitive has scope over Modal, the Modal *-r* suffix is ordered after the Repetitive *-t* suffix, as shown in the examples below in (30).

- (30) T-R
 mi udd-it-ir-ii baafal ŋgal sawru (*mi udd-ir-it-ii)
 1sg close-REP-MOD-past door det. stick
 'I closed the door with a stick again' (the same stick)
- T-R
 mi ha66-it-ir-ii ɓoggol ŋgol juude am (*mi ha66-ir-it-ii)
 1sg tie-REP-MOD-past rope det. hands 1sg
 'I tied the rope with my hands again'
- T-R
 o sok-t-ir-ii baafal ŋgal coktirgal (*o sok-r-it-ii)
 3sg lock-REP-MOD-past door det. key
 'she locked the door with a key again' (the same key)

This is the first example we have seen in which the order of suffixes does not correspond to their scope. Based on the Scope Hypothesis, we would have expected the Repetitive *-t* to be ordered after the Modal *-r* in these examples. We know that Repetitive has scope over Modal in these examples because it is understood in each example that the same instrument is used to do both the original and repeated actions. This corresponds to the application of the Modal to the verb root, and then application of the Repetitive to the verb that already has an instrument, such that the repetition of the action involves the use of the same instrument. Since this ordering is fixed and inviolable with no apparent semantic explanation for the rejection of the **-r-t* order, I assume the *-t-r* order is fixed as part of the morphological template. This will be accounted for in the analysis to be presented in section 3.3.

The ordering of the Causative *-n* with the Comprehensive *-d* depends upon their relative scope. When the Comprehensive has scope over the Causative, the Causative *-n* precedes the Causative *-d* (31), and forms with the order *-d-n* are not compatible with this reading. Below each English gloss in the examples below, I give a bracketed gloss indicating that Comprehensive has wide scope, so that the 'joint action' meaning applies to the Causativised verb, not just to the bare verb root. As can be seen, the order of the suffixes corresponds to their scope.

- (31) N-D
 ɓe jaŋŋ-in-id-ii mo
 3pl learn-CAU-COM-past 3sg
 'they taught him together'
 [[they taught him] together]
- N-D
 ɓe ñaam-n-id-ii rawaandu ndu
 3pl eat-CAU-COM-past dog det.
 'they fed the dog together'
 [[they fed the dog] together]

N-D
 6e njal-n-id-ii mo
 3pl laugh-CAU-COM-past 3sg
 ‘we all made him laugh together’
 [[we all made him laugh] together]

When Causative has scope over Comprehensive, the Causative *-n* is ordered after the Comprehensive *-d*, as predicted by the Scope Hypothesis. However, there is an added complication that the opposite ordering, *-n-d*, is also apparently compatible with this reading, as seen in the examples below in (32).

- | | | | | | | | |
|----------|----------------------------------|-----|---|-----|--------------------|-----|--|
| (32) D-N | | | | N-D | | | |
| mi | woy-d-in-ii | 6e | ~ | mi | woy-n-id-ii | 6e | |
| 1sg | cry-COM-CAU-past | 3pl | | 1sg | cry-CAU-COM-past | 3pl | |
| | ‘I made them cry together’ | | | | | | |
| D-N | | | | N-D | | | |
| a | dog-d-in-ii | min | ~ | a | dog-n-id-ii | min | |
| 2sg | run-COM-CAU-past | 1pl | | 2sg | run-CAU-COM-past | 1pl | |
| | ‘you (sg.) made us run together’ | | | | | | |
| D-N | | | | N-D | | | |
| mi | jaŋŋg-id-in-ii | 6e | ~ | mi | jaŋŋg-in-id-ii | 6e | |
| 1sg | learn-COM-CAU-past | 3pl | | 1sg | learn-CAU-COM-past | 3pl | |
| | ‘I made them learn together’ | | | | | | |

This may be due simply to the difficulty of constructing an English stimulus where Causative unambiguously has scope over Comprehensive. For example, in ‘I made them learn together,’ where the intended meaning is that the causees are made to learn with each other, there is another possible interpretation where the causees are each made separately to learn; e.g., they were taught simultaneously to do two different things. Thus, when the speaker is presented with English sentences such as these, he may interpret them such that the Comprehensive has over the Causative in these examples, and this would explain why the ordering *-n-d* is accepted in these examples. The meaning difference is too subtle and context-dependent to elicit, so I leave this issue for further investigation in conversational and narrative situations. The Scope Hypothesis predicts that if a context can be found for these verbs in which Causative has unambiguous Scope over Comprehensive, then the order of the suffixes will be *-d-n*, and *-n-d* will be disallowed.

When the Comprehensive combines with the Modal so that Comprehensive has scope over Modal, the Comprehensive *-d* is ordered after the Modal *-r*, as seen in the examples in (33). The scope relation is indicated by the fact that in

these examples, a different instrument is used to perform the action on each object.

- (33) R-D
 mi sok-r-id-ii baafe de coktirgal godɲgal
 1sg lock-MOD-COM-past doors det. key different
 ‘I locked each of the doors with a different key’
- R-D
 o udd-ir-id-ii baafe de sawru wodɲdu
 3sg close-REP-MOD-past doors det. stick different
 ‘he closed each of the doors with a different stick’

When the Modal has scope over the Comprehensive, as in the examples in (34), the Modal *-r* is ordered after the Comprehensive *-d*. The fact that the Modal has scope over the Comprehensive in these examples is evidenced by the fact that the same instrument is used to perform the action on each object. This is consistent with the application of the Modal (introducing the instrument) to a verb stem that already has the comprehensive meaning, such that the instrument applies to all of the objects referred to by the Comprehensive.

- (34) D-R
 mi sok-d-ir-ii baafe de coktirgal
 1sg lock-COM-MOD-past doors det. key
 ‘I locked all of the doors with a key’ (the same key)
- D-R
 mi ñoot-id-ir-ii simisaaji meselal
 1sg sew-COM-MOD-past shirts needle
 ‘I sewed all the shirts with a needle’ (the same needle)
- D-R
 o talɗ-id-ir-ii boggi di lafi
 3sg cut-COM-MOD-past ropes det. knife
 ‘she cut all the ropes with a knife’ (the same knife)

The final possible pairwise combination of consonantal suffixes exhibits some interesting variation that does not follow from the Scope Hypothesis. When the Causative has scope over the Modal, we expect the Causative *-n* to be ordered outside the Modal *-r*. We do find this order corresponding to this scope reading. However, as shown in (35), the opposite ordering of these suffixes can also yield the same scope reading. That is, for each example, either order of these suffixes is permitted with no apparent meaning difference corresponding to the two orderings.

- (35) R-N
 o irt-ir-in-ii kam supu o kuddu ~
 3sg stir-MOD-CAU-past 1sg soup det. spoon
- N-R
 o irt-in-ir-ii kam supu o kuddu
 3sg stir-CAU-MOD-past 1sg soup det. spoon
 ‘he made me stir the soup with a spoon’ (I used a spoon)
- R-N
 o ñoot-ir-in-ii kam simis o meselal ~
 3sg sew-MOD-CAU-past 1sg shirt det. needle
- N-R
 o ñoot-in-ir-ii kam simis o meselal
 3sg sew-CAU-MOD-past 1sg shirt det. needle
 ‘she made me sew the shirt with a needle’ (I used a needle)

This variable order is problematic with respect to the Scope Hypothesis. Since there is a clear scope relation between the Causative and Modal in these examples, the Scope Hypothesis predicts that we should find only the scope-based order.

We find the same pattern of variability when the Modal has scope over the Causative, as in (36). Here, we expect that the Modal *-r* should be ordered after the Causative *-n*, but we find that the opposite order can also be used with the same meaning.

- (36) N-R
 o irt-in-ir-ii kam supu o lafi ~
 3sg stir-CAU-MOD-past 1sg soup det. knife
- R-N
 o irt-ir-in-ii kam supu o lafi
 3sg stir-MOD-CAU-past 1sg soup det. knife
 ‘he made me stir the soup with a knife’ (he used a knife)

Again, since there is a clear scope relation between the Causative and Modal, the Scope Hypothesis predicts that we should find only the scope-based order. The fact that the opposite order is also allowed here as well as in (35) will need to be accounted for via a mechanism other than that used to generate scope-based order.

We have now seen examples of each possible pairwise combinations of the consonantal extensions. Based on the above combinations, the generalizations regarding the ordering of consonantal verb suffixes in Fuuta Tooro Pulaar are given in (37).

- (37) a. Repetitive *-t* precedes Modal *-r* regardless of scope
 b. Causative *-n* and Modal *-r* are freely ordered regardless of scope
 c. Otherwise, order is determined by scope

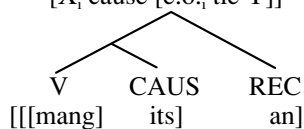
In section section 3.3 below, I provide an analysis of affix order in Fuuta Tooro that accounts for these generalizations.

3.3. A scope/template analysis of Fuuta Tooro affix order

Given the above generalizations, the order of consonantal suffixes in Fuuta Tooro can be analyzed as a mixed Scope-Template system similar to Mirror-Template system in Chichewa described by Hyman (2003), where affix order is determined via the interaction of constraints representing the Mirror Principle (Baker 1985) and a language-specific morphological template. In the case of Fuuta Tooro, the templatic constraints outrank SCOPE, since while scope determines most orderings, some specific templatic orderings apply regardless of the intended scope reading (as in the examples in (30) where the Repetitive *-t* is ordered before Modal *-r* in contradiction to the scope relation between the suffixes).

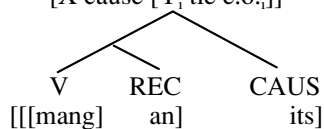
Hyman's (2003) analysis of Chichewa involves argument structure-changing affixes whose scope relations are generally very clear and whose order follows in part from the Mirror Principle, so that the order of affixation 'mirrors' the order of syntactic operations. For example, in combinations of the Causative *-its* suffix and the Reciprocal *-an* suffix, the outer suffix has scope over the inner suffix, as shown in (38) (p. 247).

- (38) a. Reciprocalized Causative
 [X_i cause [$e.o._i$ tie Y]]



'cause each other to tie'

- b. Causativized Reciprocal
 [X cause [Y_i tie e.o.,]]



'cause to tie each other'

A 'CARP' template (where the order of affixes is **C**ausative-**A**pplicative-**R**eciprocal-**P**assive) exerts a different pressure, and it is the interaction of this template with the Mirror Principle that determines Chichewa affix order. The situation is complicated by some forms where orderings obeying the CARP template can have two different scope readings (which Hyman accounts for via variable constraint ranking), but the different forms in (38a) and (38b) are selected when MIRROR(R,C) ('The morphosyntactic input [[[...] REC] CAUS] is

realized Verb-an-its') is ranked above TEMPLATE ('A morphosyntactic input is realized according to CARP'). Sample tableaux are given in (39) (modified from Hyman's tableaux on pp. 251–252).

- (39) a. mang-its-an- 'cause each other to tie'
CAUS-REC

[[[mang] C] R]	MIRROR(R,C)	TEMPLATE
☞ mang-its-an-		
mang-an-its-	*!	*

- b. mang-an-its- 'cause to tie each other'
REC-CAUS

[[[mang] R] C]	MIRROR(R,C)	TEMPLATE
mang-its-an-	*!	
☞ mang-an-its-		*

Note that Hyman (2003) collapses the CARP template into a single constraint. An alternative would be to give separate constraints for each pairwise combination, such that Causative precedes Applicative, Causative precedes Reciprocal, etc. As we will see, an analysis of Fuuta Tooro Pulaar along these lines requires breaking down the TEMPLATE constraint. The three TEMPLATE constraints for Fuuta Tooro are given in (40).

- (40) $T_{REP} > R$: Repetitive *-t* precedes Modal *-r*.
 $N > R$: Causative *-n* precedes Modal *-r*.
 $R > N$: Modal *-r* precedes Causative *-n*.

As will be shown, the ordering of the Repetitive *-t* before Modal *-r* can be enforced by the constraint $T_{REP} > R$ above. The variable ordering of Causative *-n* and Modal *-r* is selected by the constraints $N > R$ and $R > N$ shown above; as we will see, variable ranking of these two constraints with respect to the scope constraint allows us to select forms with either order, regardless of scope.

Hyman's Chichewa examples have affixes that add or change arguments, such that the order of syntactic operations can be deduced. Therefore, the Mirror

Principle and the corresponding MIRROR constraint are appropriate to account for the patterns observed in Chichewa. For Fuuta Tooro, semantic scope-based ordering (Baker 1985, Bybee 1985, Condoravdi and Kiparsky 1998, Rice 2000) seems to be a more appropriate concept to use since many of the affixes we are considering have no effect on argument structure (e.g., the Repetitive) and therefore the order of ‘syntactic operations’ is not obvious. In this analysis, therefore, I will use Condoravdi and Kiparsky’s (1998) SCOPE constraint (41) rather than Hyman’s MIRROR as the constraint that interacts with the templatic constraints.

(41) SCOPE: Morphological constituency reflects scope.

In this analysis, forms involving multiple verbal suffixes will be selected by ranking templatic constraints ahead of SCOPE, such that scope determines the ordering as long as the template is not violated. The tableau in (42) shows the need for the ranking $T_{REP} > R \gg SCOPE$. In this example, the Repetitive has scope over the Modal, as evidenced by the fact that the same instrument is used for both the original and repeated action. Thus, the SCOPE constraint will be satisfied here by the order *-r-t* (Modal precedes Repetitive). However, the optimal output corresponding to this meaning has the opposite order, *-t-r*. This is modelled by ranking $T_{REP} > R$ ahead of SCOPE, so that in examples such as this, satisfaction of the template forces a violation of SCOPE.

(42) mi udd-it-ir-ii baafal ŋgal sawru ‘I closed the door with a stick again’
 REP-MOD (same stick)

/udd, -t, -r/	$T_{REP} > R$	SCOPE
☞ udd-it-ir-		*
udd-ir-it-	*!	

As shown in the tableau in (43), the ranking $T_{REP} > R \gg SCOPE$ also successfully selects forms where the template and scope agree. In this example, the use of a different instrument to do the original vs. repeated action indicates that Modal has scope over Repetitive. Thus, the SCOPE constraint is satisfied by forms with the *-t-r* ordering (i.e., where the Modal is ordered outside the Repetitive). The candidate that satisfies SCOPE also satisfies $T_{REP} > R$, so in this case the winning candidate satisfies both constraints.

- (43) mi irt-it-ir-ii supu o kuddu god'fo 'I stirred the soup again with a different
REP-MOD spoon'

/irt, -t, -r/	T _{REP} > R	SCOPE
☞ irt-it-ir-		
irt-ir-it-	*!	*

For completeness, I demonstrate below that this constraint ranking also correctly selects forms in which the template underdetermines the order, allowing SCOPE to select the optimal candidate (44). In this example, the two suffixes that are being combined are Comprehensive and Separative. The TEMPLATE constraint T_{REP} > R has no bearing on the relative order of these two suffixes, so this constraint is satisfied by either order. Thus, the selection of the output is left to SCOPE. Here, Separative has scope over Comprehensive (evidenced by the iterative rather than simultaneous reading; see examples (21) and (22) and surrounding discussion). Thus, SCOPE is satisfied by a form with the Separative ordered outside the Comprehensive (-*d-t*), so the optimal candidate has this ordering.

- (44) mi ha66-id-it-ii 6oggi d'i 'I untied all the ropes (in sequence)'
COM-SEP

/ha66, -t, -d/	T _{REP} > R	SCOPE
ha66-it-id-		*!
☞ ha66-id-it-		

The ranking of the constraints N > R and R > N is somewhat more complex since these constraints need to produce variable affix order (recall from the examples in (35) and (36) and relevant discussion in section 3.2 that -*n* and -*r* exhibit free ordering regardless of their scope relation). I analyse this phenomenon via two different constraint rankings. In the first ranking, the constraint N > R outranks R > N and SCOPE. This selects forms in which the Causative -*n* is ordered before Modal -*r*, regardless of whether this ordering agrees with the scope of the suffixes. In (45), SCOPE is satisfied by the order -*r-n* since the Causative has scope over the Modal (indicated by the fact that the causee, not the causer, uses the instrument). However, because N > R outranks SCOPE, the form with the opposite order, -*n-r*, is selected.

- (45) o irt-in-ir-ii kam supu o kuddu 'he made me stir the soup with a spoon'
CAU-MOD (I used a spoon)

/irt, -r, -n/	N > R	R > N	SCOPE
☞ irt-in-ir-		*	*
irt-ir-in-	*!		

In (46), the SCOPE constraint is satisfied by the order *-n-r* since Modal has scope over Causative (indicated by the fact that it is the causer, not the causee, who uses the instrument). Since the candidate with the *-n-r* order also satisfies the highly ranked N > R, this candidate is selected.

- (46) o irt-in-ir-ii kam supu o labi 'he made me stir the soup with a knife'
CAU-MOD (he used a knife)

/irt, -n, -r/	N > R	R > N	SCOPE
☞ irt-in-ir-		*	
irt-ir-in-	*!	*	

In order to select forms where Modal *-r* is ordered before Causative *-n*, we need a second constraint ranking where R > N outranks N > R and SCOPE. This ranking selects forms with the ordering *-r-n* whether this order agrees with scope or not. In the example in (47), the SCOPE constraint is satisfied by a form with the order *-n-r* since the Modal has scope over the Causative (indicated by the fact that the causer, not the causee, uses the instrument). However, the highly ranked R > N constraint is satisfied here at the expense of a violation of SCOPE, and the *-r-n* order is selected even though this does not correspond to the scope of the suffixes.

- (47) o irt-ir-in-ii kam supu o labi 'he made me stir the soup with a knife'
MOD-CAU (he used a knife)

/irt, -n, -r/	R > N	N > R	SCOPE
irt-in-ir-	*!		
☞ irt-ir-in-		*	*

In (48), the winning candidate satisfies both $R > N$ and SCOPE. Here, SCOPE is satisfied by a form with the *-r-n* order since Causative has scope over Modal (indicated by the fact that the causee, not the causer, uses the instrument). Since this order also satisfies the highly ranked $R > N$, the winning candidate satisfies both constraints.

- (48) o irt-ir-in-ii kam supu o kuddu 'he made me stir the soup with a spoon'
 MOD-CAU (I used a spoon)

/irt, -r, -n/	$R > N$	$N > R$	SCOPE
irt-in-ir-	*!		*
☞ irt-ir-in-		*	

There is no conflict of $T_{REP} > R$ with $R > N$ and $N > R$. We are thus left with the two constraint rankings shown below.

- (49) Ranking A: $T_{REP} > R, N > R \gg R > N, SCOPE$
 Ranking B: $T_{REP} > R, R > N \gg N > R, SCOPE$

I assume that both rankings coexist in the speaker's grammar, and the choice between which ranking to apply for a given form is freely variable, since there appears to be no principle (semantic, pragmatic, social, or otherwise) behind the variable order of the Causative and Modal suffixes. The above analysis is not meant as an endorsement of variable ranking as the best way of modelling variation in OT; I have used variable ranking here simply because it is a relatively uncomplicated way of representing the observed variation.

In this section, I have proposed a straightforward OT account for the ordering of the consonantal suffixes in Fuuta Tooro. In this account, scope is the primary determiner of affix order, but in some specific combinations of affixes, a fixed ordering can override the scope principle. A full analysis of the order of all affixes in the language may require more morpheme-specific template constraints in addition to the three posited here, but I am claiming that the general scope-driven nature of affix ordering holds throughout the language. Since I have shown how scope drives affix order in Fuuta Tooro, in the next section I revisit Arnott's (1970) Gombe Fula examples and show that the scope generalization can replace the 'TDNR' generalization as a predictor of affix order in that dialect as well.

4. GOMBE FULA REVISITED

The results of the study of Fuuta Tooro Pulaar described above revealed that the Scope Hypothesis is very useful in accounting for the order of consonantal suffixes in that dialect. This raises the question of whether Gombe Fula may also yield to this type of analysis. In this section, I demonstrate that a scope-based analysis is not only consistent with Arnott's (1970) Gombe Fula data, but it also accounts for more of the data than did Arnott's own account involving fixed ordering.

Though the majority of Arnott's (1970) examples do obey the 'TDNR' generalization on the surface, they are all also consistent with the Scope Hypothesis. For instance, the example in (50) below obeys the 'TDNR' generalization, since the order of affixes is *-t-r*. However, this example also conforms to the Scope Hypothesis, since the adverb 'slowly,' which is introduced by the Modal suffix, applies to the Reversive action (opening), not to the original action (closing) which is being reversed. The Modal suffix applies to a verb to which Reversive has already applied, corresponding to the ordering of the Modal *-r* outside the Reversive *-t*.

- (50) T-R
 'o ma66-it-ir-ii yolnde hakkiilo
 3sg close-REV-MOD-past door slowly
 'He opened the door slowly' (p. 367)

The suffix order exhibited in (51) below also conforms to the Scope Hypothesis in addition to the 'TDNR' generalization. In this example, the order of suffixes corresponds directly to the order of logical operations performed on the root. First, the Denominative *-d* suffix attaches to the adjectival root, converting it into a verb stem meaning 'be healthy'. Then, the Repetitive *-t* suffix applies to this verb stem, yielding a new verb stem with the meaning 'be healthy again' (= 'be cured'). Next, the Causative *-n* suffix applies to this verb stem, resulting in a verb stem meaning 'make be cured' (= 'cure'). Finally, the Modal *-r* suffix attaches to this verb stem, introducing an instrument, giving the final meaning 'cure with (some new medicine)'. The order of attachment of the affixes is reflected directly in the order of the consonantal suffixes in this example: *-t-n-r* (Repetitive-Causative-Modal).

- (51) T-N-R
 'o yam-d-it-in-ir-ii mo lekki gokki kesi
 3sg_i healthy-DEN-[REP]-CAU-MOD-past 3sg_j medicine other new
 'He_i cured him_j with some new medicine' (p. 368)

Thus, the ordering generalization that Arnott (1970) accounts for using the ‘TDNR’ generalization can also be accounted for via the Scope Hypothesis. The two examples shown above are perhaps the clearest examples of this, but none of the other examples of the ‘TDNR’ order provided by Arnott (1970) contradicts the Scope Hypothesis.

Not only does the Scope Hypothesis account for Arnott’s (1970) examples obeying the ‘TDNR’ generalization, but it also accounts for the forms that disobey the ‘TDNR’ generalization, which Arnott explained away as having lexicalised stems including the first extension. Recall that Arnott provides five exceptional example types. Of these, three can be explained straightforwardly based on the scope relations between affixes. First, in the form in (52), presumably the intended meaning is for the Repetitive *-t* to have scope over the Comitative *-d*.

- (52) D-T
 mi wol-d-it-at-aa ’e maβbe
 1sg speak-COM-REP-future-negative with 3pl
 ‘I won’t speak with them again’ (p. 368)

Ignoring the Negative for simplicity, the hypothetical form *mi-wol-d-it-ii ’e maβbe* ‘I spoke with them again’ would mean that the subject ‘I’ had spoken with the indirect objects ‘them’ and did so again. This can be schematised as [[speak with] again]. The alternative, probably unintended reading, schematised as [[speak again] with], would be that the subject ‘I’ and the indirect objects ‘they’ have previously performed the act of speaking separately and will do it again together. Thus, in this example, the order of the affixes corresponds to their scope.

Similarly, in (53), the Comitative *-d* most likely has scope over Causative *-n*, since the word ‘fed’ is used in the English translation; the alternative translation would have meant something more like, ‘He made it so that they all ate’.

- (53) N-D
 ’o ñaam-n-id-ii dī
 3sg eat-CAU-COM-past 3pl
 ‘He fed them all’ (p. 368)

Therefore, once again, this apparent exceptional form is explained straightforwardly based on the scope of the suffixes.

Finally, in (54), the Retaliative *-t* must have scope over the Causative *-n*, since the term ‘frighten’ in the English gloss means ‘cause to fear’. Thus, the interpretation of this sentence is [[cause to fear] in turn], which corresponds directly to the ordering of the Retaliative outside the Causative, in violation of the ‘TDNR’ generalization.

- (54) N-T
 mi hul-n-it-oo mo
 1sg fear-CAU-RET-future 3sg
 ‘I’ll frighten him in turn’ (p. 368)

The remaining two exceptional forms, *mi yaa-r-id-ii dī* ‘I took them all’ (p. 368) and *mi war-t-ir-id-an-te dī* ‘I’ll bring them all back to you’ (p. 368) are difficult to interpret since in each case, the specific function of the Modal *-r* is unclear. However, given that three of Arnott’s exceptional forms can be explained based on scope, and that neither the two remaining exceptional forms nor any of the ‘TDNR’ forms contain orderings that violate the Scope Hypothesis, then at the very least, we can say that Arnott did not provide evidence in favour of any principle other than scope for determining the order of consonantal extensions. The scope-based reanalysis allows us to explain the ‘exceptional’ forms, which Arnott chose to ignore. It also avoids the problem that Arnott’s lexicalised stems did not behave like true lexicalised stems, since their meaning was straightforwardly derivable from their component parts rather than being idiomatic as ‘frozen’ forms often are.

We have seen in this section that the principle of scope-based ordering (Baker 1985, Bybee 1985, Condoravdi and Kiparsky 1998, Rice 2000) allows a more complete account of the order of consonantal suffixes in Gombe Fula than did Arnott’s (1970) ‘TDNR’ generalization. As we saw in the previous section, this is also true of the Fuuta Tooro dialect. In the following section, I discuss some theoretical implications of the scope-based analysis of Pulaar affix order that has been presented in this paper.

5. IMPLICATIONS

The analysis that I have presented has important consequences for at least two claims that have been advanced in the theoretical morphology literature. The first is a claim made by Rice (2000) relating to the Scope Hypothesis introduced in section 1.1, and the second is McCarthy and Prince’s (1993a,b) proposed model of the phonology-morphology interface. As I discuss in this section, the facts of Pulaar affix order suggest that aspects of both of these proposals be reconsidered.

5.1. Implications for the Scope Hypothesis

Rice’s (2000) Scope Hypothesis predicts that affix order corresponds to semantic scope. As we have seen in this paper, the facts of Gombe Fula and Fuuta Tooro Pulaar provide support for this prediction, since scope is an excellent predictor of affix order in both dialects. However, based on our analysis of Fuuta Tooro, it

appears that Rice's (2000:395) claim that templates have 'no theoretical status' is too strong. Rice claims that templates will only emerge when there is no scope relation between the affixes in question, but as was seen in section 3.2, there are examples in Fuuta Tooro where a fixed ordering between two suffixes blatantly contradicts the order expected based on their scope relation. In particular, recall that the Repetitive *-t* suffix always precedes the Modal *-r*, even in examples such as (30), where the Repetitive has scope over the Modal and therefore the opposite order is predicted. Examples such as this contradict Rice's claim that templatic ordering occurs only when there is no scope relation between the affixes in question, since in examples such as the one mentioned above, there is a clear scope relation between the affixes and yet their order is fixed.

The priority of scope-based analyses over templates therefore needs to be weakened. It may be better characterized as a methodological imperative for the linguist: look first for a semantic principle to account for affix order; failing this, make use of a template. In a sense, this follows from a more general strategy in descriptive linguistics: when a researcher first encounters a pattern, he/she first looks for a general explanation for it, and if no general explanation is readily available, then (and only then) the researcher proposes a specific mechanism in the grammar to account for the pattern. It does not follow from this strategy that there is 'no theoretical status' for arbitrary statements in grammar.

5.2. *Implications for an OT model of phonology-morphology*

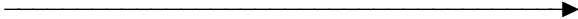
In addition to the implications for Rice's (2000) Scope Hypothesis, Pulaar affix order also has implications for a particular Optimality Theory model of the phonology-morphology interface advanced by McCarthy and Prince (1993a,b). In this section, I discuss this model and how Pulaar *might* have provided an example of a phenomenon predicted by McCarthy and Prince's model. I then show that under the present analysis of affix order in Pulaar, this language does not actually exemplify the predicted phenomenon. This negative result is problematic for McCarthy and Prince's model since, as discussed by Paster (in press), there are no other known examples of this particular phenomenon.

McCarthy and Prince (1993a,b) propose modelling phonological effects in morphology by ranking phonological (P) constraints over morphological (M) constraints in OT, yielding the ranking schema 'P \gg M'. This P \gg M schema accounts for a wide range of phonological effects in morphology, including phonologically conditioned suppletive allomorphy (Mester 1994, Dolbey 1997, Kager 1996), mobile affixes (Noyer 1994; McCarthy and Prince (1993a) suggest a P \gg M analysis), phonologically induced morphological gaps (Prince and Smolensky 1993, but see Orgun and Sprouse 1999), and infix placement (McCarthy and Prince 1993a,b, but see McCarthy 2003, Yu 2003). The P \gg M schema is also claimed to account for phonologically conditioned affix order, and is used for

this purpose by Hargus and Tuttle (1997) to account for the placement of the *s*-Negative prefix in Witsuwit'en, an Athapaskan language of British Columbia.

In the domain of phonologically conditioned affix order, the $P \gg M$ model predicts not only that the placement of an individual affix can be phonologically determined, but that all morphemes in a word made up of several morphemes can line up along some phonological scale. The $P \gg M$ model denies morphological constituency, so the input to each tableau is the set of all morphemes in the word, unordered and with no bracketing or internal structure. Therefore, a highly ranked phonological constraint can be wholly or largely responsible for the order of all morphemes in the word. This is an empirical question: Do we find languages in which a series of several morphemes is ordered along some phonological scale?

A survey of phonologically conditioned affix order reveals a few cases where phonological considerations may affect the placement of a single affix, though all of them are consistent with phonological metathesis or some other explanation that does not require the use of the $P \gg M$ ranking (Paster in press). However, the survey reveals only one case where a *series* of affixes may be claimed to be phonologically ordered: Gombe Fula (Arnott 1970). If Arnott were correct in his claim that the consonantal extensions in Gombe Fula are ordered according to the 'TDNR' generalization, then this could be interpreted as a case of phonologically driven affix order (as suggested by Paster 2001), since the 'TDNR' order corresponds to increasing sonority along the sonority scale (see, for example, Ladefoged 1982), schematised below.

(55)	t	d	n	r
	voiceless	voiced	nasals	liquids
	stops	stops		
				
	sonority			

Imagine for the sake of the argument that the 'TDNR' ordering is fixed in Gombe Fula as claimed by Arnott (1970). In this case, we can account for the ordering using a $P \gg M$ analysis. First, we need a phonological (P) constraint preventing sonority from decreasing across morphemes from left to right in the word (56).

- (56) *FALLINGSONORITY C+C: When a consonant C_1 is followed by a consonant C_2 across a morpheme boundary, C_2 may not be less sonorous than C_1 .


This sonority constraint outranks the M constraint SCOPE (Condoravdi and Kiparsky 1998) discussed in section 3.3 and reproduced below, which requires affix ordering to correspond to scope relations among affixes.

- (57) SCOPE: Morphological constituency reflects scope.

In order to ensure that this constraint affects only the order of consonantal suffixes, we also need to assume some undominated constraints to prevent non-consonantal suffixes from being reordered, and to prevent violations of the *FALLINGSONORITY C+C constraint from being repaired by consonant feature changes rather than reordering (I will not formulate these here, since the analysis being proposed here is hypothetical and is not the analysis that I endorse).

Under this analysis, the ranking of the markedness (P) constraint over the scope (M) constraint selects forms with orderings corresponding to the ‘TDNR’ generalization, even when the morpheme order violates SCOPE. For example, in (58), the order preferred by the SCOPE constraint is *-r-n* since Causative has scope over Modal (the causee, not the causer, uses the instrument). But since *FALLINGSONORITY C+C is satisfied by the opposite order and since this constraint outranks SCOPE, the *-n-r* order is selected.

- (58) *o-irt-in-ir-ii* kam supu o kuddu ‘He made me stir the soup with a spoon’
CAU-MOD (I used a spoon)

/irt, -r, -n/	*FALLINGSONORITY C+C	SCOPE
irt-ir-in-	*!	
 irt-in-ir		*

It is important to note that in this example, *o irt-in-ir-ii* ‘he made me stir...’ is not an attested form, but a hypothetical form constructed for the sake of the argument based on Arnott’s ‘TDNR’ generalization. This was necessary because, as mentioned earlier, none of Arnott’s examples violates SCOPE. In every example provided in the grammar, the order of affixes is consistent with the scope generalization; though some examples are ambiguous, none clearly contradicts what we expect based on scope.

The above ranking still allows for Arnott’s exceptional forms, since in these cases, the first extension that has been lexicalised as part of a frozen stem; that is, it is not evaluated by *FALLINGSONORITY because it is no longer analyzed as a suffix. The surface form follows straightforwardly once the existence of the lexicalised stem has been accepted, but for completeness, I show below how the surface form of the ‘exceptional’ example ‘*o nyaamnidi*’ ‘He fed all (of them)’ is selected rather than *‘*o nyaamdini*’. Since SCOPE is not relevant when only one suffix attaches to the verb, the output is selected by CONTIGUITY (McCarthy and Prince 1995), paraphrased in (59), which disallows insertion between segments of a single morpheme.

- (59) CONTIGUITY: Elements that are adjacent in the input must be adjacent in the output.

A sample derivation of an exceptional form is provided in (60).

- (60) 'o nyaam-n-id-ii 'He fed all (of them)'
CAU-COM

/nyaamn, -d/	*FALLINGSONORITY C+C	SCOPE	CONTIGUITY
☞ nyaamn-id-			
nyaam,di,n-			*!

Both candidates satisfy *FALLINGSONORITY C+C because there is only one extension in the input, so there is no consonant sequence to evaluate. Similarly, neither candidate violates SCOPE since this constraint evaluates the scope of multiple extensions with respect to one another, and so cannot be violated when there is only a single extension in the input. The losing candidate violates CONTIGUITY because [di] is inserted between the segments [m] and [n], which were adjacent in the input. The winning candidate does not violate CONTIGUITY.

In this section, I have presented a $P \gg M$ analysis of extension order in Gombe Fula based on Arnott's 'TDNR' generalization. The use of the $P \gg M$ mechanism for this purpose, while it does capture the pattern, is problematic for at least two reasons. First, the ranking $P \gg M$ is not obviously necessary, since although Arnott claims that the order of suffixes is always 'TDNR', he does not provide any examples where this requirement 'wins out' over the expected scope-based order. Second, and more importantly, there are counterexamples to the 'TDNR' order in Arnott's (1970) data. Arnott explained these as lexicalised forms, but as discussed in section 4, the scope-based analysis is able to handle these forms without having to assume that they involve any lexicalised root-suffix combinations.

The fact that this putative example of the type of comprehensive phonological affix reordering predicted by $P \gg M$ turns out not to be an example of this has an important consequence for the $P \gg M$ model. As mentioned above, Gombe Fula was the only known possible example of this phenomenon, so the fact that this language does not actually exhibit phonological affix order is a devastating negative result for the $P \gg M$ model. Since this major class of effects predicted by the model is not attested, we are led to conclude that the $P \gg M$ model is too powerful and should be abandoned (see Paster (in press) for further discussion). This conclusion converges with other findings such as those of Yu (2003) and Paster (forthcoming), which show that the $P \gg M$ model is not necessary or sufficient to account for infix placement or phonologically conditioned suppletive allomorphy, respectively. Future research will reveal the extent to which the other phenomena accounted for by the $P \gg M$ model may also be better analyzed using alternative mechanisms.

6. CONCLUSION

In this paper, I have described affix order in Fuuta Tooro Pulaar and demonstrated how it follows from semantic scope in combination with a partial morphological template. I have shown that the same scope ordering principle applies also to the Gombe Fula dialect described by Arnott (1970) and that Arnott's proposed fixed 'TDNR' ordering template is not a determining factor in affix order in either Gombe Fula or Fuuta Tooro Pulaar. I have discussed how these findings provide support for Rice's (2000) Scope Hypothesis and other previous proposals relating affix order to semantic scope (Baker 1985, Bybee 1985, Condoravdi and Kiparsky 1998); I have also shown that a specific claim made by Rice (2000) in connection with the Scope Hypothesis is too strong: namely, the claim that templates have no formal status in grammar and emerge only in cases where scope principles cannot apply. As was demonstrated in the analysis of suffix order in Fuuta Tooro, fixed or templatic ordering of affixes needs to be able to outrank or override scope-based ordering in some cases; this indicates that there is a place in the grammar for templatic ordering. Finally, I discussed implications of the reanalysis of Gombe Fula affix order for McCarthy and Prince's (1993a,b) 'P » M' model of the phonology-morphology interface, showing how Gombe Fula is a putative case of a phenomenon predicted by the P » M model that nonetheless does not turn out to exemplify this particular phenomenon. As was discussed, this is problematic for the P » M model because Gombe Fula was the only known possible example of this phenomenon; without it, the model predicts an apparently unattested class of phonological effects in morphology and should therefore be reconsidered in favour of a more restrictive model.

NOTES

* I would like to thank Sharon Inkelas, Lynn Nichols, Larry Hyman, Andrew Garrett, Juliette Blevins, Johanna Nichols, Rebecca Cover, Stefan Elders, Sebastian Ross-Hagebaum, and three anonymous reviewers for helpful comments. This research was funded in part by a National Science Foundation Graduate Research Fellowship.

¹ In Arnott's (1970) orthography, <'> represents glottal stop, <'y> is a palatal implosive, <sh> is a palatal fricative, and <c> and <j> are voiceless and voiced palatal affricates, respectively. The other symbols correspond to their standard values, except that <n> before <g> is pronounced [ŋ]. I have normalized Arnott's transcriptions by using spaces where Arnott used hyphens between subject/object clitics and the verb.

² The Denominative *d* invariably occurs immediately next to the root in all of Arnott's examples, though Arnott omits it from the 'TDNR' formula.

³ I have omitted examples where the verb root and consonantal extensions are identical to or subsumed by another form that is already listed.

⁴ In this and the examples to follow, I use the official Senegalese Pulaar orthography (Hartell 1993:250), which differs from the orthography used by Arnott (1970) in the following ways.

First, while Arnott uses the glottal stop symbol <'> before word-initial vowels, the Senegalese orthography omits these since the presence of glottal stop in this environment is predictable. Second, where Arnott transcribes the palatal nasal with <ny>, this orthography uses <ñ>. Finally, where Arnott uses <'y> for the palatal implosive, the Senegalese orthography uses a 'hooked y', which I replace with <ɟ> here since the hooked y is not available in standard linguistics font sets.

⁵ Thanks to Stefan Elders for pointing this out.

⁶ Combinations of three or more of these suffixes, though occurring occasionally, are generally dispreferred. Therefore, a thorough systematic study of the order of these suffixes is possible only in pairwise combinations.

⁷ The opposite ordering, with *-t* preceding *-d*, is also possible with this general meaning, but with an added idiosyncratic semantic nuance: namely, that the action has taken place so many times that the speaker or the subject has grown tired of it. Since this meaning difference is idiosyncratic and could not have been predicted from the meanings of the component morphemes, I assume the T-D ordering in this particular construction is idiomatic and can be factored out of the scope analysis of the Comprehensive-Repetitive ordering. It is also possible that the *-t* is understood as Intensive in these utterances, though I was unable to elicit the same 'tired of it' meaning in forms using *-t* without *-d*.

REFERENCES

- Anderson, Stephen R. (1986). Disjunctive ordering in inflectional morphology. *Natural Language & Linguistic Theory* 4, 1–32.
- Arnott, D.W. (1970). *The Nominal and Verbal Systems of Fula*. Oxford: Clarendon Press.
- Baker, Mark (1985). The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16, 373–415.
- Bloomfield, Leonard (1962). *The Menomini Language*. New Haven: Yale University Press.
- Bybee, Joan (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins.
- Condoravdi, Cleo, and Paul Kiparsky (1998). Optimal order and scope. Paper presented at Lexicon in Focus, Heinrich-Heine Universität Düsseldorf.
- de Wolf, Paul P. (1985). Some notes on five almost homonymous extensions in Noŋaare Fulani (Western Niger). *Afrika und Übersee* 68, 207–232.
- de Wolf, Paul P. (1986). Verbal extensions in Noŋaare Fulani: Causative, Dative, Celerative, Distantive and Simulative. *Afrika und Übersee* 69, 45–60.
- de Wolf, Paul P. (1987). Further verbal extensions in Noŋaare Fulani: Reciprocal, Iterative, Iterative-Reciprocal and Verbalizing extensions. *Afrika und Übersee* 70, 61–72.
- de Wolf, Paul P. (1991). On four more verbal extensions in Noŋaare Fulani: Associative, Comprehensive, Modal and Locative. *Afrika und Übersee* 74, 49–66.
- Dolbey, Andrew (1997). Output optimization and cyclic allomorph selection. In Brian Agbayani and Sze-Wing Tang (eds.), *The Proceedings of the Fifteenth West Coast Conference on Formal Linguistics*. Stanford: CSLI, 97–112.
- Ethnologue: Languages of the World*, 14th ed. Dallas, Texas: SIL International. <http://www.ethnologue.org/web.asp>.

- Good, Jeff (2003). *Strong linearity: Three case studies towards a theory of morphosyntactic templatic constructions*. Doctoral dissertation, University of California, Berkeley.
- Hargus, Sharon, and Siri G. Tuttle (1997). Augmentation as affixation in Athabaskan languages. *Phonology* 14, 177–220.
- Hartell, Rhonda L. (ed.) (1993). *Alphabets of Africa*. Dakar: UNESCO/SIL.
- Hyman, Larry M. (2003). Suffix ordering in Bantu: A morphocentric approach. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2002*. Dordrecht: Kluwer Academic, 245–281.
- Hyman, Larry M., and Sharon Inkelas (1999). Emergent templates: The unusual case of Tiene. Paper presented at the Hopkins Optimality Workshop / Maryland Mayfest, Baltimore.
- Inkelas, Sharon (1993). Nimboran position class morphology. *Natural Language & Linguistic Theory* 11, 559–624.
- Kager, Rene (1996). On affix allomorphy and syllable counting. In Ursula Kleinhenz (ed.), *Interfaces in Phonology*. Studia grammatica 41. Berlin: Akademie Verlag, 155–171.
- Ladefoged, Peter (1982). *A Course in Phonetics*. New York: Harcourt Brace Jovanovich.
- McCarthy, John J. (2003). OT constraints are categorical. *Phonology* 20(1), 75–138.
- McCarthy, John J., and Alan Prince (1993a). Generalized alignment. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1993*. Dordrecht: Kluwer, 79–153.
- McCarthy, John J., and Alan Prince (1993b). Prosodic morphology I: Constraint interaction and satisfaction. MS, University of Massachusetts, Amherst and Rutgers University.
- McCarthy, John J., and Alan Prince (1995). Faithfulness and reduplicative identity. In Jill Beckman, Suzanne Urbanczyk, and Laura Walsh (eds.), *University of Massachusetts Occasional Papers in Linguistics* 18, 249–384.
- Mester, R. Armin (1994). The quantitative trochee in Latin. *Natural Language & Linguistic Theory* 12: 1–61.
- Noyer, Rolf (1994). Mobile affixes in Huave: Optimality and morphological wellformedness. In E. Duncan, D. Farkas, and P. Spaelti (eds.), *Proceedings of the Twelfth West Coast Conference on Formal Linguistics*. Stanford: CSLI, 67–82.
- Orgun, Cemil Orhan, and Ronald L. Sprouse (1999). From MParse to Control: Deriving ungrammaticality. *Phonology* 16(2), 191–224.
- Paster, Mary (2001). What determines suffix ordering in Fula? Paper presented at the 32nd Annual Conference on African Linguistics, University of California, Berkeley.
- Paster, Mary (in press). A survey of phonological affix order with special attention to Pulaar. In Leah Bateman and Cherlon Ussery (eds.) *Proceedings of the 35th Annual Meeting of the North Eastern Linguistics Society*. Amherst, Massachusetts: GLSA Publications.
- Paster, Mary (forthcoming). *Phonological Conditions on Affixation*. PhD dissertation, UC Berkeley.
- Prince, Alan, and Paul Smolensky (1993). Optimality Theory: Constraint interaction in generative grammar. MS, Rutgers and University of Colorado, Boulder.
- Rice, Keren (2000). *Morpheme Order and Semantic Scope: Word Formation in the Athapaskan Verb*. Cambridge: Cambridge University Press.
- Simpson, Jane, and Margaret Withgott (1986). Pronominal clitic clusters and templates. In Hagit Borer (ed.), *Syntax and Semantics 19: The Syntax of Pronominal Clitics*. New York: Academic Press.

- Speas, Margaret (1990). *Phrase Structure in Natural Language*. Dordrecht: Kluwer.
- Stump, Gregory (1992). On the theoretical status of position class restrictions on inflectional affixes. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1991*. Dordrecht: Kluwer, 211–241.
- Yu, Alan C. L. (2003). *The morphology and phonology of infixation*. Doctoral dissertation, University of California, Berkeley.
- Zwicky, Arnold (1985). How to describe inflection. In M. Niepokuj, M. VanClay, V. Nikiforidou, and D. Feder (eds.), *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 372–386.

*Department of Linguistics
2816 Cathedral of Learning
University of Pittsburgh
Pittsburgh, PA 15260
USA*

mpaster@pitt.edu

Declension hopping in dialectal Croatian: Two predictions of frequency

ANDREA D. SIMS*

1. INTRODUCTION

Are certain words more likely to be affected by analogy than others? How does analogy move through the lexicon? These questions have long been of interest to linguists. Malkiel (1990, 1992) posited “leader words,” while Mańczak’s “tendencies of analogy” implicitly argued that some forms of a word will undergo analogy while others will not.¹ Sturtevant (1917), Prokosch (1939) and later Wang’s (1969) idea of lexical diffusion posit item-by-item movement through the lexicon.² Janda and Joseph’s (2003) Big Bang theory posits spread dependent upon category extension, social factors, etc.³ Many recent theories (e.g., Bybee 1985, 1988, Ogura and Wang 1996, Phillips 1984, papers in Bybee and Hopper 2001) have argued that frequency plays a role in determining which forms will undergo analogy, especially regarding changes from non-productive to productive morphological patterns (e.g. the shift of English verbs from strong to weak past tense markers). Despite the abundance of studies, however, the predictions regarding frequency are not always clear.

This paper investigates the declension of Croatian masculine a-stem nouns. These nouns refer to males, but have traditionally belonged to the a-stem declension class, which overwhelmingly contains feminine nouns. This paper shows that, contrary to previous descriptions, some of these nouns, for some speakers, can innovatively decline according to the o-stem pattern, which is associated with masculine nouns. Through this variation in declension, two predictions of the role of frequency are evaluated. It is suggested that frequency, reflected in the structure of the lexicon, does not directly affect analogical change. Rather, it appears that frequency and social factors together create an upper-boundary constraint on analogical change.

2. FREQUENCY AND LANGUAGE CHANGE

Studies of the effects of frequency on language change have fallen into essentially two groups: direct causation and correlation mediated through the lexicon. In the direct causation camp are theories such as grammaticalisation and lexical diffusion. Grammaticalisation (as formulated by, e.g., Hopper and Traugott 1993) holds that frequent forms are subject to ‘wearing down’ phonetically, eventually leading free words to become morphemes, and then phonemes.⁴ Essentially, frequency leads to (1) ‘sloppy’ pronunciation through repetition and (2) perceptual predictability (and thus less need for careful articulation on the part of the

speaker). Lexical diffusion (e.g., Hooper 1976, Ogura and Wang 1996, Phillips 1984, 2001) is argued to work similarly, with 'physiologically' motivated changes affecting the most frequent words first.

By contrast, analogical change has often been argued to affect infrequent words first, and frequency is viewed as impacting analogical change through lexical organisation. This view draws primarily from psycholinguistic research on frequency effects, which has suggested that frequency (as one instantiation of language use) plays a role in the structure of the lexicon and thus, potentially, in lexically-based language change (i.e. analogy).⁵ Research has correlated frequency with reaction time, showing (at its most basic) that mental processing time for a word is inversely correlated with that word's frequency (e.g., Alegre and Gordon 1998, Bertram et al. 2000, Butterworth 1983, Caramazza et al. 1988, Hay 2003, Schreuder and Baayen 1995, Stemberger and MacWhinney 1988, Taft and Forster 1975). It is argued that lexical structure and/or access must therefore be governed in part by frequency, although the details are debated. Some researchers argue that all words are stored fully inflected in the lexicon and accessed as such, while others argue that while common words and irregular forms may be accessed as fully inflected forms, infrequent words are composed from morphemes.

This paper follows most recent work in assuming the first type of model, which theorises that all words are stored fully-inflected and form connections to morphologically related words (among other types of connections), but the strength of those connections varies in accordance with the frequency of the word. Low frequency words have stronger connections than high frequency words, which can 'stand alone' (i.e. high-frequency forms are entrenched). This makes low frequency words more susceptible to analogical change.

Note that frequency does not cause analogical change. The causative factor in analogy is the connection that a speaker makes between two distinct lexical patterns, and that connection determines which forms are eligible for analogy. Nor is this to say that when an analogical change could happen that it must happen. Rather, usage-based models predict that when morphological change *does* occur, frequency will determine which tokens within this eligible group actually undergo analogical change, and the order in which they undergo it. Specifically, low-frequency words are expected to undergo analogy before high frequency words.

While this theory of the influence of token frequency on analogical change may seem clear, within this idea there are in fact two subtly different hypotheses regarding how frequency affects language change.

Hypothesis 1: The cognitive approach

When analogical language change occurs, it will begin with all infrequent forms, progressing to the frequent ones.

The key thing to note for this hypothesis is the direct connection between representation and production. Many usage-based models have all but removed the concept of grammatical rules. For example, Bybee (1988:119) states that she “takes a different perspective on representations, focusing on the lexicon directly and approaching rules as generalizations that arise from representations.” Due to the connections between low-frequency words and morphologically related forms (note how this is fundamentally analogical), all low-frequency forms should be equally affected by analogical change. In other words, Hypothesis 1 predicts that all things being equal, all low-frequency forms should be ‘selected’ for analogy.

It is not clear, however, that a direct connection is the correct assumption. There are many factors of language use which are not usually taken into consideration by cognitive models. Grammars are used by speakers, and those speakers live in a social world. Consideration of social factors may have a subtle but important influence on our view of how frequency influences morphological change. The cognitive perspective assumes that low frequency words are selected for change because of their connections to related words, but a social salience perspective suggests that in cases of prescriptivism it is in fact high frequency words which are ‘selected’—they are selected to *not* vary.

Salience is defined by Kerswill and Williams (2002:81) as “... a property of a linguistic item or feature that makes it in some way perceptually and cognitively prominent.” It is an extra-linguistic phenomenon, which may have several contributing causes, such as extra-linguistic factors, syntactic, prosodic or pragmatic factors (see Cheshire 1996), or frequency. For example, it is assumed that entrenchment causes high-frequency forms to be more salient. In other words, variation in these forms is more noticeable than variation in low frequency forms. Thus, while frequency contributes to lexical organisation, saliency as a concept is intended to be one step removed from the lexicon and/or grammar itself; it is perceptual. It is important to distinguish between lexical organisation and perceptions of language that may be fed by lexical organisation because conscious changes (accommodation, hypercorrection, etc.) may be driven by people’s perceptions of languages more than by linguistic structure itself.

Salience has been employed (in various ways) within the historical linguistics literature to explain why some features are adopted and others rejected in language change (e.g. Bardovi-Harlig 1987, Cheshire 1996, Kerswill 1985, Mufwene 1991 and Trudgill 1986). One argument (e.g. Yaeger-Dror 2003) is that saliency works in tandem with social markers. A social marker (in the Labovian sense, see, e.g. Labov 1994) is a change in progress that has acquired social recognition, with one variant usually being stigmatised. As such, it may be the subject of hypercorrection, linguistic insecurity, prescriptivism, etc. The hypothesis is that social markers define the candidate class for a change,⁶ and saliency determines which among the candidate forms will be affected, or more precisely, will *not* be affected. Specifically,

Hypothesis 2: The social salience approach

Where a social marker exists, salient forms will be inhibited from showing the stigmatised variant. Since high frequency promotes saliency, low frequency is thus a necessary but not sufficient condition for the appearance of a stigmatised morphological form.

In other words, low-frequency forms should behave ‘randomly’.

Note that Hypothesis 2 assumes the same lexical organisation as does Hypothesis 1. The difference lies only in the consideration of social factors, resulting in different perspectives on how forms are selected for analogical change. In summary, both Hypothesis 1 and Hypothesis 2 predict that low frequency forms will change before high frequency forms. However, by suggesting that low-frequency forms are ‘selected’ for change, a purely cognitive approach implies that infrequent forms will behave uniformly. A social saliency perspective argues that under conditions of prescriptivism, high-frequency forms are the ones ‘selected’ to *not* change, predicting that low-frequency forms will show both realisations.

The group of Croatian⁷ nouns traditionally referred to as masculine a-stems presents an interesting chance to test these two hypotheses. The masculine a-stems are nouns which were ‘caught in the crossfire’ in the development of the modern language from a thematic vowel system to a gender system. As a result, there is a perceived ‘conflict’ between declension class, natural sex of the referent and agreement. The result is a series of associations that promote an analogical shift to a new pattern. In this paper, I primarily investigate the declensions of masculine a-stem nouns among dialect speakers in Split, Croatia, but a foray into agreement will first set the stage for the type of change that is expected.

3. MOTIVATION FOR DECLENSION HOPPING: A BRIEF WORD ON AGREEMENT

Some might argue that semantic/natural gender and morphological gender are not connected beyond the arbitrary naming conventions of grammarians. However, a deeper look shows that the two in fact can be cognitively connected through agreement.

Corbett (1991) discusses the connection between natural and morphological gender for what he terms ‘hybrid nouns’. According to his definition, hybrid nouns are nouns that take mixed gender agreement patterns because the gender assigned by semantic criteria is in conflict with the gender assigned by morphological criteria.⁸ An example from Russian is given in (1) (taken from Corbett 1991). In Russian, the word *vrač* ‘doctor’ may refer to either a male or a female, but the word belongs to a declension class which overwhelmingly contains masculine nouns. When referring to a male, agreement is always masculine. When referring to a female, however, as in (1) (Ivanova is a female name), agreement can be either masculine or feminine, although feminine agreement for attributives is less

common than for other agreement targets and verges on substandard (Corbett 1991:184, Timberlake 1993:837).

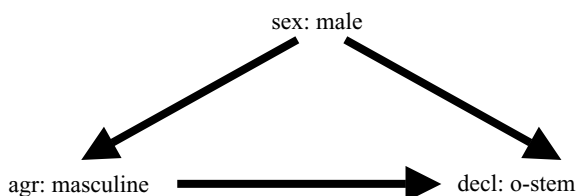
- (1) a. Ivanova—xoroš-ij vrač.
 Ivanova good.MASC doctor.O-STEM
 ‘Ivanova is a good doctor.’
- b. Ivanova—xoroš-aja vrač.
 Ivanova good.FEM doctor.O-STEM
 ‘Ivanova is a good doctor.’

Mixed agreement patterns often arise in the situation in which grammatical gender differs from natural gender, suggesting that speakers make a connection between the two.

Furthermore, we might expect that if speakers connect natural gender and gender agreement, as in (1b), that they might also make a connection between natural gender and morphological realisations of grammatical gender, especially if preceded by a change in agreement. Corbett (1991) captures this idea in an Agreement Hierarchy. According to the Agreement Hierarchy, attributives are least likely to take semantic agreement, and personal pronouns are most likely to. This suggests that the connection between natural and grammatical gender exists most strongly for pronouns, which is not surprising considering their deictic role, and then has the possibility to spread. In other words, Corbett’s Agreement Hierarchy seems to predict that a change in agreement could precede, and possibly trigger, a change in morphology. Slavic masculine a-stem nouns, as examples of hybrid nouns, are no exception.

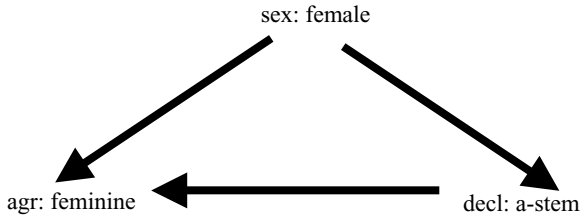
In Slavic, there is a correlation between noun declension, agreement and sex when the noun refers to an animate being. Although the situation varies slightly from language to language, we can take Croatian as an example. In Croatian, nouns referring to male humans almost always take masculine agreement and have o-stem declension.⁹ (The specific morphological endings that define this class are given in section 4.) Masculine agreement also implies o-stem declension, but not the reverse since o-stems may be either masculine or neuter. This may be represented by the following diagram, in which the arrows indicate a generally held implicational relationship.

- (2) Normal pattern for a Croatian noun denoting a male human



Similarly, words denoting female humans are strongly associated with both feminine agreement and a-stem declension. A-stem declension implies feminine agreement, but not the reverse since two distinct declension types trigger feminine agreement.¹⁰

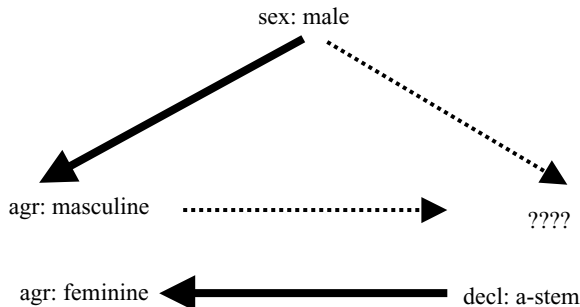
(3) Normal pattern for a Croatian noun denoting a female human



In short, there is usually a high degree of correlation between declension class, gender agreement and sex for nouns that refer to humans.

The pattern for the masculine a-stem nouns is different, however. These nouns refer either to males or are common gender (i.e. refer to both males and females). According to the above diagrams, male reference implies both masculine agreement and o-stem declension. While masculine agreement is indeed normal for these nouns (this issue is discussed later in this section), the nouns have traditionally declined in Croatian and the other Slavic languages according to the a-stem pattern. In other words, the nouns historically have prototypically 'feminine' morphology. Since a-stem declension implies feminine agreement, the pattern of the masculine a-stems may be represented as follows. The question marks and dashed arrows indicate a relationship that might be expected according to the pattern indicated in (2) and (3), but which historically did not obtain.

(4) Pattern for Croatian masculine a-stem nouns



Looking at this relationship, it should not be surprising to discover that in several of the Slavic languages, including Czech, Polish, Byelorussian and non-standard Croatian, masculine a-stem nouns have undergone a partial replacement of

their historically a-stem morphological endings in favour of o-stem counterparts. These analogical changes can be termed declensional gender resolution because the changes were specifically triggered by the ‘conflict’ between masculine reference (and the o-stem declension that it implies) and a-stem (implied feminine) morphology. In short, a change from the a-stem to the o-stem paradigm is not accidental but rather results from the influence of natural gender upon grammatical gender.

Tension between natural and grammatical gender for these nouns in Croatian has been noted by Belić (1924b), Glavan (1928–29) and Marković (1954), among others. This tension has led to an unusual gender agreement pattern. Prescriptively, masculine a-stem nouns have masculine agreement in the singular, and feminine agreement in the plural. As the diagram in (4) indicates, feminine agreement stems from the implicational relationship between a-stem declension and agreement, while masculine agreement stems from the implicational relationship between natural sex and agreement. The singular/plural split is probably the result of animacy marking (see section 4.5.1 for details).¹¹ Examples of this prescriptive agreement pattern for masculine a-stem nouns are given in (5) below (modified from Weschler and Zlatić 2000:813). Example (5a) shows the noun *sudija* ‘judge’ in the singular. Having masculine agreement, it must refer to a male. Example (5b) is the corresponding plural sentence. Although *sudije* ‘judges’ may refer to female judges, feminine agreement is also used to refer to an exclusively male group.

- (5) a. T-aj star-i sudij-a je dobro sudi-o.
 that.MASC old.MASC judge.A-STEM AUX well judged.MASC
 ‘That old (male) judge judged well.’
- b. T-e star-e sudij-e su dobro sudi-le.
 those.FEM old.FEM judges.A-STEM AUX well judged.FEM
 ‘Those old (male) judges judged well.’

Several sources (e.g. Belić (1924a,b), Browne (1993), Corbett (1991), Glavan (1928–29), Marković (1954) and Weschler and Zlatić (2000)) note that the agreement pattern in reality is not this simple, with both feminine and masculine agreement being possible in both the singular and the plural. Corbett, Belić and Marković state that the greater tendency is for masculine (i.e. natural gender) agreement to be extended beyond its prescribed domain, with increasing likelihood the further removed the agreement target. An example of feminine agreement in the singular is given in (6), agreeing with the masculine a-stem noun *sluge* ‘servant.GEN.SG’. Examples of masculine agreement in the plural are given in (7) and (8), modifying the nouns *sluge* ‘servants.ACC.PL’ and *krvopije* ‘bloodsuckers.NOM.PL’. These examples are taken from Glavan (1928–29) and Marković (1954), who cite the work of several classic Čakavian¹² writers.

- (6) Da vidite još rvanje od Lovrinca božje sluge.
 that see.2.PL still wrestling from Lovrinac god's.FEM servant.A-STEM
 'If you could only see the wrestling of God's servant from Lovrinac.'
- (7) Da budemo i mi zvani verni sluge kako krstjani.
 that will also we called true.MASC servants.A-STEM as christians
 'That we too, as Christians, would be called faithful servants.'
- (8) I ov-e krvopij-e narodn-e nisu li se u početku
 and these.FEM bloodsuckers.A-STEM folk.FEM not.AUX Q REFL in beginning
 pokaziva-li kao dobri ljudi.
 demonstrate.MASC as good people
 'And didn't these folk bloodsuckers not in the beginning show themselves
 to be good people?'

In short, in Serbian and Croatian, agreement in the plural, where it is prescriptively feminine, can in reality be feminine or masculine. And with agreement being so variable in the Slavic languages, it might be expected that where we find semantic (i.e. masculine) agreement, we might also find variation in the declension of the noun between the prescriptive and historical a-stem (prototypically feminine) and innovative o-stem (prototypically masculine) paradigms. Such a change would represent an alignment of natural gender and morphological declension class, thus resolving the declensional gender 'conflict' represented in (4).

4. CROATIAN

Although it is commonly accepted that Croatian masculine a-stems decline according to the a-stem pattern, fieldwork suggests that for some speakers an innovative o-stem pattern is possible.

Examples of the masculine and common gender a-stem nouns of Serbian and Croatian are given in (9) and (10) below. Some are of Common Slavic origin (e.g. *sluga* 'servant', *sudija* 'judge'), other roots and suffixes are borrowed from Turkish (e.g. words with *-džija* 'one who sells X'), and still others represent borrowings from other languages (e.g. *kolega* 'colleague'). In Serbian and Croatian the split between masculine a-stem nouns and common gender nouns is partially semantically based; derogatory terms tend to be common gender. On the other hand, some Croatian masculine a-stem nouns refer only to males because they have lexicalised feminine counterparts, e.g. the equivalents of *sluga* 'male servant' and *gazda* 'landlord' are *sluškinja* 'female servant' and *gazdarica* 'landlady'. Still other words simply lack female reference for social or other reasons.

(9) Serbian and Croatian masculine a-stem nouns

<i>gazda</i>	‘landlord’	<i>štediša</i>	‘thrifty person’
<i>vođa</i>	‘leader’	<i>papa</i>	‘Pope’
<i>ubojica</i>	‘warrior; murderer’	<i>sluga</i>	‘servant’
<i>kolega</i>	‘colleague’	<i>Sarajlija</i>	‘Sarajevo resident’
<i>buregdžija</i>	‘maker or seller of <i>burek</i> ’ <i>ždera</i> ‘glutton’		

(10) Serbian and Croatian common gender a-stem nouns

<i>mušterija</i>	‘customer’	<i>izb(j)eglica</i>	‘refugee’
<i>zloća</i>	‘evil/ill-tempered person’	<i>luda</i>	‘crazy person’
<i>pristaša</i>	‘follower’	<i>neplatiša</i>	‘defaulter (on debt)’
<i>varalica</i>	‘imposter’	<i>sudija</i>	‘judge’

The traditional description (e.g., Benson 1994, Browne 1993, Ivić 1994) of these nouns is that they decline according to the a-stem declension in both the singular and plural of all cases. The o-stem and a-stem declensions are given in (11) and (12), followed by the prescriptive declension of masculine and common gender a-stem nouns in (13).

(11) o-stem animate declension (*junak* ‘hero’)

Nom	<i>junak</i>	<i>junac-i</i> ¹³
Voc	<i>junač-e</i> ¹⁴	<i>junac-i</i>
Acc	<i>junak-a</i>	<i>junak-e</i>
Gen	<i>junak-a</i>	<i>junak-a</i>
Dat-Loc	<i>junak-u</i>	<i>junac-ima</i>
Inst	<i>junak-om</i>	<i>junac-ima</i>

(12) a-stem declension (*žena* ‘woman’)

Nom	<i>žen-a</i>	<i>žen-e</i>
Voc	<i>žen-o</i>	<i>žen-e</i>
Acc	<i>žen-u</i>	<i>žen-e</i>
Gen	<i>žen-e</i>	<i>žen-a</i>
Dat-Loc	<i>žen-i</i>	<i>žen-ama</i>
Inst	<i>žen-om</i>	<i>žen-ama</i>

(13) Masculine and common gender a-stem declension (*mladoženja* ‘bridegroom’)

Nom	<i>mladoženj-a</i>	<i>mladoženj-e</i>
Voc	<i>mladoženj-o</i>	<i>mladoženj-e</i>
Acc	<i>mladoženj-u</i>	<i>mladoženj-e</i>
Gen	<i>mladoženj-e</i>	<i>mladoženj-a</i>
Dat-Loc	<i>mladoženj-i</i>	<i>mladoženj-ama</i>
Inst	<i>mladoženj-om</i>	<i>mladoženj-ama</i>

As noted in section 3, variability in *agreement* for these masculine a-stems is widely recognized. However, no grammar, dialect study or other source known to me claims that in Serbian or Croatian there can be *declensional* variation on the noun itself. Marković (1954:87) specifically states that there is no morphological variation.

We can divide all nouns with the ending *a* in the nominative singular into two groups, *among which there is not a morphological difference*, but only a syntactic one . . . The most numerous and most typical representatives are nouns in *-a* which are feminine both in form and in meaning, both in the singular and in the plural. In the other group we could include all those nouns with the ending *a* which are different from the “normal” nouns of the first group. (translation and emphasis mine)

Yet such morphological variation might well be expected, if there is a perceived ‘conflict’ between declension and agreement. And Marković does in fact state that such a conflict exists. “. . . The meaning of [masculine a-stem] nouns comes into conflict with the form which demands feminine gender and thus—in the linguistic feeling of people appears a hesitation . . .” (1954:89, translation mine). In section 4.2 it is shown that morphological variation is possible, for some speakers and some words, as would be expected.

4.1. *Data collection methodology*

The following data was gathered primarily in the fall of 2002 in the city of Split, Croatia. Split is the second largest city in Croatia; it is located on the Adriatic coast in central Dalmatia. The traditional dialect of the city and surrounding area is known as Čakavian. It differs significantly from Standard Croatian, which is based on the Štokavian dialect, and from the dialect of the capital, Zagreb, which traditionally has a third dialect—Kajkavian. The modern dialect in Split is a mix of Čakavian and Štokavian. Although this dialect mixing is not directly relevant to the data presented here, Čakavian and other dialect features that appeared during data collection are occasionally noted in endnotes.

Data was collected from several residents of the neighbourhood known as Radunica. Radunica is an old neighbourhood in the centre of Split. Once known as an area of “Fetivi Splici” (real Split residents), the neighbourhood is now primarily occupied by people who are described and who describe themselves as Vlaji—immigrants from the mountainous villages behind Split, or children, grandchildren, etc., of immigrants from those regions. I interviewed 11 native Croatian speakers ranging in age from 22 to 65, and in education from 4 formal years of schooling to 2 years of college (14 total years of schooling). The younger generation (below age 40) had all attended at least some college, while

no informant over 40 had completed more than 8 years of formal schooling. Informants of the older generation were themselves immigrants to the city (from the mainland Croatian cities of Imotski, Kljaci, Katuni and Omiš), while most of the younger generation had been born and raised in the city of Split. The 11 informants represent four families, and the families are neighbours and close friends, living on the same street. Informants were contacted through a friend who is a member of the community and an informant in this study.

Questionnaires were used in which informants were asked to fill in lexical frames with all possible forms for a given word which they themselves use. Each case/number combination which contrasts the o-stem and a-stem paradigms,¹⁵ excluding the vocative, was included in the questionnaire. The questionnaire included both masculine and feminine a-stem forms, both of which were controlled for frequency. The long version of the questionnaire included 30 lexemes (10 frequent masculine a-stems, 10 infrequent masculine a-stems, 5 frequent feminine a-stems and 5 infrequent feminine a-stems). A short version was half this length. Each form was elicited twice (although that does not necessarily mean that each informant provided each form twice, for independent and uncontrollable reasons), and forms were checked for accuracy as much as possible. If the informant did not know a word or if he/she claimed that he/she never uses a particular word, it was excluded from the data set. Some informants were additionally asked in follow-up sessions to judge sentences which varied according to masculine vs. feminine agreement.

Frequency of masculine a-stem nouns in Croatian was determined in the following way. A dictionary search of Benson (1994) was made to gather masculine a-stem nouns. All nouns ending in *-a* in the nominative singular and marked in the dictionary as being grammatically either masculine or masculine and feminine were considered “masculine a-stem nouns”. This search yielded 336 nouns. Archaic forms were included. A random sample of 160 nouns was gathered as a control group. Lexeme frequencies of both the random sample and the masculine a-stems were taken from an online corpus.¹⁶ For the masculine a-stems, mean frequency = 5.1/million, median = 0.25/million. For the random sample, mean frequency = 18.2/million; median = 0.22/million. Although the mean for the random sample group is more than triple that of the masculine a-stems, this is the result of three extremely high frequency nouns in the random sample group (*čovjek* ‘man’, *kuća* ‘house’, and *situacija* ‘situation’) which skew the average. For example, the Hrvatski Nacionalni Korpus lists *čovjek* ‘man’ as the 177th most frequent word of the Croatian language. If these three super-high frequency nouns are removed, the remaining 157 random sample nouns have a mean frequency = 5.53/million and median frequency = 0.22/million, which are quite close to the masculine a-stem numbers. Thus, it is assumed that the masculine a-stems do not contain any words of super-high frequency, but that generally a frequent masculine a-stem is a frequent word overall, while an infrequent masculine a-stem is an infrequent word overall.

Token frequencies were determined using the same corpus. Homophonous case/number combinations¹⁷ were counted by hand. Dative and locative were combined since they are essentially homophonous in all paradigms and have the same agreement patterns. In other words, the distinction between the dative and locative is a historical one, and is not reflected in the modern language.

4.2. Data

Of the 11 informants, only the oldest 4 allowed variation among the masculine a-stems. The fact that not all speakers allowed variation is significant, and in fact research previously conducted using young, well-educated informants from Bosnia and Serbia who were living in Columbus, OH also showed no variation. It is thus unclear whether the influential factor in the following data is age (representing a variation pattern which is being lost), or education. It is possible that the younger speakers do not allow variation in the declension of masculine a-stems because they have been more influenced by prescriptivist norms, and in fact there was some evidence to this effect during a few interviews. Some younger speakers acknowledged that their parents or other people that they know use the non-standard declension, but admitted that to them it sounded uneducated, even ignorant. Unfortunately, however, the confluence of education and age among my informants did not allow the factors of age and education to be teased apart to a satisfactory degree. Regardless, the nature of the variation is interesting. The following data is intended to *only* describe the declension used by *those speakers who allow variation*, and is not a description of the entire Split community, or even a representative sample of the community.

The informants who allowed an innovative o-stem declension for the masculine a-stem nouns (and in some cases even selected the innovative form as primary) did not allow variation on all nouns. Typical examples are given below. In examples (14)–(17) only the prescriptive a-stem declension was produced/allowed, despite the fact that the nouns can only refer to males. In general, where there was a potential for ambiguity between male and female natural gender, masculine agreement was used to disambiguate.

(14)¹⁸ Star-e kamiondžij-e/*kamiondžij-i, koj-e su jučer
 old.FEM truckdrivers.A-STEM/*O-STEM, who AUX.3.PL yesterday
 radi-le, spavaju danas.
 worked.FEM, sleep today.
 ‘The old truck drivers, who worked yesterday, are sleeping today.’

(15) Njen-e lijen-e slug-e/*sluz-i/*slug-i¹⁹ nisu
 her.FEM lazy.FEM servants.A-STEM/*O-STEM/*O-STEM AUX.NEG.3.PL
 htje-le raditi.
 wanted.FEM work
 ‘Her lazy servants didn’t want to work.’

- (16) Moj gazd-a/*gazd/*gazad, ko-mu posuđujem²⁰
 my.MASC landlord.A-STEM/*O-STEM/*O-STEM, to whom.MASC owe.1.SG
 novac, nije simpatičan čovjek.
 money not.is kind man
 ‘My landlord, to whom I owe money, is not a kind man.’
- (17) Dao sam pismo svoj-im koleg-ama²¹/*koleg-ima.
 gave AUX.1.SG letter self’s colleagues.A-STEM/*O-STEM
 ‘I gave the letter to my colleagues.’

By contrast, in (18) to (22), o-stem and a-stem declensions were equally palatable. Each noun in question must be interpreted as masculine. For example, the lumberjack in (18) could not be interpreted as being female because the participle agreement target *obarao* ‘cut’ has masculine form. The same holds true even more obviously for (19) and (22) because bridegrooms and bachelors semantically cannot be female.

- (18) Posmatrao sam jak-og drvosječ-a/drvosječ-u dok je
 watched AUX.1.SG strong.MASC lumberjack.O-STEM/A-STEM while AUX.3.SG
 obara-o drvo.
 cut.MASC tree
 ‘I watched the strong lumberjack while he cut down the tree.’
- (19) Obred vjenčanja pokazuje ljubav lijep-og
 wedding ceremony shows love handsome.MASC
 mladoženj-a/mladoženj-e prema svojoj ženi.
 bridegroom.O-STEM/A-STEM towards self’s wife
 ‘The wedding ceremony demonstrates the love of the handsome bridegroom towards his wife.’
- (20) Poslat ću pismo lovokradic-ima²²/lovokradic-ama.
 send FUT letter lumber poachers.O-STEM/A-STEM
 ‘I will send a letter to the lumber poachers.’
- (21) Ima vlast nad sveznalic-ima/sveznalic-ama.
 has power over know-it-alls.O-STEM/A-STEM
 ‘He has power over the know-it-alls.’
- (22) Govorimo o neženj-u/neženj-i.
 talk.1.PL about bachelor.O-STEM/A-STEM
 ‘We are talking about the bachelor.’

Finally, as shown in (23) through (26), speakers occasionally provided/allowed only o-stem declension for some forms of a word, suggesting a mixed paradigm. These cases are perhaps the most interesting among the data because the informants did not provide the Standard Croatian forms and when questioned, consistently judged them as ungrammatical.

- (23) Hoće naći lovokradivc-a/*lovokradic-u, da...
 want.3PL find lumber poacher.**O-STEM/*A-STEM** so that...
 ‘They want to find the poacher so that...’
- (24) Tražim paralažc-a/*paralaž-u.
 search for.1.SG liar.**O-STEM/*A-STEM**
 ‘I am looking for the consummate liar.’
- (25) Ne moraš vjerovati t-om paralaž-u/*paralaž-i.
 not should.2.SG believe that.MASC liar.**O-STEM/*A-STEM**
 ‘You shouldn’t believe that consummate liar.’⁶
- (26)²³ U zoološkom vrtu moja kćerka voli gledati goril-a/*goril-u.
 At zoo my daughter like watch gorilla.**O-STEM/*A-STEM**
 ‘At the zoo my daughter likes to watch the gorilla.’

Note that in (23) and (24), the reanalysis involves not only a jump to the o-stem paradigm, but also a reformation of the stem—nominative singular *lovokradic-a* becomes *lovokradivac-Ø*, and *paralaž-a* becomes *paralažac-Ø*. Thus, while there is a paradigm change, this is not the only strategy of speakers when confronted with gender conflict. In Croatian, it is quite uncommon to have an o-stem noun with the stem ending /ž/ or /v/. By actively reforming the stem through adding /ac/, speakers are normalising the word form so that it appears to be a more typical member of the o-stem group. All four informants used this strategy. Comparing these forms with (25) and (20), however, we can see that the traditional stem (with both prescriptive and innovative suffixes) is still possible for these speakers. In short, the pattern of variation is lexically specific, and even more than that, it appears to be determined item-by-item, with declension varying even by case/number combination. Could frequency in Croatian be playing a role in this item-by-item pattern?

4.3. Analysis

Figure 1 shows that there is clearly a relationship between frequency and declension type.²⁴ Words of high frequency uniformly follow the standard a-stem declension. Tokens with intermediate lexeme frequencies are overwhelmingly

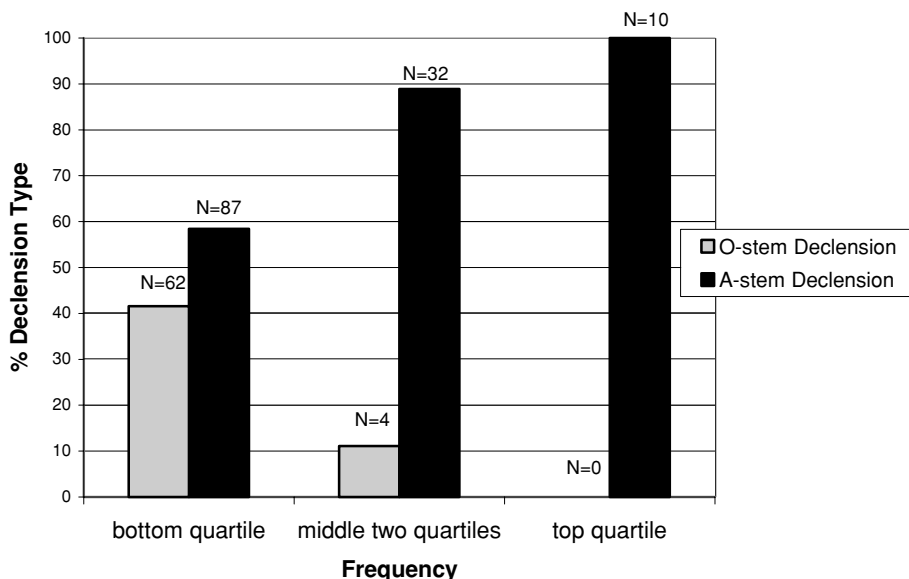


Figure 1. Croatian data according to frequency

declined according to the standard language pattern, but some words do decline according to the innovative o-stem pattern. Most important, however, is that in the lowest frequency range, there are a significant number of both types of declension for the masculine a-stem nouns. This suggests that in the lowest frequency range, speakers were choosing (almost) randomly between the two patterns. Furthermore, as can be seen from Figure 2, *the same informant* often allowed *both* the innovative o-stem declension and the standard a-stem declension for *the same word*. Figure 2 represents the 66 tokens from Figure 1 that allowed innovative declension.

Among the lowest frequency tokens, some forms take only the o-stem form, while others showed both declension types. All of the varying forms in the middle frequency range allow both the innovative o-stem and traditional a-stem declensions. Where both declensions were allowed, informants varied (from word to word, day to day and speaker to speaker) as to which they considered primary. This pattern further confirms that declension type for masculine a-stems correlates with frequency, and that especially among the lowest frequency tokens, declension type is best described as lexically specific, even random.

Returning to our two predictions regarding frequency, the Croatian data seems to support Hypothesis 2, and not Hypothesis 1. These are repeated below.

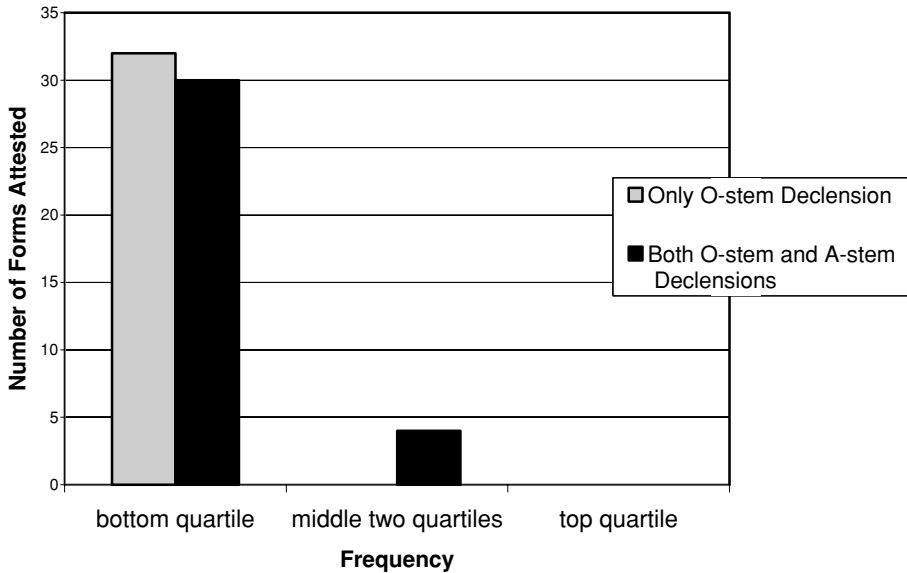


Figure 2. Forms showing o-stem declension

Hypothesis 1: The cognitive approach

When analogical language change occurs, it will begin with all infrequent forms, progressing to the frequent ones.

Hypothesis 2: The social salience approach

Where a social marker exists, salient forms will be inhibited from showing the stigmatised variant. Low frequency is thus a necessary but not sufficient condition for the appearance of a stigmatised morphological form.

If frequency straightforwardly determines which tokens are affected by analogy, we would expect a clear inverse correlation between frequency and o-stem declension. However, what we get overall, and for individual speakers, is a more chaotic pattern. This chaos is the specific problem for Hypothesis 1. As we have seen, each individual speaker produced multiple forms for the same paradigmatic cell of the same word. It seems that among low-frequency tokens, it simply does not matter which form is produced. This result is predicted by a social perspective. This behaviour is not expected in a solely cognitive view, which predicts that production will directly result from lexical structure.

All other things assumed to be equal, it appears that frequency is serving as an upper boundary constraint on change, rather than as a determiner of the path

of the change. I speculate that the enforcement of speech community norms and salience of variation among high-frequency words (as a result of the structure of the lexicon) lead to the upper boundary. In other words, the real constraint may be *social*, with frequency correlating only as a by-product.

However, this conclusion rests on two prerequisites. The first is that the variation among the masculine a-stems is, in fact, stigmatised, and thus a social marker. After all, in the social/communicative approach the prediction changes depending upon social forces. Hypothesis 2 holds only if the a-stem form is considered overtly prestigious, and the relevant community views the o-stem form as stigmatised. The second prerequisite is that all things are equal—that factors other than frequency are not affecting the masculine a-stem declensional pattern. These two issues are addressed in the following sections.

4.4. Anecdotal evidence for social markers

There is some anecdotal evidence that masculine a-stems have, in fact, become social markers. For example, during the course of this fieldwork it was common for my less educated informants (the same speakers who provided the forms above) to try to refer me to more educated people to find out how the words are ‘supposed’ to be. This indicates that they do not view their own speech as prestigious, but they do value overt prestige. Second, when discussing the above data with educated speakers of the language, many had strong negative reactions to the o-stem forms. Many fieldworkers have had the experience that when a prestigious group is unusually aggressive in denying the acceptability of a form, it is because that form does exist—in the speech of a stigmatised group. Third, as noted above, some educated speakers conceded that the innovative o-stem forms are possible, but associated them with uneducated speech. Finally, the less educated speakers readily identified themselves as Vlaji. Since this term is widely acknowledged to be derogatory (Magner 1978), it suggests that the speakers do, in fact, view themselves as less (overtly) prestigious than another group. Putting all of these little pieces of evidence together, the signs seem to be pointing towards the status of the masculine a-stems as social markers, and of the o-stem declension as the stigmatised realisation.

4.5. Are all things equal?

To judge Hypothesis 1 versus Hypothesis 2, we must also assume that all factors other than frequency are equal, to ensure that the frequency data is not skewed by external factors. There are two good candidates in this regard—animacy and agreement.

4.5.1. Animacy

Are some cases being affected before others, regardless of frequency? We might expect that animacy would cause o-stem forms to appear in the genitive and accusative singulars more than in other paradigmatic cells.

In Serbian and Croatian animacy is semantically motivated for masculine nouns only, and is morphologically reflected in the o-stem paradigm and in agreement. It is marked as genitive—accusative syncretism in the singular, e.g. *junak* ‘hero.NOM.SG’, *junaka* ‘hero.ACC.SG’, *junaka* ‘hero.GEN.SG’. Non-animate masculine o-stem nouns have nominative—accusative syncretism in the singular, e.g. *grad* ‘city.NOM.SG’, *grad* ‘city.ACC.SG’, *grada* ‘city.ACC.SG’.

Masculine a-stems have not traditionally marked animacy morphologically, but agreement targets show animacy where appropriate (also marked as genitive—accusative synchrony). Furthermore, Belić (1924b:26) argues that the use of masculine agreement in the singular for masculine a-stem nouns (see section 3 above) resulted from a desire to mark animacy. He claims that masculine agreement was first used in the accusative singular and later spread to the singular generally. Masculine agreement did not spread as effectively to the plural because animacy is not marked in the plural.

Since the desire to mark animacy seems to have had a great influence on the use of masculine agreement, we might wonder whether there is a similar effect on declension in Croatian. Specifically, the question is whether among my informants in Split, Croatia, the innovative o-stem endings appear more in the accusative and genitive singulars than in other case number combinations. If animacy is a contributing factor in the observed ‘declension hopping’, we would expect o-stem declension to be more prevalent in these paradigm cells than in the others. As can be seen in (27), the accusative singular and genitive singular do *not* in fact have proportionally more innovative forms.

In this table, the total number of occurrences of the o-stem form (either as one of two possibilities, or as the only possibility) is divided by the total number of elicited masculine a-stems for that case/number combination. The percentage thus represents the percent occurrence of the innovative o-stem form.

(27) Innovation by case/number combination

	Singular	Plural
Nominative	11/36 (31%)	9/25 (36%)
Accusative	12/36 (33%)	*
Genitive	10/29 (34%)	*
Dative/Locative	11/30 (37%)	8/21 (38%)
Instrumental	*	5/18 ²⁵ (28%)

* = o-stem and a-stem paradigms are homophonous.

Note that in the accusative singular 33% of the forms matched the o-stem paradigm, and in the genitive singular this number was 34%. In other words, these cells behaved very similarly to the other case number combinations. This lack of significance suggests that animacy is *not* playing a role in the Croatian declensional change and that only frequency has an obvious effect.

4.5.2. Agreement

Returning to the idea that morphological variation among these nouns represents declensional gender resolution, we might also wonder if masculine agreement triggers o-stem declension. This hypothesis still needs to be researched, but there is at least some preliminary suggestion that agreement may be indirectly, but not directly, influential.

Janda and Varela-García (1991) have noted that agreement can be subject to a ‘recency’ effect. They note examples of the following type from Spanish, in which nouns such as *el ave* ‘the bird’ (i.e. masculine nouns with typically feminine morphology) can take masculine agreement to the left of the noun, but feminine agreement to the right.

- (28) el otr-o ave está enferm-a
 the.MASC other.MASC bird is sick.FEM
 ‘The other bird is sick’.

If morphology can have this sort of immediate effect on agreement, we might wonder if agreement in Croatian could affect morphology in a similar way. The examples in (29a) and (29b) show that, when asked to make syntactic judgments, informants did not report sensing a grammaticality difference between the sentence with an immediately preceding attributive, and one without.

- (29) a. Vjenčanje pokazuje ljubav ∅ mladoženj-a prema svojoj
 wedding demonstrates love ∅ bridegroom.**O-STEM** towards self’s
 ženi.
 wife.
 ‘The wedding ceremony demonstrates the love of the bridegroom
 towards his wife.’
- b. Vjenčanje pokazuje ljubav *lijep-og* mladoženj-a
 wedding demonstrates love handsome.MASC bridegroom.**O-STEM**
 prema svojoj ženi.
 towards self’s wife.
 ‘The wedding ceremony demonstrate the love of the handsome
 bridegroom towards his wife.’

However, there seems to be a more general effect for agreement. Informants who accepted declensional variation were (subjectively) more likely to accept non-standard masculine agreement in the plural than were informants who did not allow any declensional variation. Examples are given in (30) and (31).

(30) Typical judgments of informants that did *not* allow morphological variation:

a. Njen-e lijn-e slug-e nisu htje-le raditi.
her.FEM lazy.FEM servants.A-STEM NEG.AUX wanted.FEM work.
'Her lazy servants didn't want to work.'

b. *Njen-i lijn-i slug-e nisu htje-li raditi.
her.MASC lazy.MASC servants.A-STEM NEG.AUX wanted.MASC work.
'Her lazy servants didn't want to work.'

(31) Typical judgments of informants that did allow morphological variation:

a. Njen-e lijn-e slug-e nisu htje-le raditi.
her.FEM lazy.FEM servants.A-STEM NEG.AUX wanted.FEM work.
'Her lazy servants didn't want to work.'

b. Njen-i lijn-i slug-e nisu htje-li raditi.
her.MASC lazy.MASC servants.A-STEM NEG.AUX wanted.MASC work.
'Her lazy servants didn't want to work.'

While this difference of judgement still needs investigating, it suggests that the extension of masculine agreement may be linked to the use of o-stem declension for these a-stems. If this proves to be the case, a more detailed study of the interacting effects of frequency and context/social patterns would be needed.

Still, this does not contradict the frequency effects seen above. In fact, if a social salience perspective on variation is taken, the juxtaposition of conflicting agreement and declension may be salient just as variation among high-frequency nouns is salient, suggesting agreement and frequency may not be competing influences, but rather complementary ones.

5. CONCLUSIONS

Considering that analogy relies on connections drawn by the speaker between two lexical items, and thus its domain is the lexicon, we should not be surprised that frequency affects morphological change. Many theories of the lexicon and lexical access have incorporated frequency effects and historical linguistics has recognised that language change, and especially analogy, is strongly affected by social and other extralinguistic factors. In short, the experience of the speaker and the generalisations that that speaker makes cannot be separated. And while

frequency has often been incorporated into theories of language change, only the broad implications of frequency have, in general, been investigated.

This paper suggests that one view of the lexicon could result in two different predictions regarding frequency, depending upon whether a cognitive or a social perspective is taken. The first prediction is rooted in the structure of the lexicon itself as a guiding factor in analogical change. The second considers the fundamental constraining factor to be social norms rather than the structure of the lexicon in any direct way. In this paper I have presented evidence primarily from Croatian masculine a-stems which shows that declensional gender resolution is possible for these nouns. It is concluded that the item-by-item nature of this variation tips the scale in favour of viewing the role of the social perspective.

NOTES

* I owe thanks to the many people who have commented on parts of this work, but most of all Brian Joseph, Dan Collins and Keith Johnson. I also thank Bojan Belić, Robin Dodsworth, Charles Gribble, Jiří Hana, Christina Kramer, Grant McGuire, Jeff Mielke, Kate Minnis, Bill Raymond, Joseph Schallert, Kat Tancock, three anonymous reviewers and the members of Changelings (OSU), the Slavic Linguistics Forum (OSU), and Beth Hume O’Haire’s Frequency-based Models seminar (OSU). All errors are mine. This work was supported in part by a G. Micheal Riley Scholarship, a Graduate Student International M.A. Thesis/Dissertation Research Grant, the Kenneth E. Naylor Professorship and the Department of Linguistics, all at Ohio State University.

¹ For example, the Eighth and Ninth tendencies: “If there is a difference between the inflection of a geographic noun and a common noun, which otherwise are similar, the local cases generally present an archaic character, while in the non-local cases innovations are more common” (cited in Hock 1991:232–233).

² Wang’s work is primarily concerned with sound change, but the spread of sound change is considered to be fundamentally analogical.

³ Janda and Joseph are also primarily concerned with changes in sound rather than with morphological analogy, but they include analogy among the main mechanisms of spread.

⁴ Grammaticalisation may involve semantic bleaching, but need not according to Hopper and Traugott. “... [E]arly stages of grammaticalization do not show bleaching. Rather there is a balance between loss of older, typically more concrete meanings, and development of newer, more abstract ones that at a minimum cancel out the loss” (1993:96). Where semantic loss is found in later stages, Hopper and Traugott seem to assume that it runs in parallel with loss of form, rather than one being causative of the other.

⁵ Taking a broad definition of ‘lexical’. The use of this term is not meant to imply that grammatical endings and other morphemes cannot participate in analogy.

⁶ This paper is not concerned with the origin of markers. It might be suggested that a particular example of variable language use develops into a marker because it is salient, but this creates a circularity in the concept of salience. There is not space to explore this problem here.

⁷ The term ‘Croatian’ is used here because all informants for this study were from the city of Split. It is not meant to imply that similar changes could not happen in Serbian. When describing features that are common to both variants, the term ‘Serbian and Croatian’ is used.

⁸ Formal syntactic analyses of mixed gender agreement are presented in Kathol (1999) and Weschler and Zlatić (2000).

⁹ Slavists will recognise that this terminology is not entirely standard. 'o-stem' and 'a-stem' are terms which are often reserved for discussion of Proto-Slavic and Old Church Slavonic because of later changes to the paradigms in many languages. However, this paper is concerned with natural sex, agreement and declension class, all of which have something to do with gender. The terms 'o-stem' and 'a-stem' are used throughout to clarify the important difference between declension class assignment and syntactic gender agreement; I reserve the terms 'masculine' and 'feminine' for agreement. This follows the usage of Comrie and Corbett (1993).

¹⁰ Nouns of the second type, traditionally known as either i-stems, class III nouns or feminine II nouns, are much fewer in number, and not of concern for this paper.

¹¹ Common gender a-stem nouns pattern like masculine a-stems when referring to males, but have entirely feminine agreement when referring to females. This is to be expected when the influence of male reference is removed.

¹² Čakavian is a dialect of the Croatian language spoken along the Dalmatian coast. It differs significantly from the standard language, and has a strong literary tradition dating from the 16th century.

¹³ According to the Second Palatalization of Velars /k/ alternates with /č/, with the latter appearing before /i/, although some exceptions exist.

¹⁴ According to the First Palatalization of Velars /k/ alternates with /ć/, with the latter appearing before /e/ in the vocative singular, but not in the accusative plural.

¹⁵ It can be seen from the paradigms in (11) and (12) that the Instrumental singular, genitive plural and accusative plural are indeterminate since the o-stem and a-stem forms are homophonous for these case/number combinations.

¹⁶ *Hrvatski Nacionalni Korpus/Croatian National Corpus*, a 9.1-million-word, balanced textual corpus. <www.hnk.ffzq.hr>

¹⁷ For a-stem declension, the following combinations are homophonous: (a) genitive singular and nominative plural, (b) instrumental plural and dative/locative plurals and (c) nominative singular and genitive plural in writing. In speech the nominative singular and genitive plural are distinguishable because the genitive plural has post-tonic lengthening, but this distinction is not recorded in the written language.

¹⁸ An anonymous reviewer rightly asked whether agreement was controlled for in examples (14) and (15), i.e. whether the feminine agreement might have influenced the judgment that o-stem declension was ungrammatical. Although most of the plural examples contained feminine agreement (as is prescriptive), some contained masculine agreement. Furthermore, informants were asked in follow-up sessions for judgements regarding masculine agreement in these examples. The oldest informants allowed masculine agreement in examples such as (14) and (15) but still reported that o-stem declension was ungrammatical. The issue of the immediate influence of agreement upon declension is addressed in section 4.5.2.

¹⁹ It was observed that these informants often use velars where palatals would be expected according to the Second Palatalization of Velars, e.g. in front of nominative plural /-i/. For this reason, both the palatal and velar variants were tested where relevant.

²⁰ Words with final [m] in Standard Croatian were alternately pronounced by my informants as either [m] or [n]. $m \rightarrow n / _ \#$ was noted by Magner (1978) as a dialectal feature of Split, and is widely observable in dialects along the Dalmatian coast of Croatia.

²¹ The alternate feminine dative/locative/instrumental plural endings *-an* and *-aman* (e.g. *žen(am)an* 'women.DAT/LOC/INST.PL.') were attested (for this and other words) alongside the

Standard Croatian *-ama* (e.g. *ženama* ‘women.DAT/LOC/INST.PL’). Jutronić (2001) and Magner (1978) note *-an* and *-aman* as dialect features of Split. The non-homophonous instrumental plural *-ami* (e.g. *ženami* ‘women.INST.PL’) and locative plural *-ah* (e.g. *ženah* ‘women.LOC.PL’), which were noted by Finka (1971) for Čakavian generally are now considered archaic, and were not present in my data.

²² Instrumental plural *-iman* was also attested for this and other words, alongside the standard language *-ima*. This ending has been noted by Jutronić (2001) and Magner (1978) as a feature of the Split dialect.

²³ Although *gorila* ‘gorilla’ is semantically distinct from other masculine a-stems in not denoting a human, it is listed as a masculine a-stem in Benson (1994), and the variation according to the natural sex of the gorilla suggests that speakers view this noun as being of the same type as other masculine a-stems.

²⁴ It is unlikely that the usage of the speakers described here is exactly the same as listed in the Hrvatski Nacionalni Korpus. Nonetheless, the corpus is the best available estimate of frequency.

²⁵ This cell is not significant ($p = .40$).

REFERENCES

- Alegre, Maria and Peter Gordon (1998). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language* 40, 41–61.
- Bardovi-Harlig (1987). Markedness and salience in second-language acquisition. *Language Learning* 37(3), 385–407.
- Belić, A (1924a). Prilozi istoriji slovenskih jezika. In *Južnoslovenski filolog* 4. Beograd: Srpska Akademija Nauke, 1–28.
- Belić, A (1924b). Napomene o jednoj sintaksičko-morfološkoj osobini srpskohrvatskog jezika. *Južnoslovenski filolog* 4.
- Benson, Morton (1994). *Serbo-Croatian-English dictionary*, 3rd edn. Cambridge: Cambridge University Press.
- Bertram, Raymond, Robert Schreuder and R. Harald Baayen (2000). The Balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy and productivity. *Journal of Experimental Psychology: Learning, Memory and Cognition* 26(2): 1–23.
- Browne, Wayles (1993). Serbo-Croat. In Bernard Comrie and Greville Corbett (eds.), *The Slavonic Languages*. New York: Routledge, 306–387.
- Butterworth, Brian (1983). Lexical representation. In Brian Butterworth (ed.), *Language Production, Vol. 2: Development, Writing and Other Language Processes*. London: Academic Press, 257–294.
- Bybee, Joan (1985). *Morphology: A Study of the Relation Between Meaning and Form*. Philadelphia: John Benjamins.
- Bybee, Joan (1988). Morphology as lexical organization. In Hammond, Michael and Michael Noonan (eds.), *Theoretical Morphology: Approaches in Modern Linguistics*. San Diego, CA: Academic Press, 119–141.
- Bybee, Joan and Paul Hopper (2001). *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins.

- Caramazza, Alfonso, Alessandro Laudanna and Christina Romani (1988). Lexical access and inflectional morphology. *Cognition* 28(3), 297–332.
- Cheshire, Jenny (1996). Syntactic variation and the concept of prominence. Juhani Klemola, M. Kytö and M. Rissanen (eds.), *Speech Past and Present: Studies in English Dialectology in Memory of Ossi Ihalainen*. Frankfurt: Peter Lang, 1–17.
- Comrie, Bernard and Greville Corbett (eds.) (1993). *The Slavonic Languages*. New York: Routledge.
- Corbett, Greville (1991). *Gender*. Cambridge: Cambridge University Press.
- Finka, Božidar (1971). Čakavsko narečje. *Čakavska rič* 1(1), pp. 11–71.
- Glavan, Vjenceslav (1928–29). Kongruencija u jeziku starih čakavskih pisaca. *Južnoslovenski filolog* 7, 111–159.
- Hay, Jennifer (2003). *Causes and Consequences of Word Structure*. New York: Routledge.
- Heim, Michael (1982). *Contemporary Czech*. Columbus, OH: Slavica.
- Hock, Hans Henrich (1991). *Principles of Historical Linguistics*. New York: Mouton de Gruyter.
- Hooper, Joan (1976). Word frequency in lexical diffusion and the source of morphophonological change. In William Christie (ed.), *Current Progress in Historical Linguistics: Proceedings of the Second International Conference on Historical Linguistics*. New York: North-Holland, 95–106.
- Hopper, Paul and Elizabeth Traugott (1993). *Grammaticalization*. New York: Cambridge University Press.
- Ivić, Pavle (1994). *Srpskohrvatski dijalekti: njihova struktura i razvoj*. Sremski Karlovci: Knjižnica Zorana Stojanovića.
- Janda, Richard and Brian Joseph (2003). Reconsidering the canons of sound change: Towards a Big Bang theory. In Barry Blake and Kate Burridge (eds.), *Historical Linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics*, Melbourne, 13–17 August 2001. Amsterdam: John Benjamins, 205–219.
- Janda, Richard and Fabiola Varela-García (1991). On lateral hermaphroditism and other variation in Spanish ‘feminine’ *el*. *Papers from the 27th Regional Meeting of the Chicago Linguistic Society* 27(1), 276–290.
- Jutrović, Dunja (2001). Morphological changes in the urban vernacular of the city of Split. *International Journal of the Sociology of Language* 147, 65–78.
- Kathol, Andreas (1999). Agreement and the syntax-morphology interface in HPSG. In Robert Levine and Georgia Green (eds.), *Studies in Head-Driven Phrase Structure Grammar*. Cambridge: Cambridge University Press, 223–274.
- Kerswill, Paul (1985). A sociophonetic study of connected speech processes in Cambridge. English: an outline and some results. *Cambridge Papers in Phonetics and Experimental Linguistics* 4, 17–39.
- Kerswill, Paul and Ann Williams (2002). ‘Salience’ as an explanatory factor in language change: Evidence from dialect leveling in urban England. In Mari C. Jones and Edith Esch (eds.), *Language Change: The Interplay of Internal, External and Extra-Linguistic Factors*. New York: Mouton de Gruyter, 81–109.
- Labov, William (1994). *Principles of Linguistic Change: Social Factors*. Cambridge, MA: Blackwell.
- Malkiel, Yakov (1990). *Diachronic Problems in Phonosymbolism*. Amsterdam: John Benjamins.

- Malkiel, Yakov (1992). *Diachronic Studies in Lexicology, Affixation, Phonology*. Amsterdam: John Benjamins.
- Magner, Thomas (1978). Diglossia in split. *Folia Slavica* 1(3), 400–443.
- Marković, Svetozar (1954). O kolebljivosti slaganja u rodu kod imenica čiji se prirodni i gramatički rod ne slažu (i o rodu ovih imenica). *Pitanja književnosti i jezika* 1, 87–110.
- Mufwene, Salikoko (1991). Pidgins, creoles, typology, and markedness. In Francis Byrne and Thom Huebner (eds.), *Development and Structure of Creole Languages: Essays in Honor of Derek Bickerton*. Amsterdam: John Benjamins, 123–143.
- Ogura, Mieko and William S.-Y. Wang (1996). Snowball effects in lexical diffusion: The development of -s in the third person singular present indicative in English. In Derek Britton (ed.), *English Historical Linguistics 1994: Papers from the 8th International Conference on English Historical Linguistics*. Philadelphia, PA: John Benjamins, 119–141.
- Phillips, Betty (1984). Word frequency and the actuation of sound change. *Language* 60, 320–342.
- Phillips, Betty (2001). Lexical diffusion, lexical frequency, and lexical analysis. In Joan Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*. Philadelphia: John Benjamins, 123–136.
- Prokosch, Eduard (1939). *A Comparative Germanic Grammar*. Philadelphia: Linguistic Society of America.
- Schreuder, Robert and R. Harald Baayen (1995). Modeling morphological processing. In Laurie Beth Feldman (ed.), *Morphological Aspects of Language Processing*. Hillsdale, NJ: Lawrence Erlbaum, 131–154.
- Stemberger, Joseph Paul and Brian MacWhinney (1988). Are inflected forms stored in the lexicon? In Michael Hammond and Michael Noonan (eds.), *Theoretical Morphology: Approaches in Modern Linguistics*. San Diego: Academic Press, 101–116.
- Sturtevant, Edgar (1917). *Linguistic Change: An Introduction to the Historical Study of Language*. New York: G.E. Stechert & Co.
- Taft, Marcus and Kenneth Forster (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior* 14(6), 638–647.
- Timberlake, Alan (1993). Russian. In Bernard Comrie and Greville Corbett (eds.), *The Slavonic Languages*. New York: Routledge, 827–886.
- Trudgill, Peter (1986). *Dialects in Contact*. Oxford: Blackwell.
- Wang, William S.-Y. (1969). Competing changes as cause of residue. *Language* 45, 9–25.
- Weschler, Stephen and Larisa Zlatić (2000). A theory of agreement and its application to Serbo-Croatian. *Language* 76(4), 799–832.
- Yaeger-Dror, Malcah (2003). Linguistic analysis of dialect ‘correction’ and its interaction with cognitive salience. *Language Variation and Change* 5(2), 189–224.

The Ohio State University

asims@ling.ohio-state.edu

Referrals and morphemes in Sora verb inflection

GREGORY T. STUMP

A persistent challenge for grammatical theory is that of accounting for the varied mismatches which may exist between a word's syntacticosemantic content and its inflectional form. In this paper, I analyse a body of data from the Sora language which proves particularly challenging in this respect. In the inflection of Sora verbs, a high degree of morphological economy is achieved by employing the same morphology to express very different kinds of content. The Sora facts exclude a simple, morpheme-based conception of inflectional morphology, and instead necessitate a conception of morphology in which a lexeme's inflectional realisation is defined through the interface of the lexicon with a richly articulated system of rules.

In section 1, I present a description of the Sora system of verb inflection and draw attention to three inflectional suffixes that have more than one function (a fact which excludes a morpheme-based analysis of this system): (i) a suffix which functions in some paradigms as an expression of voice, but in other paradigms as a mere mark of inflection-class membership; (ii) a suffix which functions in some paradigms as an expression of first-person exclusive subject agreement, in other paradigms as an expression of first-person plural inclusive object agreement, and in still other paradigms as an expression of speaker-directed motion; and (iii) a suffix which functions in some paradigms as an expression of third-person plural subject agreement but in other paradigms as an expression of third-person plural object agreement. In sections 2 and 3, I argue that the mismatches between form and content engendered by these suffixes can be resolved within the framework of an inferential-realisation theory of morphology by postulating (a) rules of referral that relate contrasting uses of the same suffix, and (b) rules of semantic interpretation that are sensitive to a word's paradigmatic context. In section 4, I develop these ideas in a formal analysis of the Sora facts; in this analysis, the versatility of the suffixes in (i)–(iii) is accounted for through the interaction of rules of types (a) and (b). I summarise my conclusions in section 5.¹

1. AN OVERVIEW OF SORA VERB INFLECTION

Sora, an Austro-Asiatic language belonging to the Munda subgroup, is spoken by approximately 288,000 people (as of 1997), principally in the Indian states of Orissa, Andhra Pradesh and Madhya Pradesh. Verbs in Sora fall into several distinct inflection classes (Biligiri 1965, Ramamurti 1931:25–45, 118–146); the three principal classes are represented in Table 1 by the verbs *DE* 'get up', *BƏD* 'make' and *DE* 'become'. (Some additional members of these three classes are given in (1).) The *DE*-class and the *BƏD*-class differ quite straightforwardly: in

all person/number combinations, members of the DE-class exhibit an *-n* suffix which members of the BƏD-class lack; and in the realization of all person/number combinations except the first-person plural inclusive, members of the BƏD-class involve an *-ε* suffix which members of the DE-class lack. The simplicity of this contrast between the DE- and BƏD-classes is obscured by two morphophonological modifications, namely (a) the prevocalic elision of suffix-final vowels, as in *bət-t-ε* ‘s/he makes’ (where the vowel in the non-past-tense suffix *-te* is elided before *-ε*) and *bət-t-ay* ‘I make’ (where elision causes the suffix sequence *-te -ε -ay* to be realised as *-t-ay*); and (b) the realisation of the tense suffixes *-te* and *-le* as *-tə* and *-lə* before consonants other than /n/, as in *bət-tə-be* ‘we (incl.) make’. In his analysis of Sora verb morphology, Biligiri (1965) labels the DE- and BƏD-classes as classes $+N^b$ and $-N^b$, respectively, where the $+N/-N$ specification reflects the presence or absence of the *-n* suffix; the significance of the superscript ‘b’ in Biligiri’s labels will be explained shortly.

Table 1. Partial inflectional paradigms of three Sora verbs

			DE ‘get up’	BƏD ‘make’	DE ‘become’	
Non-past Affirmative	Sg	1 st	de-te-n-ay	bət-t-ay	dε-t-iñ	
		2 nd	de-te-n	bət-t-ε	dε-t-əm	
		3 rd	de-te-n	bət-t-ε	dε-t-e	
	Pl	1 st incl	de-te-n-be	bət-tə-be	dε-t-ay	
		1 st excl	ə-de-te-n-ay	ə-bət-t-ay	dε-tə-lən	
		2 nd	ə-de-te-n	ə-bət-t-ε	dε-tə-ben	
Past Affirmative	Sg	3 rd	de-te-n-ji	bət-t-ε-ji	dε-tə-ji	
		1 st	de-le-n-ay	bəd-l-ay	dε-l-iñ	
		2 nd	de-le-n	bəd-l-ε	dε-l-əm	
	Pl	3 rd	de-le-n	bəd-l-ε	dε-l-e	
		1 st incl	de-le-n-be	bəd-lə-be	dε-l-ay	
		1 st excl	ə-de-le-n-ay	ə-bəd-l-ay	dε-lə-lən	
			2 nd	ə-de-le-n	ə-bəd-l-ε	dε-lə-ben
			3 rd	de-le-n-ji	bəd-l-ε-ji	dε-lə-ji
			Biligiri’s (1965) classification:			$+N^b$

- (1) a. $+N^b$: ÑA ‘walk’, BER ‘speak’, DER ‘believe’, DAKU ‘stay’, ...
 b. $-N^b$: GA ‘drink’, ÑI ‘buy’, GU ‘call’, JUM ‘eat’, Gʃ ‘see’, TEM ‘sell’, ...
 c. I: DA ‘feel’, MʔEŋ ‘live’, ASU ‘pain’, KƏJED ‘die’, RʔA ‘be in blossom’, ʔO ‘bear fruit’, MANA ‘be tasty’, BAJA ‘be mad’, GITTA ‘appear’ ...

The verb DE in Table 1 belongs to a third class, which Biligiri labels class I. In the inflection of transitive members of this class, agreement morphology encodes properties of the object rather than of the subject; in this respect, class

i differs from classes $+N^b$ and $-N^b$, whose members uniformly inflect for subject agreement. Although the verb's agent argument may be unexpressed in a clause headed by a class *i* verb, such a clause cannot be seen as passive, for if the verb's agent is expressed, it is as the verb's subject; the examples in (2) illustrate. (Nevertheless, I shall often use an English passive as the nearest approximation for the translation of a Sora impersonal form; I shall mark such translations with the '≈' symbol.)

- | | | | | | |
|--------|--------------------|----------------|----|---------------------|--------------------|
| (2) a. | amən | gij-t-iñ. | b. | ənin | gij-t-əm. |
| | you.SG | see-NONPAST-me | | he | see-NONPAST-YOU.SG |
| | 'you (sg.) see me' | | | 'he sees you (sg.)' | |
- (Biligiri 1965:240f)

Single-argument members of class *i* tend to have a non-agentive (patient or experiencer) argument; if such verbs are thought of as having an expletive (non-thematic) subject, then these, too, exhibit object rather than subject agreement.² Verbs belonging to class *i* are therefore "impersonal" in the sense that as a class, they fail to inflect for subject agreement. (As will be seen below, however, this last statement must be qualified.)

Although the three inflection classes represented in Table 1 are fundamentally distinct, some verbs inflect as members of more than one of these classes. Consider first the verb *kuŋ* 'shave' in Table 2. Biligiri assigns this verb to the inflection class $\pm N^b$, indicating that it inflects both as a class $+N^b$ verb (like *DE* in Table 1) and as a class $-N^b$ verb (like *BƏD* in Table 1). Significantly, the difference between class $+N^b$ morphology and class $-N^b$ morphology in the inflection of class $\pm N^b$ verbs expresses a voice contrast: thus, as a class $+N^b$ verb, *kuŋ* has the reflexive meaning 'shave oneself', while as a class $-N^b$ verb, it has the non-reflexive meaning 'shave (someone else)'. Some other verbs belonging to class $\pm N^b$ are given in (3).

- (3) $\pm N^b$: *SO* 'hide', *B#B* 'drown', *ABUMA* 'bathe', ...

These facts raise the question of whether verbs belonging to the $+N^b$ class should be seen as inherently reflexive (and those belonging to the $-N^b$ class, as inherently non-reflexive). While one might try to make a case for this assumption in individual instances, it is clear that it cannot, in general, be maintained. For example, according to Ramamurti (1931:102ff), the verbs *DER* 'believe', *GUMDA* 'gargle', *JER* 'run', and *GUSAI* 'envelop' are members of class $+N^b$, but the verbs *APPADE* 'imagine', *G7A* 'drink', *ŊA* 'walk', and *PALLUD* 'enclose' are members of class $-N^b$; in these instances, it seems unlikely that the observed difference in morphology can be attributed to a systematic difference in meaning. Moreover, there are near-synonyms which differ with respect to their membership in the classes $+N^b$ and $-N^b$, e.g. *OPPUŋ* 'talk' ($-N^b$) and *BER* 'talk' ($+N^b$), *ALəŋ* 'thatch' ($-N^b$)

Table 2. Partial inflectional paradigm of the Sora verb *kuŋ* ‘shave’

			[‘shave oneself’]	[‘shave (s. o. else)’]
Non-past Affirmative	Sg	1 st	<i>kuŋ-te-n-ay</i>	<i>kuŋ-t-ay</i>
		2 nd	<i>kuŋ-te-n</i>	<i>kuŋ-t-ε</i>
		3 rd	<i>kuŋ-te-n</i>	<i>kuŋ-t-ε</i>
	Pl	1 st incl	<i>kuŋ-te-n-be</i>	<i>kuŋ-tə-be</i>
		1 st excl	<i>ə-kuŋ-te-n-ay</i>	<i>ə-kuŋ-t-ay</i>
		2 nd	<i>ə-kuŋ-te-n</i>	<i>ə-kuŋ-t-ε</i>
Past Affirmative	Sg	1 st	<i>kuŋ-le-n-ay</i>	<i>kuŋ-l-ay</i>
		2 nd	<i>kuŋ-le-n</i>	<i>kuŋ-l-ε</i>
		3 rd	<i>kuŋ-le-n</i>	<i>kuŋ-l-ε</i>
	Pl	1 st incl	<i>kuŋ-le-n-be</i>	<i>kuŋ-lə-be</i>
		1 st excl	<i>ə-kuŋ-le-n-ay</i>	<i>ə-kuŋ-l-ay</i>
		2 nd	<i>ə-kuŋ-le-n</i>	<i>ə-kuŋ-l-ε</i>
		3 rd	<i>kuŋ-le-n-ji</i>	<i>kuŋ-l-ε-ji</i>
Biligiri’s (1965) classification:			$\pm N^b$	

and *JUMAL* ‘thatch’ ($+N^b$), and so on. This, then, is one instance of morphological versatility in the Sora system of verb inflection: in the inflection of verbs belonging to class $\pm N^b$, the suffix *-n* functions as an exponent of reflexivity;³ in the inflection of verbs belonging only to class $+N^b$, by contrast, *-n* has the effect of nothing more than a mark of inflection-class membership.

An additional complication in the Sora system of verb inflection relates to a class of fifty or so motion verbs exemplified by the verbs *DUŋ* ‘set out to go’ and *YER* ‘go’ in Table 3. The inflection of these verbs differs strikingly from that of verbs belonging to classes $+N^b$, $-N^b$, $\pm N^b$ and *I*; the different uses to which they put the suffix *-ay* is particularly remarkable. In the inflection of a verb belonging to class $+N^b$, $-N^b$ or $\pm N^b$, the appearance of the first-person suffix *-ay* is conditioned by the verb’s subject-agreement properties: *-ay* appears in the verb’s first-person singular and first-person plural exclusive forms (cf. Tables 1 and 2). In the inflection of a verb belonging to class *I*, the appearance of *-ay* is conditioned by the verb’s object-agreement properties: *-ay* appears in the verb’s first-person plural inclusive forms (cf. Table 1). But in the inflection of a verb of the type exemplified by *DUŋ* or *YER*, the appearance of the first-person suffix *-ay* is instead conditioned by the direction of the motion to which the verb refers: the presence of *-ay* expresses motion towards the speaker (e.g. *duŋ-te-n-ay* ‘s/he sets out to come’), and the absence of *-ay* expresses motion in some other direction (*duŋ-te-n* ‘s/he sets out to go’); as a consequence, verbs of this type exhibit no overt person agreement with a first-person exclusive subject. Thus, the varied

uses of the suffix *-ay* constitute a second instance of morphological versatility in Sora verb inflection.

Table 3. Partial inflectional paradigms of two Sora verbs

		duŋ ‘set out to go’		YER ‘go’		
		[‘set out to go’]	[‘set out to come’]	[‘go’]	[‘come’]	
Non-past Affirmative	Sg	1 st	duŋ-te-n	duŋ-te-n-ay	yer-t-ε	yer-t-ay
		2 nd	duŋ-te-n	duŋ-te-n-ay	yer-t-ε	yer-t-ay
		3 rd	duŋ-te-n	duŋ-te-n-ay	yer-t-ε	yer-t-ay
	Pl	1 st incl	duŋ-te-n-be	duŋ-te-n-ay-be	yer-tə-be	yer-t-ay-be
		1 st excl	ə-duŋ-te-n	ə-duŋ-te-n-ay	ə-yer-t-ε	ə-yer-t-ay
		2 nd	ə-duŋ-te-n	ə-duŋ-te-n-ay	ə-yer-t-ε	ə-yer-t-ay
Past Affirmative	Sg	1 st	duŋ-le-n	duŋ-le-n-ay	yer-r-ε	yer-r-ay
		2 nd	duŋ-le-n	duŋ-le-n-ay	yer-r-ε	yer-r-ay
		3 rd	duŋ-le-n	duŋ-le-n-ay	yer-r-ε	yer-r-ay
	Pl	1 st incl	duŋ-le-n-be	duŋ-le-n-ay-be	yer-rə-be	yer-r-ay-be
		1 st excl	ə-duŋ-le-n	ə-duŋ-le-n-ay	ə-yer-r-ε	ə-yer-r-ay
		2 nd	ə-duŋ-le-n	ə-duŋ-le-n-ay	ə-yer-r-ε	ə-yer-r-ay
		3 rd	duŋ-le-n-ji	duŋ-le-n-a-ji	yer-r-ε-ji	yer-r-a-ji
Biligiri’s (1965) classification:		+N ^a		-N ^a		

Some of the verbs which inflect according to the pattern in Table 3 take the *-n* suffix (*duŋ-te-n* ‘s/he sets out to go’), while others do not (*yer-t-ε* ‘s/he goes’; the verbs in (4) behave similarly). Accordingly, Biligiri (1965) labels them as members of class +N^a or class -N^a, respectively. The superscript ‘a’ in these labels signals the use of *-ay* to express direction of motion; by contrast, the superscript ‘b’ in the labels +N^b, -N^b and ±N^b signals the use of *-ay* to express subject agreement.

- (4) -N^a: ɪɪ ‘go’, URUŋ ‘take’, AD ‘drive (cattle)’, ADU ‘reach’, JUNJUN ‘escort’, ...

Since the class of verbs that are semantically reflexivisable and the class of verbs that refer to directional motion intersect, we naturally find some verbs whose inflection expresses both reflexiveness and directionality; an example is the verb ɸAŋ ‘take/bring’ in Table 4. Biligiri assigns this verb to class ±N^a in order to account for its multiple inflectional possibilities: thus, ɸAŋ may be reflexive but not involve speaker-oriented motion (‘take for oneself’); it may be reflexive and also involve speaker-oriented motion (‘bring for oneself’); it may be non-reflexive and not involve speaker-oriented motion (‘take for someone else’); or it may be non-reflexive yet involve speaker-oriented motion (‘bring for someone else’).

Table 4. Partial inflectional paradigm of the Sora verb paŋ ‘take/bring’

		['take for oneself']	['bring for oneself']	['take (for s. o. else)']	['bring (for s. o. else)']		
Non-past Affirmative	Sg	1 st	paŋ-te-n	paŋ-te-n-ay	paŋ-t-ε	paŋ-t-ay	
		2 nd	paŋ-te-n	paŋ-te-n-ay	paŋ-t-ε	paŋ-t-ay	
		3 rd	paŋ-te-n	paŋ-te-n-ay	paŋ-t-ε	paŋ-t-ay	
	Pl	1 st incl	paŋ-te-n-be	paŋ-te-n-ay-be	paŋ-t-ε-be	paŋ-t-ay-be	
		1 st excl	ə-paŋ-te-n	ə-paŋ-te-n-ay	ə-paŋ-t-ε	ə-paŋ-t-ay	
		2 nd	ə-paŋ-te-n	ə-paŋ-te-n-ay	ə-paŋ-t-ε	ə-paŋ-t-ay	
	Past Affirmative	Sg	3 rd	paŋ-te-n-ji	paŋ-te-n-a-ji	paŋ-t-ε-ji	paŋ-t-a-ji
			1 st	paŋ-le-n	paŋ-le-n-ay	paŋ-t-ε	paŋ-l-ay
			2 nd	paŋ-le-n	paŋ-le-n-ay	paŋ-t-ε	paŋ-l-ay
Pl		3 rd	paŋ-le-n	paŋ-le-n-ay	paŋ-t-ε	paŋ-l-ay	
		1 st incl	paŋ-le-n-be	paŋ-le-n-ay-be	paŋ-t-ε-be	paŋ-l-ay-be	
		1 st excl	ə-paŋ-le-n	ə-paŋ-le-n-ay	ə-paŋ-t-ε	ə-paŋ-l-ay	
			2 nd	ə-paŋ-le-n	ə-paŋ-le-n-ay	ə-paŋ-t-ε	ə-paŋ-l-ay
			3 rd	paŋ-le-n-ji	paŋ-le-n-a-ji	paŋ-t-ε-ji	paŋ-l-a-ji

±N^a

Bilgiri's (1965) classification:

As the examples in Tables 2–4 show, the three basic verb classes exemplified in Table 1 (= classes +N^b, -N^b and I) are supplemented by four variations on these classes: the class ±N^b of reflexivisable verbs exemplified in Table 2, the classes +N^a and -N^a of directional motion verbs exemplified in Table 3, and the class ±N^a of reflexivisable directional motion verbs exemplified in Table 4. What is striking about the latter four classes is that their members employ exactly the same morphological resources as members of the three basic classes, but use these resources to express special content. That is, it would be wrong to say that Sora has three -ay suffixes—a first-person exclusive subject-agreement suffix -ay₁, a first-person plural inclusive suffix -ay₂, and a speaker-directed motion suffix -ay₃—since one would then have to regard their mutual exclusion, their syntagmatic parallelism, and their shared component of meaning (‘first person’) as a coincidence. For the same reason, it would be wrong to say that Sora has two -n suffixes—an inflection-class suffix -n₁ and a reflexive suffix -n₂.

Processes of derivation and compounding allow members of one class to give rise to members of another class. For instance, a transitive verb of class +N^b, -N^b or ±N^b, whose personal inflection encodes its subject, can give rise to an impersonal verb (class I), whose personal inflection encodes its object; this sort of derivation (examples of which are given in Table 5) is regularly exploited for the expression of pronominal objects in Sora (Biligiri 1965:240f). It is also possible to form new verbs—personal as well as impersonal—by means of noun incorporation; as the examples in Table 6 show, both object and subject incorporation are possible.

Table 5. Examples of impersonalisation in Sora

Operation	Base lexeme	Inflected forms	Derivative lexeme	Inflected forms
-N ^b → I	DER ‘believe’	derrenay ‘I believed’	DER ≈ ‘be believed’	derriñ ≈ ‘I was believed’
+N ^b → I	Gʝ ‘see’	gijlay ‘I saw’	Gʝ ≈ ‘be seen’	gijliñ ≈ ‘I was seen’
±N ^b → I	Bʙ ‘drown’	biblenay ‘I drowned myself’ biḅlay ‘I drowned (s.o. else)’	Bʙ ≈ ‘be drowned’	bibliñ ≈ ‘I was drowned’

Impersonal transitive verbs arising by word-formation processes of these sorts present a final inflectional complication. Thus, consider the paradigms of the impersonal transitive verbs Gʝ (≈ ‘be seen’) and ŃAM-KID (≈ ‘be seized by tiger’) in Table 7. These paradigms exhibit exactly the inflectional markings found in the paradigm of the basic impersonal verb DE ‘become’ in Table 1. The forms in Table 7 do, however, allow additional inflectional markings; in particular, they allow the prefix ə- and the suffix -ji to appear as expressions of subject agreement.

Table 6. Examples of noun incorporation in Sora

Operation	Sample lexeme	Inflected form
$V + N_{\text{obj}} \rightarrow V [+N^b]$	$\check{N}AM + KID \rightarrow \check{N}AM-KID$ catch tiger catch-tiger	$\check{n}amkittenay$ ‘I shall catch the tiger’
$V + N_{\text{obj}} \rightarrow V [\pm N^b]$	$KU\eta + B\check{E}B \rightarrow KU\eta-B\check{E}B$ shave head shave-head	$ku\eta b\check{e}bte$ ‘you shave s.o.’s head’ $ku\eta b\check{e}bten$ ‘you shave your head’
$V + N_{\text{subj}} \rightarrow V [I]$	$\check{N}AM + KID \rightarrow \check{N}AM-KID$ catch tiger catch-tiger	$\check{n}amkitti\check{n}$ \approx ‘I shall be caught by the tiger’

(Recall that, by contrast, the agreement markings in Table 7 themselves encode object agreement.) Thus, the form *gij-t-iñ* (\approx ‘I am seen’) appears with a second-person singular subject in (2a); in (5a), however, this same form exhibits the prefix $\check{\theta}$, which encodes the plural number of the subject. (Compare the inflection of the class $+N^b$ and class $-N^b$ verbs in Table 1, where $\check{\theta}$ appears as a default mark of plural subject agreement.) Similarly, the form *gij-t-əm* (\approx ‘you (sg.) are seen’) appears with a third-person singular subject in (2b); in (5b), however, this same form exhibits the suffix *-ji*, which encodes third-person plural subject agreement. (Compare again the inflection of the class $+N^b$ and class $-N^b$ verbs in Table 1, where *-ji* appears as a mark of third-person plural subject agreement.) But notice, too, that *-ji* appears as the mark of third-person plural object agreement in forms such as *gij-tə-jj* (\approx ‘they are seen’), and in such forms, the use of an additional *-ji* as a mark of third-person plural subject agreement is not allowed: **gij-tə-jj-ji*. Thus, just as in the cases of *-n* and *-ay* examined above, here too we find the same morphology being used in different functions; it would be wrong to say that Sora has two *-ji* suffixes (one expressing subject agreement, the other expressing object agreement), since one would then have to regard their mutual exclusion, their syntagmatic parallelism, and their shared component of meaning (‘third person plural’) as purely coincidental.

- (5) a. $\check{\theta}mben$ $\check{\theta}gij-t-i\check{n}$. b. $\check{\theta}ninji$ $gij-t-əm-ji$.
 you.PL PL-SEE-NONPAST-me they see-NONPAST-YOU.SG-3PL
 ‘you (pl.) see me’ ‘they see you (sg.)’

(Biligiri 1965:240f)

In summary, the Sora system of verb inflection presents an important challenge for morphological theory: that of accounting for the fact that a single morphological marking may vary in the exponence relations in which it participates. Thus, consider the verb-forms in (6). These forms are identical in their morphology, yet the morphology varies in the content that it expresses. In all forms, the suffix *-te* is an exponent of past tense. The *-n* suffix, by contrast, seems to serve as a mere inflection-class marker in (6a,c) but as an expression of reflexivity in

Table 7. Partial inflectional paradigms of two derived verbs in Sora

			Gɣ ≈ ‘be seen’	ŋAM-KID ≈ ‘be seized by tiger’
Nonpast Affirmative	Sg	1 st	gij-t-iñ ≈ ‘I am seen’	ñam-kit-t-iñ
		2 nd	gij-t-əm	ñam-kit-t-əm
		3 rd	gij-t-e	ñam-kit-t-e
	Pl	1 st incl	gij-t-ay	ñam-kit-t-ay
		1 st excl	gij-tə-lɛn	ñam-kit-tə-lɛn
		2 nd	gij-tə-ben	ñam-kit-tə-ben
		3 rd	gij-tə-ji	ñam-kit-tə-ji
Past Affirmative	Sg	1 st	gij-l-iñ	ñam-kil-l-iñ
		2 nd	gij-l-əm	ñam-kil-l-əm
		3 rd	gij-l-e	ñam-kil-l-e
	Pl	1 st incl	gij-l-ay	ñam-kil-l-ay
		1 st excl	gij-lə-lɛn	ñam-kil-lə-lɛn
		2 nd	gij-lə-ben	ñam-kil-lə-ben
		3 rd	gij-lə-ji	ñam-kil-lə-ji
Biligiri’s (1965) classification:			(-N→)I	I

(6b,d). The first-person suffix *-ay* expresses exclusive subject agreement in (6a,b) but expresses speaker-directed motion in (6c,d); moreover, this same suffix expresses inclusive object agreement in impersonal forms such as (7). Similarly, the third-person plural suffix *-ji* expresses subject agreement in (8a) but object agreement in (8b).

- (6) a. *de-te-n-ay* ‘I will get up’
 b. *kuŋ-te-n-ay* ‘I will shave myself’
 c. *duŋ-te-n-ay* ‘s/he will set out to come’
 d. *paŋ-te-n-ay* ‘s/he will bring for her/himself’
- (7) *gij-t-ay* ≈ ‘we (incl.) are seen’
- (8) a. *gij-t-əm-ji* ‘they see you (sg.)’
 b. *gij-tə-ji* ≈ ‘they are seen’

The Sora facts presented here do not admit a credible morpheme-based analysis, since the inflectional markings *-ay*, *-n* and *-ji* do not behave like morphemes: although they are constant in their phonological shape and in their morphological distribution, they are not constant in the content that they express. A morpheme-based analysis of Sora would therefore have to involve three first-person *-ay*

suffixes, two *-n* suffixes, and two third-person plural *-ji* suffixes, and would have to portray as coincidental the fact that the members of each group of identical suffixes are mutually exclusive, are identical in their syntagmatic potential, and have a shared component of meaning. Clearly a richer theory is needed to account for the fact that in Sora, sameness of morphological form does not necessitate sameness of content.

In the following two sections, I discuss two key characteristics of a theoretical framework capable of accounting for the Sora facts: inward rules of referral (section 2) and contrast-dependent morphemes (section 3). These characteristics imply a general theory of inflectional morphology that is both realisational and inferential (Stump 2001: Chapter 1): *REALISATIONAL* in the sense that it deduces the form of a word *w* realising a lexeme *L* from *w*'s morphosyntactic properties and *L*'s lexical properties; *INFERENCEAL* in the sense that it expresses associations between grammatical properties and their inflectional exponents as rules rather than as lexical entries.⁴

2. INWARD REFERRALS IN SORA VERB MORPHOLOGY

The Sora suffix *-ay* serves different purposes in different paradigms. In the inflection of personal non-directional verbs (members of Biligiri's classes $+N^b$, $-N^b$, and $\pm N^b$), it realises first-person exclusive subject agreement, as in (6a,b); in the inflection of personal directional verbs (members of Biligiri's classes $+N^a$, $-N^a$, and $\pm N^a$), it realises speaker orientation, as in (6c,d); and in the inflection of impersonal verbs (members of Biligiri's class 1), it realises first-person plural inclusive object agreement, as in (7). These facts are summarised in Table 8. What unifies these three uses of *-ay* is the fact that in each instance, *-ay* realises a set-valued feature⁵ whose value includes the property 'first person'. This generalisation is expressible by means of rules of referral.

Table 8. Uses of the Sora suffix *-ay*

Property set realised by <i>-ay</i>	Verb class	Example
{AGR(su):{1 st excl}}	[+PERS, -DIR] verbs	(6a,b)
{ORI:{1 st excl}}	[+PERS, +DIR] verbs	(6c,d)
{AGR(ob):{1 st pl incl}}	[-PERS] verbs	(7)

In general, a language's rules of inflectional realisation are of two sorts (Stump 2001:36ff, Zwicky 1985): a *RULE OF EXPONENCE* directly associates a particular morphosyntactic property set with a particular morphological marking; by contrast, a *RULE OF REFERRAL* allows the exponence of one property set to be stipulated to be identical to that of some contrasting property set (whatever the

exponence of the latter set might happen to be). Rules of the latter sort have been shown to be well suited for capturing generalisations about syncretism (Stump 1993, 2001; Baerman 2004). For instance, a rule of referral can be used to capture the generalisation that in Sanskrit, a neuter noun's nominative forms are morphologically identical to its accusative forms, whatever the morphology of the latter forms might be; in this instance, the relevant rule of referral stipulates that the exponence of the property sets associated with one group of word-forms (a neuter noun's nominative word-forms) is identical to the exponence of the property sets associated with a distinct group of word-forms (that noun's accusative word-forms).

Rules of referral can, however, account for more than just syncretism (Stump 2001:244f), and indeed, the rules of referral that I shall propose to account for the properties of Sora's *-ay* suffix are rather different in character from the Sanskrit neuter nominative rule; they are what I shall call rules of inward referral. Where *F* is a set-valued feature, a rule of INWARD REFERRAL realises the property set $\{F:\tau\}$ through the realisation of the property set τ .⁶ Three such features are at issue here. The values of the subject-agreement feature *AGR(su)* and of the object-agreement feature *AGR(ob)* are sets of properties specifying person, number, and (in the first person) inclusiveness; the values of the feature *ORI* of directional orientation are $\{1^{st} \text{ excl}\}$ and $\{\text{other}\}$. The three uses of *-ay* schematised in Table 8 differ insofar as they involve the realisation of the three distinct features *AGR(su)*, *AGR(ob)* and *ORI*; but they are alike insofar as the value of each of the realised features contains the property 'first person'. Thus, one can envisage an analysis of Sora verb morphology in which the three uses of *-ay* in Table 8 reflect the operation of three distinct rules of inward referral all of which converge on a single rule of exponence realising the property 'first person' through the suffixation of *-ay*.

The versatility of the third-person plural suffix *-ji* admits the same sort of analysis. This suffix serves two different purposes: in the inflection of all verbs, it realises third-person plural subject agreement, as in (8a), and in the inflection of impersonal verbs (members of Biligiri's class 1), it realises third-person plural object agreement, as in (8b). These facts are summarised in Table 9.

Table 9. Uses of the Sora suffix *-ji*

Property set realised by <i>-ji</i>	Verb class	Example
$\{\text{AGR(su)}:\{3^{rd} \text{ pl}\}\}$	all verbs	(8a)
$\{\text{AGR(ob)}:\{3^{rd} \text{ pl}\}\}$	$[-\text{PERS}]$ verbs	(8b)

Here, too, we see a common realisation for distinct set-valued features whose values are alike. Thus, one can assume an analysis of Sora verb morphology in which the two uses of *-ji* in Table 9 reflect the operation of two rules of inward

referral which converge on a single rule of exponence realising the properties ‘third-person plural’ through the suffixation of *-ji*.

One might be tempted to construe this analysis as involving the neutralisation of certain morphosyntactic distinctions; neutralisation, however, ordinarily refers to a nullification of the contrast between values of the same feature, and that isn’t what is involved here. Instead, this analysis involves a common realisation for distinct set-valued features whose values are alike. For this same reason, this analysis cannot be seen as involving an “impoverishment” rule, by which the deletion of a single feature’s value allows that feature to assume its default value (Noyer 1998).

3. A CONTRAST-DEPENDENT MORPHEME IN SORA VERB MORPHOLOGY

Like the suffixes *-ay* and *-ji*, the suffix *-n* serves different purposes in different paradigms. In the paradigms of reflexivisable verbs such as *KUŋ* ‘shave’ and *PAŋ* ‘take/bring’, it serves to distinguish those forms having a reflexive interpretation from those having a non-reflexive interpretation. In the paradigms of personal non-reflexivisable verbs, by contrast, it has the effect of an inflection-class marker: as a matter of pure stipulation, some such verbs (e.g. *DE* ‘get up’, *DUŋ* ‘set out to go’) exhibit the *-n* suffix throughout their paradigm, while others (e.g. *BƏD* ‘make’, *YER* ‘go’) uniformly fail to exhibit it.

The simplest way of accounting for these facts is to assume (i) that personal verb-forms with the *-n* suffix have the property ‘REFL:yes’ and those lacking it have the contrasting property ‘REFL:no’; but (ii) that these properties are only semantically significant in paradigms in which they stand in direct contrast with one another. On this approach, a form realising the property set {REFL:yes X} only receives a reflexive interpretation if there is a form realising the property set {REFL:no X} in the same paradigm; in the same way, a form realising {REFL:no X} only receives an explicitly non-reflexive interpretation if there is a form realising {REFL:yes X} in the same paradigm.

One might question this approach on the grounds that it is incompatible with both (a) the assumption that a given morphosyntactic property is either always meaningful or never meaningful, and (b) the assumption that a word-form’s semantic interpretation is fully determined by its morphosyntactic properties together with the lexical semantic content of the lexeme that it realises. But assumptions (a) and (b) are by no means necessary; indeed, there is no logical or empirical obstacle to the assumption that a morphosyntactic property’s significance in a word-form’s interpretation may depend on the presence of contrasting forms in the same paradigm.

In the terminology of Aronoff (1994), the morphosyntactic properties ‘REFL:yes’ and ‘REFL:no’ are MORPHEMES in Sora: words that share the same

specification for the feature REFL do not thereby necessarily share any syntactic or semantic property; they are simply alike in their morphology. Even so, the properties 'REFL:yes' and 'REFL:no' have specific semantic consequences when they appear in the right sort of paradigm. In particular, they are what I shall call CONTRAST-DEPENDENT MORPHOMES: a morpheme of this sort is a morphosyntactic property whose semantic significance in a given paradigm depends on whether it directly contrasts (in that paradigm) with some other property belonging to the same inflectional category. The existence of such morphemes⁷ provides yet another argument for the centrality of paradigms in the definition of a language's grammar.⁸

4. FORMAL ANALYSIS OF A FRAGMENT OF SORA VERB MORPHOLOGY

In this section, I develop a formal analysis of Sora verb inflection in which the postulation of rules of inward referral and contrast-dependent morphemes affords a straightforward account of the versatility of the suffixes *-ay*, *-n*, and *-ji*. In this analysis, I make a number of uncontroversial assumptions about Sora verbal lexemes:

- (i) Each lexeme has a paradigm consisting of a set of cells: each such cell is the pairing of the lexeme's root with a particular set of morphosyntactic properties,⁹ and each has a particular word-form as its realisation.
- (ii) Every verbal lexeme is lexically specified as either personal [+PERS] or impersonal [-PERS]. The former class comprises Biligiri's classes $+N^b$, $-N^b$, $\pm N^b$, $+N^a$, $-N^a$, and $\pm N^a$, and the latter coincides with Biligiri's class I.
- (iii) A given verbal lexeme may be lexically specified as possessing semantic directionality. For present purposes, I shall use the lexical feature [\pm DIR] to represent a lexeme's specification for directionality; the [+DIR] class comprises Biligiri's classes $+N^a$, $-N^a$, and $\pm N^a$, and the [-DIR] comprises the remaining classes.
- (iv) A given verbal lexeme may be lexically specified as possessing reflexivability. The lexical feature [\pm REFLEXIVIZABLE] will be used to represent a lexeme's specification for reflexivability; [+REFLEXIVIZABLE] comprises Biligiri's classes $\pm N^b$ and $\pm N^a$, and [-REFLEXIVIZABLE], the remaining classes.
- (v) The dimensions of a lexeme's finite paradigm depend on the class to which it belongs. Every cell in every finite verbal paradigm is specified for the morphosyntactic features of subject agreement (AGR(SU)), tense (TNS), and polarity (POL). On the other hand, only cells in the

finite paradigms of [+PERS] verbs are specified for the morphosyntactic feature REFL: in the finite paradigm of a [+REFLEXIVIZABLE] verb, half the cells have the property ‘REFL:yes’ and the remainder have ‘REFL:no’; in the finite paradigm of a personal verb that is [−REFLEXIVIZABLE], either all cells have the property ‘REFL:yes’ (= Biligiri’s classes +N^b and +N^a) or all cells have ‘REFL:no’ (= classes −N^b and −N^a). In addition, only cells in the finite paradigms of [+DIR] verbs are specified for the morphosyntactic feature of directional orientation (ORI): in such paradigms, half the cells have the property ‘ORI:{1st excl}’ of speaker-oriented motion, and the remainder have the property ‘ORI:{other}’ of motion not toward the speaker. Only the finite paradigms of impersonal verbs have cells specified for the object-agreement feature AGR(ob); in addition, the cells in the finite paradigm of the impersonal verb DE ‘become’ have the expletive subject-agreement specification ‘AGR(su):{3rd sg}’. These facts are summarised in Table 10.

Table 10. Lexemic properties and paradigm dimensions of eight Sora verbs

Sample lexeme	Biligiri’s class	Lexemic properties	Dimensions of the lexeme’s inflectional paradigm
DE ‘get up’	+N ^b	[+PERS, −DIR, −REFLEXIVIZABLE]	AGR(su) × <u>REFL: yes</u> × TNS × POL
BƏD ‘make’	−N ^b	[+PERS, −DIR, −REFLEXIVIZABLE]	AGR(su) × <u>REFL: no</u> × TNS × POL
KUŋ ‘shave’	±N ^b	[+PERS, −DIR, +REFLEXIVIZABLE]	AGR(su) × REFL × TNS × POL
PAŋ ‘take/bring’	±N ^a	[+PERS, +DIR, +REFLEXIVIZABLE]	AGR(su) × REFL × ORI × TNS × POL
DUŋ ‘set out to go’	+N ^a	[+PERS, +DIR, −REFLEXIVIZABLE]	AGR(su) × <u>REFL: yes</u> × ORI × TNS × POL
YER ‘go’	−N ^a	[+PERS, +DIR, −REFLEXIVIZABLE]	AGR(su) × <u>REFL: no</u> × ORI × TNS × POL
DE ‘become’	I	[−PERS, −DIR, −REFLEXIVIZABLE]	<u>AGR(su):{3rd sg}</u> × AGR(ob) × TNS × POL
GH ≈ ‘be seen’	I	[−PERS, −DIR, −REFLEXIVIZABLE]	AGR(su) × AGR(ob) × TNS × POL

N.B.: Underlined properties are lexically stipulated.

The realisation of a given cell $\langle r, \sigma \rangle$ in a Sora verb’s paradigm is determined by the paradigm function PF_{Sora} ; in particular, PF_{Sora} applies to $\langle r, \sigma \rangle$ to yield that pairing $\langle Z, \sigma \rangle$ such that Z is the realisation of $\langle r, \sigma \rangle$. The relevant clause in the

definition of this paradigm function specifies the interaction of five blocks of realisation rules, as in (9).

- (9) Definition of the Sora paradigm function (clause relating to verbs)
 $PF_{\text{Sora}}(\langle r, \sigma \rangle) = (((\langle r, \sigma \rangle : \mathbf{Block1}) : \mathbf{Block2}) : \mathbf{Block3}) : \mathbf{Block4}) : \mathbf{Block5}$

In this definition, the notation “ $\langle X, \sigma \rangle : \text{Block } n$ ” means “ $\langle Y, \tau \rangle$, where $\langle Y, \tau \rangle$ is the result of applying the narrowest applicable rule in **Block** n to the pairing $\langle X, \sigma \rangle$ ”. The pairing $\langle Y, \tau \rangle$ will ordinarily be such that $\sigma = \tau$, since most realisation rules apply to a pairing $\langle X, \sigma \rangle$ to yield a value $\langle Y, \sigma \rangle$. A rule of referral, however, may apply to a pairing $\langle X, \sigma \rangle$ to yield a value $\langle Y, \tau \rangle$ such that $\sigma \neq \tau$; for instance, the rules referral to be defined in (12f,g,k,r,s) below have this effect. Nevertheless, the definition of the “ $\langle X, \sigma \rangle : \text{Block } n$ ” notation causes the property set with which Y is paired to be “reset” to that of the pairing to which **Block** n applies.

The conception of rule narrowness presupposed by the “ $\langle X, \sigma \rangle : \text{Block } n$ ” notation is defined in (10) (cf. Stump 2001:52); this notation embodies Pāṇini’s principle.

- (10) Rule **P** IS NARROWER THAN rule **Q** iff
- (i) the class of lexemes in whose inflection **P** applies is a proper subset of the class in whose inflection **Q** applies; or
 - (ii) both rules apply in the inflection of the same class of lexemes but the set of morphosyntactic properties realised by **P** is a proper extension¹⁰ of that realised by **Q**.

Each of the realisation rules constituting a rule block has the format in (11). A rule in this format is to be read as follows: given the pairing $\langle X, \sigma \rangle$ of a stem X with a morphosyntactic property set σ such that σ is an extension of τ and X belongs to class C , this rule (a member of **Block** m) realises $\langle X, \sigma \rangle$ as Y . Y is in all instances the pairing $\langle Z, \rho \rangle$ of a form Z with a morphosyntactic property set ρ . In a rule of exponence, the identity of the pairing Y is stipulated as part of the formulation of the rule itself; in all such instances, $\rho = \sigma$. In a rule of referral, the identity of Y is invariably specified by means of the “ $\langle X, \sigma \rangle : \text{Block } n$ ” notation.

- (11) Format for realisation rules: **Block** m : $\langle X, \sigma \rangle ; C, \tau = Y$

In these formats, the blocks of realisation rules for Sora verb inflection may be formulated as in (12). The rules in **Block1** introduce the tense suffixes $-tE$ and $-lE$; the morphophoneme E in these rules is realised as $[e]$ before $[n]$ (as in *delenbe* ‘we (incl.) got up’) and otherwise as $[\emptyset]$ (as in *bədləbe* ‘we (incl.) made’).

(12) Sora rule blocks:

- Block1:** a. $\langle X, \sigma \rangle ; V, \{\text{TNS:non-past}\}$ = $\langle Xt\varepsilon, \sigma \rangle$
 b. $\langle X, \sigma \rangle ; V, \{\text{TNS:past}\}$ = $\langle Xl\varepsilon, \sigma \rangle$
- Block2:** c. $\langle X, \sigma \rangle ; V, \{\text{REFL:yes}\}$ = $\langle Xn, \sigma \rangle$
 d. $\langle X, \sigma \rangle ; V, \{\text{REFL:no}\}$ = $\langle X\varepsilon, \sigma \rangle$
 e. $\langle X, \sigma \rangle ; V, \{\text{AGR(su)}:\{1^{\text{st}} \text{ pl incl}\}, \text{REFL:no}\}$ = $\langle X, \sigma \rangle$
- Block3:** f. $\langle X, \sigma \rangle ; V[+\text{PERS}, -\text{DIR}], \{\text{AGR(su)}:\{1^{\text{st}} \text{ excl}\}_i\}$ = $\langle X, \tau_i \rangle$: **Block3**
 g. $\langle X, \sigma \rangle ; V[+\text{PERS}, +\text{DIR}], \{\text{ORI}:\{1^{\text{st}} \text{ excl}\}_i\}$ = $\langle X, \tau_i \rangle$: **Block3**
 h. $\langle X, \sigma \rangle ; V[-\text{PERS}], \{\text{AGR(ob)}:\{1^{\text{st}} \text{ sg}\}\}$ = $\langle Xiñ, \sigma \rangle$
 i. $\langle X, \sigma \rangle ; V[-\text{PERS}], \{\text{AGR(ob)}:\{2^{\text{nd}} \text{ sg}\}\}$ = $\langle X\varepsilon m, \sigma \rangle$
 j. $\langle X, \sigma \rangle ; V[-\text{PERS}], \{\text{AGR(ob)}:\{3^{\text{rd}} \text{ sg}\}\}$ = $\langle Xe, \sigma \rangle$
 k. $\langle X, \sigma \rangle ; V[-\text{PERS}], \{\text{AGR(ob)}:\{1^{\text{st}} \text{ pl incl}\}_i\}$ = $\langle X, \tau_i \rangle$: **Block3**
 l. $\langle X, \sigma \rangle ; V[-\text{PERS}], \{\text{AGR(ob)}:\{1^{\text{st}} \text{ pl excl}\}\}$ = $\langle Xl\varepsilon n, \sigma \rangle$
 m. $\langle X, \sigma \rangle ; V[-\text{PERS}], \{\text{AGR(ob)}:\{2^{\text{nd}} \text{ pl}\}\}$ = $\langle Xben, \sigma \rangle$
 n. $\langle X, \sigma \rangle ; V, \{1^{\text{st}}\}$ = $\langle Xay, \sigma \rangle$
- Block4:** o. $\langle X, \sigma \rangle ; V, \{\text{AGR(su)}:\{\text{pl}\}\}$ = $\langle \varepsilon X, \sigma \rangle$
 p. $\langle X, \sigma \rangle ; V[+\text{PERS}], \{\text{AGR(su)}:\{1^{\text{st}} \text{ pl incl}\}\}$ = $\langle Xbe, \sigma \rangle$
 q. $\langle X, \sigma \rangle ; V, \{\text{AGR(su)}:\{3^{\text{rd}} \text{ pl}\}\}$ = $\langle X, \sigma \rangle$
- Block5:** r. $\langle X, \sigma \rangle ; V, \{\text{AGR(su)}:\{3^{\text{rd}} \text{ pl}\}_i\}$ = $\langle X, \tau_i \rangle$: **Block5**
 s. $\langle X, \sigma \rangle ; V[-\text{PERS}], \{\text{AGR(ob)}:\{3^{\text{rd}} \text{ pl}\}_i\}$ = $\langle X, \tau_i \rangle$: **Block5**
 t. $\langle X, \sigma \rangle ; V, \{3^{\text{rd}} \text{ pl}\}$ = $\langle Xji, \sigma \rangle$

The rules in **Block2** introduce the suffixes *-n* and *-ε*, which realise the respective morphemes *REFL:yes* and *REFL:no*; their application is blind to whether these morphemes are semantically significant in the cell being realised. In accordance with the definition of narrowness in (10), rule (12e) overrides (12d), guaranteeing the absence of *-ε* from first-person plural inclusive forms. As noted earlier, a rule of Sora morphophonology causes the first of two successive vowels to be elided; thus, in forms such as *bədlε* ‘s/he made’, the *-ε* suffix causes the elision of the vowel of the tense suffix *-lε*.

The rules in **Block3** introduce most of the agreement suffixes, including all suffixes serving purely as realisations of object agreement in the inflection

of impersonal verbs. **Block3** also houses three rules of inward referral (12f,g,k), each of which causes a set-valued feature whose value includes the property ‘first person’ to be realised through the application of rule (12n). In the formulation of each rule of referral, τ_i is bound to the subscripted property set on the left-hand side of the rule; in (12f), for instance, $\tau_i = \{1^{\text{st}} \text{ excl}\}$. Thus, rule (12f) causes the realisation of the pairing (13a) to be that of (13b), namely *detenay* ‘I get up’.

- (13) a. $\langle \text{deten}, \{\text{AGR}(\text{SU}): \{1^{\text{st}} \text{ sg excl}\}, \text{REFL:yes}, \text{TNS:non-past}, \text{POL:aff}\} \rangle$
 b. $\langle \text{deten}, \{1^{\text{st}} \text{ sg excl}\} \rangle$

Some of the suffixes in **Block3** create additional contexts for the elision of the first of two successive vowels; thus, for instance, the $-\varepsilon$ suffix is elided in *bədlay* ‘I made’.

The rules in **Block4** introduce the agreement affixes \varnothing and *-be*. Rule (12o) defines the \varnothing prefix as a default mark of plural subject agreement. This rule is overridden by the rule (12p) of *-be* suffixation (in personal verb-forms with first-person plural inclusive subject agreement) and by the identity function (12q) (in forms with third-person plural subject agreement).

Block5 houses two rules of inward referral (12r,s), each of which causes a set-valued feature whose value includes the properties ‘third person plural’ to be realised through the application of rule (12t). For instance, rule (12r) causes the realisation of the pairing (14a) to be that of (14b), namely *detenji* ‘they get up’. Rule (12t) induces the truncation of a preceding [y], as in *yertaji* ‘they come’.

- (14) a. $\langle \text{deten}, \{\text{AGR}(\text{SU}): \{3^{\text{rd}} \text{ pl}\}, \text{REFL:yes}, \text{TNS:non-past}, \text{POL:aff}\} \rangle$
 b. $\langle \text{deten}, \{3^{\text{rd}} \text{ pl}\} \rangle$

I assume that each of the five rule blocks in (12) contains as its ultimate default a rule which is applicable to members of any class and which realises the empty morphosyntactic property set, in accordance with the Identity Function Default (15). The ultimate default rule in each block is broader in applicability than any other rule in that block, and therefore loses in any instance of rule competition.

- (15) Identity Function Default (cf. Stump 2001:143)
 Universally, the following rule acts as the least narrow member of any rule block:
 $\langle X, \sigma \rangle$; any, $\{\} = \langle X, \sigma \rangle$

This analysis provides a complete account of the versatility of the *-ay* and *-ji* suffixes. Thus, consider again the examples in (6). The cells realised by the verb-forms in (6a,b) have the same morphosyntactic property set, as in (16a,b); those realised by the verb-forms in (6c,d) have a different property set, as in (16c,d). Even so, their realisation differs in only one of the five rule blocks, namely **Block3**: here, the property set in (16a,b) is realised by means of the rule of inward referral in (12f), while the property set in (16c,d) is realised by means of the rule of inward referral in (12g). But because these rules of referral converge on the rule of *-ay* suffixation in (12n), all four of the forms in (6) end up with the same morphology. Moreover, the rule of inward referral in (12k) also induces the application of rule (12n) in realising cell (16e) as in (7). In this way, the *-ay* suffix, though introduced by a single rule, realises first-person exclusive subject agreement in (6a,b), speaker-directed motion in (6c,d), and first-person plural inclusive object agreement in (7).

(16) The realisation of cells in the paradigms of five Sora verbs

- | | | |
|----|--------------------|--|
| a. | Cell: | $\langle de, \{AGR(su):\{1^{st} \text{ sg excl}\}, REFL:yes, TNS:non-past, POL:aff\}\rangle$ |
| | Realisation rules: | Block1 : (12a); Block2 : (12c); Block3 : (12f), (12n); Block4 : (15); Block5 : (15) |
| | Realisation: | detenay ‘I will get up’ |
| b. | Cell: | $\langle kun, \{AGR(su):\{1^{st} \text{ sg excl}\}, REFL:yes, TNS:non-past, POL:aff\}\rangle$ |
| | Realisation rules: | Block1 : (12a); Block2 : (12c); Block3 : (12f), (12n); Block4 : (15); Block5 : (15) |
| | Realisation: | kuntenay ‘I will shave myself’ |
| c. | Cell: | $\langle duŋ, \{AGR(su):\{3^{rd} \text{ sg}\}, ORI:\{1^{st} \text{ excl}\}, REFL:yes, TNS:non-past, POL:aff\}\rangle$ |
| | Realisation rules: | Block1 : (12a); Block2 : (12c); Block3 : (12g), (12n); Block4 : (15); Block5 : (15) |
| | Realisation: | duŋtenay ‘s/he will set out to come’ |
| d. | Cell: | $\langle paŋ, \{AGR(su):\{3^{rd} \text{ sg}\}, ORI:\{1^{st} \text{ excl}\}, REFL:yes, TNS:non-past, POL:aff\}\rangle$ |
| | Realisation rules: | Block1 : (12a); Block2 : (12c); Block3 : (12g), (12n); Block4 : (15); Block5 : (15) |
| | Realisation: | paŋtenay ‘s/he will bring for her/himself’ |
| e. | Cell: | $\langle gij, \{AGR(su):\{3^{rd} \text{ sg}\}, AGR(ob):\{1^{st} \text{ pl incl}\}, TNS:non-past, POL:aff\}\rangle$ |
| | Realisation rules: | Block1 : (12a); Block2 : (15); Block3 : (12k), (12n); Block4 : (15); Block5 : (15) |
| | Realisation: | gijtay \approx ‘we (incl.) are seen’ |

Similar consequences hold true for the verb-forms in (8). Although the cells realised by the two verb-forms in (8) have rather different property sets (as in (17a,b)), each is realised through the application of one of the rules of inward referral in **Block5**: the verb-form in (8a) is realised through the application of (12r), and the verb-form in (8b) is realised through the application of (12s); thus, in both forms, the third-person plural agreement properties are realised through the application of (12t).

(17) The realisation of two cells in the paradigm of the transitive impersonal verb *gɥj* ‘see’ in Sora

- a. Cell: $\langle g\mathfrak{H}, \{AGR(su):\{3^{rd} pl\}, AGR(ob):\{2^{nd} sg\}, TNS:non-past, POL:aff\}\rangle$
 Realisation rules: **Block1**: (12a); **Block2**: (15); **Block3**: (12i); **Block4**: (12q); **Block5**: (12r), (12t)
 Realisation: *gɥjtəmji* ‘they see you (sg.)’
- b. Cell: $\langle g\mathfrak{H}, \{AGR(su):\{3^{rd} sg\}, AGR(ob):\{3^{rd} pl\}, TNS:non-past, POL:aff\}\rangle$
 Realisation rules: **Block1**: (12a); **Block2**: (15); **Block3**: (15); **Block4**: (15); **Block5**: (12s), (12t)
 Realisation: *gɥjtəji* \approx ‘they are seen’

As these examples show, the proposed analysis accounts both for the range of uses to which the suffixes *-ay* and *-ji* are put and for the fundamental morphological unity of each suffix’s uses.

The versatility of the *-n* suffix hinges on the semantics of the proposed analysis. I assume an event-based theory of thematic roles (cf. Dowty 1989), in which (for example) the semantic interpretation of the verb-form *delen* ‘s/he got up’ entails that there is an event *e* that is a getting up, that *e* took place in the past, and that the agent of *e* is the referent of the verb-form’s subject. More generally, each form *X* of a verbal lexeme *L* is represented semantically as a predicate of events in which the argument structure of *L* is represented by means of indexed variables destined to be bound by the referents of *X*’s syntactic arguments.

Central to the proposed semantics are the related notions of predicate-set and semantic representation defined in (18):

(18) Every inflected verb-form *X* has a non-empty PREDICATE-SET each of whose members is an event predicate; for any inflected verb-form *X* whose predicate-set is

$$\{\text{pred}_1, \dots, \text{pred}_n\},$$

the SEMANTIC REPRESENTATION of *X* is logically equivalent to

$$\lambda e[\text{pred}_1(e) \ \& \ \dots \ \& \ \text{pred}_n(e)].^{11}$$

For example, the predicate-set of the Sora past-tense verb-form *paŋtenay* is the set of event predicates in (19a); accordingly, the semantic representation of *paŋtenay* is logically equivalent to (19b).

(19) a. The predicate-set of the verb-form *paŋtenay* ‘brings for oneself’:

$$\left\{ \begin{array}{l} \lambda e[\text{Taking}(e)], \\ \lambda e[\text{Nonpast}(e)], \\ \lambda e[\text{Agent}(x_{\text{SUBJ}}, e)], \\ \lambda e[\text{Theme}(x_{\text{OBJ}}, e)], \\ \lambda e[\text{Speaker-directed}(e)], \\ \lambda e[\text{Reflexive}(e)] \end{array} \right\}$$

b. The semantic representation of the verb-form *paŋtenay*:
 $\lambda e[\text{Taking}(e) \ \& \ \text{Non-past}(e) \ \& \ \text{Agent}(x_{\text{SUBJ}}, e) \ \& \ \text{Theme}(x_{\text{OBJ}}, e) \ \& \ \text{Speaker-directed}(e) \ \& \ \text{Reflexive}(e)]$

I assume that the interpretation of the event predicate $\lambda e[\text{Reflexive}(e)]$ appearing in (19) varies from one class of verbs to another, in accordance with meaning postulates such as those in (20). According to (20a), if a given event of taking is reflexive, then that event’s agent is also its beneficiary; according to (20b), by contrast, a reflexive event of shaving is an event whose agent and theme are identical.

(20) Sample meaning postulates for the interpretation of $\lambda e[\text{Reflexive}(e)]$:

- a. Where $\alpha = \text{Taking}, \dots$,
 $\forall e \forall x \square [[\alpha(e) \ \& \ \text{Agent}(x,e) \ \& \ \text{Reflexive}(e)] \rightarrow [\text{Beneficiary}(x,e)]]$
- b. Where $\alpha = \text{Shaving}, \dots$,
 $\forall e \forall x \forall y \square [[\alpha(e) \ \& \ \text{Agent}(x,e) \ \& \ \text{Theme}(y,e) \ \& \ \text{Reflexive}(e)] \rightarrow [x=y]]$

I assume that the predicate-set of a verb-form X is determined by (i) the lexically listed predicate-set of the lexeme L that X realises, together with (ii) one or more rules of inflectional semantics. Thus, where the verb-form X is an inflected form of the verbal lexeme L, L’s lexical entry specifies a predicate-set that is itself a subset of X’s predicate-set; for instance, the lexical entries of the lexemes listed in the left-hand column of Table 11 specify the predicate-sets in the right-hand column.

The predicate-set of a verb-form realising a lexeme L is deduced from L’s predicate-set by means of one or more of the rules of inflectional semantics in (21). For instance, the predicate-set of *paŋtenay* ‘brings for oneself’ in (19a) is deduced from that of *paŋ* (in Table 11) by means of rules (21a,b,d,e).

Table 11. Predicate-sets of eight Sora verbal lexemes

Lexeme	Biligiri's class	Predicate-set
DE 'get up'	+N ^b	{ λe[Getting.up(e)], λe[Agent(x _{SUBJ} , e)] }
BƏD 'make'	-N ^b	{ λe[Making(e)], λe[Agent(x _{SUBJ} , e)], λe[Theme(x _{OBJ} , e)] }
KUŋ 'shave'	±N ^b	{ λe[Shaving(e)], λe[Agent(x _{SUBJ} , e)], λe[Theme(x _{OBJ} , e)] }
PAŋ 'take/bring'	±N ^a	{ λe[Taking(e)], λe[Agent(x _{SUBJ} , e)], λe[Theme(x _{OBJ} , e)] }
DUŋ 'set out to go'	+N ^a	{ λe[Setting.out.to.go(e)], λe[Agent(x _{SUBJ} , e)] }
YER 'go'	-N ^a	{ λe[Going(e)], λe[Agent(x _{SUBJ} , e)] }
DE 'become'	I	{ λe[Becoming(e)], λe[Theme(x _{OBJ} , e)] }
Gŋ ≈ 'be seen'	I	{ λe[Seeing(e)], λe[Experiencer(x _{SUBJ} , e)], λe[Goal(x _{OBJ} , e)] }

(21) Rules of inflectional semantics:

- a. If a verb-form X realises a cell ⟨L, σ⟩ in the paradigm of the verbal lexeme L, then the predicate-set of L is a subset of the predicate-set of X.
- b. If a verb-form X realises ⟨L, {TNS:non-past...}⟩, then λe[Nonpast(e)] belongs to the predicate-set of X.
- c. If a verb-form X realises ⟨L, {TNS:past...}⟩, then λe[Past(e)] belongs to the predicate-set of X.
- d. If X, Y realise the respective cells ⟨L, {REFL:yes...}⟩, ⟨L, {REFL:no...}⟩, then λe[Reflexive(e)] belongs to the predicate-set of X and λe¬[Reflexive(e)] belongs to the predicate-set of Y.
- e. If X realises the cell ⟨L, {ORI:{1st excl}...}⟩, then λe[Speaker-directed(e)] belongs to the predicate-set of X.
- f. If X realises the cell ⟨L, {ORI:{other}...}⟩, then λe¬[Speaker-directed(e)] belongs to the predicate-set of X.

The rules in (21) account for the semantic versatility of the *-n* suffix. Thus, consider first cells (22a,b) from the paradigm of KUŋ 'shave'. These two cells are alike except that (22a) is specified 'REFL:yes' while (22b) is specified 'REFL:no'; rule (21d) therefore applies, assigning a reflexive interpretation to (22a) and a non-reflexive interpretation to (22b). In this instance, the morphemes 'REFL:yes' and 'REFL:no' are semantically significant. The same is not true, however, of the cells

in (23a,b). Because the lexemes *DE* ‘get up’ and *BƏD* ‘make’ are [–REFLEXIVIZABLE], neither has a paradigm in which contrasting values of the feature *REFL* coexist. Instead, all of the cells in the finite paradigm of *DE* are specified ‘REFL:yes’ as a matter of lexical stipulation (cf. Table 10); in the same way, all of the cells in the finite paradigm of *BƏD* are specified ‘REFL:no’. As a consequence, rule (21d) is inapplicable in the interpretation of the cells in (23a,b), whose specifications for the feature *REFL* are semantically inert. Thus, the rules in (21) account for the incidence of cells (such as (22a) and (23a), or (22b) and (23b)) whose morphosyntactic property sets are identical but which differ in their inflectional semantics.

- (22) The semantic interpretation of two cells in the paradigm of *KUŋ* ‘shave’
- | | |
|----------------------------------|---|
| a. Cell: | $\langle kuŋ, \{AGR(su):\{1^{st} \text{ sg excl}\},$
REFL:yes, TNS:non-past, POL:aff}\rangle |
| Realisation: | <i>kuŋtenay</i> ‘I shave myself’ |
| Rules of inflectional semantics: | (21a,b,d) |
| b. Cell: | $\langle kuŋ, \{AGR(su):\{1^{st} \text{ sg excl}\},$
REFL:no, TNS:non-past, POL:aff}\rangle |
| Realisation: | <i>kuŋtay</i> ‘I shave (someone else)’ |
| Rules of inflectional semantics: | (21a,b,d) |
- (23) The semantic interpretation of cells in the paradigms of two Sora verbs
- | | |
|----------------------------------|--|
| a. Cell: | $\langle de, \{AGR(su):\{1^{st} \text{ sg excl}\},$
REFL:yes, TNS:non-past, POL:aff}\rangle |
| Realisation: | <i>detenay</i> ‘I get up’ |
| Rules of inflectional semantics: | (21a,b) |
| b. Cell: | $\langle bəd, \{AGR(su):\{1^{st} \text{ sg excl}\},$
REFL:no, TNS:non-past, POL:aff}\rangle |
| Realisation: | <i>bəttay</i> ‘I make’ |
| Rules of inflectional semantics: | (21a,b) |

5. CONCLUSIONS

In morpheme-based theories of morphology, inflectional formatives are assumed to exist as minimal signs or ‘vocabulary items’. Such theories are therefore fundamentally irreconcilable with the remarkable mismatch between Sora verb morphology and the syntacticosemantic content that it is used to express; in particular, they are inherently incapable of accounting for the fact that the same affix may express different content in different contexts, and indeed may be semantically significant in some contexts but not in others.

By contrast, inferential-realisation theories of morphology such as that embodied by the analysis proposed above are suited to accounting for this fact. According to such theories, word-forms realise cells in paradigms, and the application of the rules introducing a word-form's inflectional formatives is sensitive to the morphosyntactic information associated with the cell which that word-form realises. This sort of theory therefore accommodates the postulation of rules of inward referral, and as I have shown here, such rules afford a straightforward account of the versatility of the first-person suffix *-ay* and the third-person plural suffix *-ji* in Sora. In addition, inferential-realisation theories of morphology are compatible with the assumption that a word-form's semantic interpretation depends on rules of inflectional semantics applying to the cell which that word-form realises. If one assumes, as I have here, that the applicability of a rule of inflectional semantics to a given cell may be conditioned by the properties of other cells in the same paradigm, then the versatility of the Sora *-n* suffix can be attributed to its status as a contrast-dependent morpheme.

What is needed at this point is further investigation of inflectional systems which deviate from the canonical pattern in which morphological form and content are in perfect correspondence. Some such deviations can be explained by postulating rules of inward referral and rules of inflectional semantics whose application is sensitive to cell contrasts. But what other sorts of deviations are possible? And how far can such deviations go—that is, how indirect can the relation be between a word's inflectional form, the morphosyntactic property set that it realises, and its semantic interpretation? A more explicit delineation of the dimensions and extent of such deviations is essential not only for grammatical theory, but for understanding the processes of language learning and language change as well.

NOTES

¹ An earlier version of this paper was presented at the CID/M3 Workshop on Possible Words, University of Surrey, in July 2004. For extremely helpful discussion, I wish to thank the other workshop participants as well as the following commentators on an earlier manuscript version: Matthew Baerman, Dunstan Brown, Greville Corbett, Raphael Finkel, David Stampe, Jaap van Marle, and three anonymous referees for *Yearbook of Morphology*.

² Thus, as Ramamurti (1931:38f) observes, Sora *argaldaliñ* 'I was thirsty' might be better translated as 'it thirsted me'; cf. German *ihn hungerte*.

³ In the inflection of some class $\pm N^b$ verbs, *-n* signals a middle rather than a reflexive interpretation; thus, as a member of class $+N^b$, GAD 'cut' means 'be cuttable' (*gatten* 'it can be cut'). This recalls the close relation between middles and reflexives in many other languages, as e.g. in French *Cela se fait*.

⁴ For additional arguments in favor of a general theory of this sort, see Matthews (1972), Zwicky (1985), Anderson (1992), Sadler and Spencer (2001), Stump (2001), Blevins (2003), Ackerman and Stump (2004), and Luís and Spencer (2005).

⁵ A set-valued feature is a feature whose value is a set of morphosyntactic properties (Gazdar et al. 1985:25; Stump 2001:40).

⁶ Rules of inward referral provide a natural means of accounting for the properties of parallel position classes; see, for example, the analysis of Lingala verb morphology in Stump (2001:144ff).

⁷ Instances of contrast-dependent morphemes are widely observable. Another clear example is that of the active and middle voices in Sanskrit. In Sanskrit, many verbal lexemes have paradigms in which active cells directly contrast with middle cells; in these instances, active and middle voice are semantically significant. Other lexemes, however, have only active forms, or only middle forms, or active forms in one tense but middle forms in another. Thus, as a class, verb-forms belonging to the same voice are neither syntactically nor semantically uniform—their commonality is purely morphological; but in the right sort of paradigm, the voice morphemes take on semantic significance.

⁸ For other arguments, see Stump (2005) and the references listed in footnote 4.

⁹ In order to account for certain kinds of phenomena, it is useful to assume that a language's morphology defines two sorts of paradigms: CONTENT PARADIGMS (whose cells are pairings of a lexeme with a morphosyntactic property set) and FORM PARADIGMS (whose cells are pairings of a root with a property set); see Stump (2002, in press), Ackerman and Stump (2004), Stewart and Stump (in press). Because this distinction is not essential for the analysis proposed here, I formulate this analysis purely in terms of form paradigms.

¹⁰ For present purposes, extensions can be equated with supersets; see Stump (2001:41) for a more precise definition.

¹¹ The notation $\lambda e[\text{pred}(e)]$ represents the set of events of which 'pred' can be truthfully predicated.

REFERENCES

- Ackerman, Farrell, and Gregory Stump (2004). Paradigms and periphrastic expression: A study in realisation-based lexicalism. In L. Sadler and A. Spencer (eds.), *Projecting Morphology*. Stanford: CSLI, 111–157.
- Anderson, Stephen R. (1992). *A-morphous Morphology*. Cambridge: Cambridge University Press.
- Aronoff, Mark (1994). *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press.
- Baerman, Matthew (2004). Directionality and (un)natural classes in syncretism. *Language* 80, 807–827.
- Biligiri, Hemmige Shrinivasarangachar (1965). The Sora verb: a restricted study. In G. B. Milner and Eugénie J. A. Henderson (eds.), *Indo-Pacific Studies, Part II: Descriptive Linguistics [Lingua 15]*. Amsterdam: North-Holland, 231–250.
- Blevins, James P. (2003). Stems and paradigms. *Language* 79, 737–767.
- Dowty, David R. (1989). On the semantic content of the notion “thematic role”. In G. Chierchia, B. Partee, and R. Turner (eds.), *Properties, Types, and Meaning*. Dordrecht: D. Reidel, 69–129.
- Gazdar, G., E. Klein, G. Pullum, and I. Sag (1985). *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

- Luís, Ana, and Andrew Spencer (2005). A Paradigm Function account of 'mesoclis' in European Portuguese (EP). In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2004*, Dordrecht: Springer, 177–228.
- Matthews, P. H. (1972). *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge: Cambridge University Press.
- Noyer, Rolf (1998). Impoverishment theory and morphosyntactic markedness. In Lapointe, S., D. Brentari and P. Farrell (eds.), *Morphology and Its Relation to Phonology and Syntax*. Stanford: CSLI, 264–285.
- Ramamurti, G. V. (1931). *A Manual of the So:ra: (or Savara) Language*. Madras: Government Press.
- Sadler, Louisa, and Andrew Spencer (2001). Syntax as an exponent of morphological features. In G. Booij and J. van Marle (eds.), *Yearbook of Morphology 2000*. Dordrecht: Kluwer, 71–96.
- Stewart, Tom, and Gregory Stump (in press). Paradigm Function Morphology and the morphology/syntax interface. In Gillian Ramchand and Charles Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press.
- Stump, Gregory T. (1993). On rules of referral. *Language* 69, 449–479. [Reprinted in *Morphology: Critical Concepts in Linguistics*, ed. by Francis Katamba (London: Routledge, 2003).]
- Stump, Gregory T. (2001). *Inflectional Morphology*. Cambridge: Cambridge University Press.
- Stump, Gregory T. (2002). Morphological and syntactic paradigms: Arguments for a theory of paradigm linkage. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2001*, Dordrecht: Kluwer, 147–180.
- Stump, Gregory T. (in press). Heteroclis and paradigm linkage. *Language* 82.
- Stump, Gregory T. (2005). Rules about paradigms. In C. Orhan Orgun and Peter Sells (eds.), *Morphology and the Web of Grammar: Essays in Memory of Steven G. Lapointe*. Stanford: CSLI, 49–82.
- Zwicky, Arnold M. (1985). How to describe inflection. In Mary Niepokuj, Mary Van Clay, Vassiliki Nikiforidou, and Deborah Feder (eds.), *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*. Berkeley Linguistics Society, 372–386.

Department of English
University of Kentucky
1215 Patterson Office Tower
Lexington, KY 40506-0027
USA

gstump@uky.edu

Affixes, stems and allomorphic conditioning in paradigm function morphology

ANDREW CARSTAIRS-McCARTHY*

1. INTRODUCTION

Gregory T. Stump's *Inflectional Morphology: A Theory of Paradigm Structure* (2001; henceforth *IM*) is the longest and most ambitious contribution to the theory of inflection since P.H. Matthews's *Inflectional Morphology* (1972). (Stephen R. Anderson's *A-Morphous Morphology* (1992) is comparable in scale but is not so narrowly focussed on inflection.) As such, *IM* certainly deserves to be reviewed in a periodical such *Yearbook of Morphology*. At the same time, reviewing it thoroughly would be a daunting task. On the one hand, Stump examines a wide range of linguistic data in terms of an intricate network of definitions and an elaborate notation. On the other hand, teasing out the empirical implications of Stump's approach (what precisely is and is not likely to happen in inflectional morphology, in his view) is often difficult. Every theoretical proposal and every analysis in the book is likely to provoke questions in the mind of any reader who knows something about morphology, and an adequate discussion of all of these would yield a review as long as the book itself.

That being so, what I offer here is not a review of *IM*. Rather, it is a discussion of aspects of Stump's Paradigm Function Morphology (PFM) in the version conveniently encapsulated in *IM*, concentrating on certain areas where Stump's views about morphology impinge on or contrast with my own. Thus I will disappoint any reader hoping for a summary of the book's contents and a discussion of each topic proportionate to the space that Stump devotes to it. For example, I will say nothing about headedness and syncretism, which occupy chapters 4 and 7. Instead, I will concentrate on certain issues whose treatment by Stump seems to me problematic, all involving his view of inflectional exponence and the relationship between affixal and non-affixal inflection.

I will begin with a short survey of morphological theory in the last few decades. Putting Stump's approach into a historical context in this way may help to shed light on points of similarity and difference with other recent approaches.

2. BACKGROUND: MORPHOLOGICAL THEORY IN RELATION TO SYNTACTIC THEORY

The re-establishment of morphology as a subfield of grammar distinct from both phonology and syntax began with publications such as Aronoff's *Word Formation in Generative Grammar* (1976) and Jackendoff's pioneering study of

morphological relationships within the lexicon (1975), both of them from scholars who were at or had recently left MIT. One might have expected, therefore, that a dominant or mainstream approach to morphology would establish itself in parallel to the MIT-influenced mainstream (or, rather, what was widely regarded as the mainstream) in syntax. The fact that that did not happen is interesting. There is no single reason for it. An important factor is a continuing belief that the morphology of any language, unlike its syntax, really can be learned solely by exposure to positive evidence and by application of inductive generalisations, so that poverty-of-stimulus arguments for innate constraints on morphology are lacking. A second factor is the tendency to view morphology as really just the syntax of elements below the level of the word. This was first illustrated in Chomsky (1957) 'affix-hopping' approach to English verb inflection, and persisted both in the exploration of 'functional heads' in the Principles and Parameters framework and in the morphological work of Selkirk (1982) and Lieber (1992). A distinctive MIT school of morphology did not arise until the early 1990s, with Distributive Morphology (DM) (Halle and Marantz 1993); and one of the central planks of DM is that morphological phenomena are 'distributed' between syntax, phonology and the lexicon, so that there is still (in their view) little that is distinctively morphological about morphology.

During the period I am describing (from the 1970s to the early 1990s), most American linguists continued to accept as central the notion 'morpheme', however defined. In Britain, however, Robins (1959) and Matthews (1972) were reviving an approach to inflectional morphology that had already been christened 'Word and Paradigm' by Hockett (1954). In this approach, morphemes were downplayed in favour of paradigmatic relationships between wordforms belonging to the same lexeme, expressing different combinations of associated grammatical properties. In the USA, this approach was taken up by Stephen R. Anderson and colleagues (Anderson 1992, Thomas-Flinders 1981), while an approach somewhat like it was being pursued by Zwicky (1985).

Work by Stump on inflectional theory began to appear in the late 1980s, and his book represents a provisional summing-up within a research program which has also given rise to an impressive quantity of articles, both previously and since. His work belongs solidly in the 'Word and Paradigm' tradition, building on ideas of Zwicky and Anderson. Stump emphasises the centrality of the paradigm as a theoretical notion, as his subtitle implies. He also denies that words have any internal hierarchical structure (apart from compounds) (*IM* 11–12). These views put him squarely at odds with Distributed Morphology, which claims that complex words consist of morphemes organised in a hierarchical structure like that of syntax, and that paradigms are mere epiphenomena, so that the notion 'paradigm' deserves no central place in morphological theory (see e.g. Bobaljik 2002).

Distributed Morphology, having been invented at MIT, is probably the morphological theory favoured by most syntacticians who work in the MIT-based

Principles-and-Parameters or Minimalist frameworks—that is, to the extent that such syntacticians take notice of morphology at all as in some degree separate from both syntax and phonology. However, morphology of Stump's kind has not been ignored entirely by syntacticians. It is popular among some practitioners of Lexical-Functional Grammar (see, e.g., Luís and Spencer 2005). Stump also sees PFM as closely compatible with work on derivational paradigms by Bauer (1997) and Booij (1996) (*IM* 252–254), and it has mutually acknowledged similarities with the Network Morphology framework developed and implemented computationally at the University of Surrey (Corbett and Fraser 1993, and subsequent work, discussed in *IM* 260–276).

In Stump's account of how his work fits into the morphological landscape, there are two other approaches that are less prominent. These are Natural Morphology (Dressler et al. 1987, Kilani-Schoch 1988, Mayerthaler 1981, Wurzel 1984;) and my own approach. Both of these, in different ways, take as their starting point the pattern of 'one form, one meaning' (OFOM). Natural Morphology sees grammar as shaped by compromises between competing principles of naturalness, both between components (as between morphology and phonology) and within components (for example, within morphology, between OFOM on the one hand and a clear indexical relationship between affixes and stems on the other (Dressler 1985)). This kind of competition recalls the competing constraints of Optimality Theory, and indeed some convergence between these two approaches has begun (Elgersma and Houseman 1999, Nessel and Enger 2002). My own approach begins with the recognition that all conceivable kinds of deviation from OFOM exist (Carstairs 1987:12–18), and seeks empirically motivated constraints on such deviation that are of a purely morphological kind (belonging to that part of UG that is responsible for 'morphology by itself', one might say) or that are rooted in developmental psycholinguistic factors. Stump's and my interest in 'morphology by itself' has however led us in rather different directions, as will become clear.

3. CORE ASSUMPTIONS OF PFM: THEORETICAL HOMOGENEITY AND THE RESTRICTION TO MORPHOSYNTAX

Stump sees inflectional morphology as being concerned with 'the nature of [the] associations between an inflected word's morphosyntactic properties and its morphology' (*IM* 1). He maps out the territory for debate by categorising approaches to inflection in terms of two intersecting parameters: 'incremental' versus 'realisational' and 'lexical' versus 'inferential' (*IM* 1–3). In incremental approaches the morphosyntactic properties of a wordform are built up gradually as morphological processes (e.g. affixation) apply to it or as properties (such as 'plural' or 'past tense') are added to it, by whatever means. By contrast, in realisational approaches, all the morphosyntactic properties associated with a lexeme in a given

context are supplied independently of their morphological realisation, and the task of morphology is to spell them out, or realise them. As regards the other parameter, lexical approaches are ones in which morphological material such as the English third person singular suffix *-s* are lexical items with lexical entries, combining with roots or stems so as to either contribute to the wordform's morphosyntactic properties (in incremental approaches) or realise those properties (in realisational approaches). By contrast, inferential approaches are ones in which morphological properties are spelled out by rules, which in turn either simultaneously add (incrementally) the properties that they spell out, or else apply (realisationally) to properties that are already specified.

Examples of all four approaches exist, thus:

- lexical–incremental: Lieber's (1992) approach to morphology as word-internal syntax
- lexical–realisational: Distributed Morphology (Halle and Marantz 1993)
- inferential–incremental: Articulated Morphology (Steele 1995)
- inferential–realisational: the approaches of Matthews (1972), Zwicky (1985) and Anderson (1992), as well as Stump's own PFM

Distributed Morphology, the most prominent current rival to Paradigm Functional Morphology, resembles it in being realisational. Perhaps unfairly, incremental approaches such as Lieber's and Steele's have not figured prominently in morphological discussions during the last 10 years, and I will not discuss them further here. In terms of Stump's fourfold distinction, then, the principal issue is whether, among realisational approaches, it is lexical or inferential ones that are superior.

Stump makes it clear that his preference for an inferential–realisational approach is empirically based. He thinks that other approaches do not make available natural ways of describing both what can and (importantly) what cannot happen in morphology. 'A theory of inflectional morphology', he says, 'must be preferred to the extent that it minimizes any dependence on theoretical distinctions which are not empirically motivated' (*IM* 9). He argues that his own approach is more restrictive than others in certain areas, notably with respect to rule competition (chapter 3) and headedness (chapter 4). It is important to emphasise this because one of Stump's reviewers (Spencer 2003:638) has argued that 'the question of restrictiveness is irrelevant' for morphology because its finite nature precludes poverty-of-stimulus arguments. Yet the unavailability of poverty-of-stimulus arguments does not entail that language faculty imposes no constraints on morphological behaviour. Whether such constraints exist or not is an empirical question. Stump appears to think that they do exist, and so do I. We both thus disagree with Spencer, who suggests that 'unattested or very rare patterns' in morphology can all be accounted for on the basis of 'other principles [lying outside morphology] (e.g. processing limitations or plausible grammaticalisation paths)'.

What distinguishes Stump's version of inferential–realisational morphology from others is the notion 'paradigm function'. A paradigm function is the relationship between, on the one hand, all the appropriate pairings of roots and morphosyntactic property combinations in a language and, on the other hand, the corresponding wordforms. It is defined in terms of three kinds of rule: rules of exponence, rules of referral and morphological metageneralisations. Rules of exponence are similar to Anderson's (1992) word-formation rules: they specify the phonological changes (affixal or non-affixal) that affect a wordform on the basis of its morphosyntactic properties. Rules of referral constitute the main mechanism for dealing with syncretism, although they have other applications too. Morphological metageneralisations cover two rather different phenomena, namely morphophonological changes in both affixes and stems such as are observed in Bulgarian verbs, and some stem distribution patterns of the kind observed in Sanskrit nouns.

In the rest of section 3, I will discuss Stump's approach to a range of analytical problems where my research interests overlap with his. My misgivings about what he says all arise from one or the other of two assumptions, or fundamental claims, that emerge from Stump's opening exposition of the issues to be explored. The first of these assumptions is that all inflectional phenomena are best described in terms of how they contribute to the realisation of morphosyntactic properties. The second assumption is that the best theory will be exclusively lexical or exclusively inferential.

The first assumption may seem self-evidently true. Surely the realisation of morphosyntactic properties is what inflectional morphology is all about, whatever one's theoretical framework! But in objecting, I emphasise the 'syntactic' in 'morphosyntactic'. Aronoff (1994) and Maiden (1992, 2005) have demonstrated the importance of 'morphomic' phenomena (patterns within inflectional morphology that have nothing to do with syntax), and Lass (1990) gives a beautiful illustration, from Afrikaans adjectives, of the bizarre functions that inflection can serve, quite unrelated to syntax. In the light of this, the domain of inflectional theory cannot be limited to 'the nature of [the] associations between an inflected word's morphosyntactic properties and its morphology' (*IM* 1). At first sight, this complaint may seem unfair, in that Stump devotes a whole chapter (chapter 6) to morphomic aspects of stem alternation. However, I will argue in section 3.2 that Stump's treatment there is less morphomic than it should be.

Stump's second assumption allows no room for the possibility that some aspects of inflection fit an inferential approach while others are best described if elements such as affixes are treated as having at least some of the characteristics of lexical items. His exclusively inferential theory goes along with an insistence that 'there is no theoretically significant difference between concatenative [i.e. affixal] and nonconcatenative inflection' (*IM* 9), and that there is no theoretically significant difference between 'properties of content' and 'properties of context', including between principal and secondary exponents in Carstairs's (1987) sense,

(*IM* 11). Stump thus assesses the evidence as counting in favour of a theory of inflection that is to this extent streamlined and homogeneous. (I emphasise ‘to this extent’ because in other respects Stump’s theory is quite heterogeneous, involving a ‘rich hierarchy’ of six rule types (*IM* 242).)

There is a widespread view among linguists that a theory that is homogeneous, so as to narrow the analytical options, always makes stronger claims than a theory that provides a wider range of analytical options. Stump nowhere espouses this view himself. However, it is possible that some readers may think that the homogeneity that I have mentioned confers conceptual advantages on PFM, independent of the empirical advantages that Stump claims. Therefore it seems worth drawing attention to an important demonstration by Dresher (1981) that homogeneity by itself does not guarantee predictive power.

From its inception, generative grammar has laid stress on the idea that a theory is bad if it is insufficiently restrictive, in the sense of providing too many ways to analyse a given body of data. Conversely, it has been widely believed that a theory has merit if it makes available only a small range of potential analyses, ideally a range of just one. Now, this analytical restrictiveness is indeed likely to be characteristic of a theory that is genuinely explanatory, in the sense that observations about what does and does not happen in seemingly independent areas emerge as consequences a single more fundamental principle. But the converse does not follow: that is, a theory may be highly restrictive in the analyses that it permits and yet have no empirical consequences at all.

Dresher takes issue with a claim made on behalf of Natural Generative Phonology (NGP) by Hooper (1976). Hooper proposes that phonological theory should constrain tightly the way in which phonological data can be analysed, as follows (1976:13):

A very strong constraint on rules would be one that does not allow abstract rules at all. It would require that all rules express transparent surface generalizations, generalizations that are true for all surface forms . . . We will call this the True Generalization Condition [TGC].

She goes on to say (1976:14) that a theory incorporating the True Generalization Condition ‘makes STRONGER claims about what is a possible natural language [than a theory without the Condition does] because it restricts considerably the number of possible grammars for any given body of data’ (Hooper’s emphasis). But Dresher shows that Hooper’s reasoning is incorrect. He says (1981:84)

To see this, consider a theory which restricts the class of grammars even further than does the TGC, say, a theory which requires that all words be stored in their phonetic representations. This theory rules out even more potentially false generalizations than NGP, and for any language it allows only one grammar—a list of surface forms in all their manifestations. But

such a theory does not make strong claims . . . about possible languages; on the contrary, it makes virtually no claims, because it embodies an extremely impoverished theory of UG [i.e. universal grammar]. It makes no predictions about likely vs. unlikely rules or languages . . . This example shows how it is possible to sharply restrict the number of grammars available while at the same time greatly reducing the explanatory power of the theory. . . . For a constraint to result in an increase of explanatory power, it must be shown that the proposed constraint sheds light on certain empirical phenomena.

This is relevant to PFM in the following way. According to PFM, no language can exhibit any important difference between affixes and non-concatenative phenomena in their behaviour or function. Any analysis that asserts such a difference is therefore excluded. Likewise, any analysis is excluded if it exploits a distinction between morphosyntactic content and context. But Drescher's argument against Hooper shows that such analytical restrictiveness does not automatically ensure that PFM makes strong empirical predictions. To assess the empirical richness of PFM, there is no substitute for working out in detail what it predicts about possible and impossible morphological phenomena (which is not always easy) and exploring whether these predictions are correct. I will argue that some of these predictions are incorrect. Moreover, because the analytical restrictions that I have mentioned are so central to PFM, they make it difficult even to formulate in its terms certain lines of inquiry that have shown themselves to be promising.

3.1. First problem: affixes, inflection classes, and allomorphic conditioning

While emphasising strongly the paradigmatic dimension in inflection, Stump has little to say about certain morphological phenomena to which the term 'paradigm' is often applied, namely inflection classes. Inflectional differences that are ascribable to the syntagmatic phonological context, ranging from common-or-garden morphophonology to what Carstairs (1988a) called 'phonologically conditioned suppletion', are handled by Stump in terms of morphological meta-generalisations. Inflectional differences that are not phonologically tractable in that way are handled in terms of 'L-indices'; for example, it is L-indices that distinguish the root *lie* belonging to the lexeme LIE_1 'recline' (past tense *lay*) and the homophonous root belonging to LIE_2 'prevaricate' (past tense *lied*) (*IM* 44). Other examples of L-indices are the features $\pm T$ and $\pm C$ by means of which Stump distinguishes the four conjugation classes of Bulgarian verbs (*IM* 34–46). But he does not suggest any restrictions on how L-indices that define conjugation and declension classes can be used. That is, he has almost nothing to say about the issue of possible constraints on the distribution of inflectional resources among inflection classes—an issue first posed in terms of 'paradigm economy' (Carstairs

1983, 1987). This is surprising, because if there is indeed a requirement that inflection class organisation should be ‘economical’ in some sense, the fact that this requirement must rely crucially on the paradigmatic dimension of morphological structure should be grist to Stump’s mill. On the other hand, if ‘economy’ constraints are illusory, it would be helpful to have this clearly demonstrated.

A possible reason for this omission suggests itself, involving the distinction between affixal and non-affixal inflection. English is not rich in inflection, but even so one may distinguish at least three or four conjugation classes for verbs. How many classes, exactly? Consider the two English verbs *give* (past participle *given*) and *spoke* (past participle *spoken*). Each clearly belongs to a different class from *bake* (past participle *baked*) and *bend* (past participle *bent*). But do they belong to different classes from each other? So far as affixal inflection goes, they are identical. Only in respect of non-affixal inflection do they differ, in that in *given* the stem vowel is the same as in the present (*give*) rather than the past (*gave*), whereas in *spoken* that pattern is reversed. Should we pay attention to the affixes alone, and so assign them to the same class, or should we pay attention to the stem differences too, and assign them to different classes? If one is not exploring restrictions on inflection class behaviour, this may seem a trivial question. But if one is exploring the possibility that inflection class organisation is as economical as possible, or is restricted in some related way (Carstairs-McCarthy 1994), then the decision makes a difference to what one will predict.

A variety of evidence has accumulated in support of a proposal that inflection class organisation is indeed constrained, so that the upper limit on the total of inflection classes that a language may in fact have has to be well below the total that it could conceivably have, given its array of inflectional resources (Carstairs 1983, 1987). The most recent version of this proposal (the No Blur Principle: Cameron-Faulkner and Carstairs-McCarthy 2000, Carstairs-McCarthy 1994) derives the relevant constraint from the Principle of Contrast proposed by the psycholinguist Eve Clark, which relates to the acquisition of vocabulary (Clark 1987, 1993:64): ‘Speakers take every difference in form to mark a difference in meaning’. It turns out that, if the information content of inflectional affixes is interpreted as potentially including precise, unequivocal information about the inflectional behaviour of the lexemes to which they attach (i.e. precise, unequivocal identification of the inflection class to which they belong), they obey a version of Clark’s Principle of Contrast neatly. A crucial point here, however, is that it is only affixes that are distinguished in this way (such as *-en* in our English verb example), not instances of extended exponence involving an affix and a stem alternant (such as *-en* plus *giv-* or *-en* plus *spok-*). Unless inflectional difference is defined for this purpose on the basis of affixes alone, inflection class behaviour emerges as much less orderly and more arbitrary. It is not that excluding non-affixal inflection for this purpose makes the No Blur Principle looser and thus harder to falsify; it merely makes it different (looser in some respects, tighter in others), and it happens to be the affixes-only version that

is better supported by the evidence (Carstairs-McCarthy 1994:757–761). This is consistent with an earlier finding in the context of exploring paradigm economy (Carstairs 1987:221–232; 1988b, Carstairs-McCarthy 1991:237–247).

It should be evident now why the No Blur Principle does not fit easily in Stump's framework. It runs counter to his view that the difference between affixal and non-affixal inflection has no theoretical importance. Also, in view of its link to lexical acquisition, the No Blur Principle may be considered to count in favour of a lexical rather than an inferential analysis for at least some inflectional affixes—that is, an analysis of them as lexical items or vocabulary items, just like *cat* or *prevaricate*, as in Lieber's (1992) approach and in Distributed Morphology. Evidence in favour of the No Blur Principle is thus to that extent evidence against PFM.

It turns out that the paradigmatic aspect of the No Blur Principle may have application in another domain superficially quite independent of inflection class organisation, namely the question of whether allomorphic conditioning (inward or outward) is constrained, and if so how (Carstairs 1987:147–206, Carstairs-McCarthy 2001, 2003). Yet even to ask this question is impossible in Stump's framework, because it presupposes the possibility of distinguishing between principal and secondary exponence, or (in Stump's terms) properties of content and properties of context. Here again, therefore, are claims which, if substantiated, count against Stump's approach. I will not rehearse here all the evidence for constraints on allomorphic conditioning, but discuss rather two of the reasons that Stump gives for rejecting any distinction between 'content' and 'context'.

Carstairs (1987:150–151) introduces a distinction between principal and secondary exponence as follows. Let us suppose that, in some inflectional paradigm, morphosyntactic properties P and Q are realised. Suppose that Q is always or usually realised unambiguously by some inflection *x*, whether or not P is present. Suppose also that P is realised by a range of inflections including *a* and *b*, but that its realisation is consistently *b* in wordforms where Q is realised by *x*. In a sense, *b* is an exponent of Q as well as of P, inasmuch as Q triggers the choice of *b* to realise P. Nevertheless, in wordforms containing *b*, Q has an exponent distinct from *b*, namely *x*, because (ex hypothesi) *x* realizes Q unambiguously, independently of the presence or absence of P. In this circumstance, we can call *x* the 'principal exponent' of Q, while *b* is merely a secondary exponent of it. Stump objects that this distinction is poorly motivated for two reasons (*IM* 156–166). Firstly, he rejects the notion of 'feature discharge' which Noyer (1992) invokes in a Distributed Morphology analysis of Tamazight Berber verb inflection and which (Noyer says) relies on the notion 'principal exponent'. Secondly, he rejects Carstairs's (1987) Peripherality Constraint, which crucially relies on the notion of 'pure sensitivity', which in turn depends on the principal-secondary distinction.

The first objection is not relevant to evaluating uses to which Carstairs puts the notion 'principal exponent', however, because Noyer's version of the notion is not the same as Carstairs's. Noyer (1992:18) says, 'I will defend the view that

each [morphosyntactic] property has a (unique) principal exponent (in any given word)'. However, Carstairs (1987:152) explicitly rejects that view. For example, compare the French wordforms (*nous*) *parlerons* 'we will speak' (future) and (*nous*) *parlerions* '(we) would speak' (conditional). No one suffix unambiguously realises either 'future' or 'conditional' here: each is realised only by the combination of *-er-* and the person-number suffixes *-ons* and *-ions*, of which the first is shared with the present tense and the second with the imperfect. Thus neither 'future' nor 'conditional' has any principal exponent in these wordforms. In fact, Carstairs deliberately makes the criteria for principal exponent status quite stringent, so that any generalisations which emerge about allomorphic sensitivity or conditioning in terms of these criteria will have a solid basis.

As for Stump's second objection, it relies on a supposed counterexample to the Peripherality Constraint in Bulgarian (Stump 1997:226–231). The Peripherality Constraint incorporates a claim that the realisation of a morphosyntactic property may be sensitive inwards to another individual property, that is to an individual property with a principal exponent that is more central or closer to the root, but not outwards, that is to an individual property with a principal exponent that is more peripheral or further from the root. Stump suggests that, in Bulgarian, the realisation of 'preterite', a property shared by the imperfect and aorist tenses, violates this constraint. However, that charge is difficult to sustain, if the definition of 'principal exponent' is properly applied. Consider the following forms of the verb *IGRÁJ* 'play' (*IM* 39):

(1)	Present	Imperfect	Aorist
	1 sg <i>igráj-ə</i>	<i>igrá-e-x</i>	<i>igrá-x</i>
	2 sg <i>igrá-e-š</i>	<i>igrá-e-š-e</i>	<i>igrá</i>
	3 sg <i>igrá-e</i>	<i>igrá-e-š-e</i>	<i>igrá</i>
	1 pl <i>igrá-e-m</i>	<i>igrá-e-x-me</i>	<i>igrá-x-me</i>
	2 pl <i>igrá-e-te</i>	<i>igrá-e-x-te</i>	<i>igrá-x-te</i>
	3 pl <i>igráj-ət</i>	<i>igrá-e-x-a</i>	<i>igrá-x-a</i>

Eight of the twelve preterite forms contain a suffix *-x-*, which it seems reasonable to regard as an unambiguous realisation of 'preterite', and indeed a principal exponent of it. But what about the four other forms? Is 'preterite' realised there in a way that is sensitive to some other property or properties with a more peripheral principal exponent?

In respect of the second and third singular aorist forms, the answer seems clearly to be no. Because these forms contain no overt realisation of person, number or tense, the question of whether one property's realisation is more peripheral than another's does not arise. Even if we interpret these two forms as exhibiting 'significative absence', due to the override of a default rule (such as, informally, 'Rewrite *igrá* as *igrá-e*') by an identity rule (informally, 'Rewrite *igrá* as *igrá*'), Stump's argument is not helped; an identity rule does not locate a

realization in a linear string at all, so it clearly does not locate a realisation in any position relative to other material in the string. That leaves for consideration the form *igrá-e-š-e*, shared by the second and third singular imperfect. Let us assume, as Stump does (IM 46), that the second person form is derived from the third person form by a rule of referral. It can thus be ignored for our present purposes, and we can focus solely on the third singular form. The Peripherality Constraint is violated if (a) the final *-e* is a principal exponent of ‘third singular’ and (b) *-š-*, which replaces *-x-* in this form only, is a realisation of ‘preterite’. Let us assume that (b) is correct, that is that one can legitimately segment *igrá-e-š-e* into a stem plus three suffixes, as Stump’s hyphens imply. Nevertheless, (a) is doubtful. If we compare the third singular present and imperfect forms of IGRÁJ only, it does indeed seem reasonable to say that the same suffix *-e* appears in both forms, and that it is a principal exponent of ‘third singular’, independent of tense. However, that analysis seems less attractive when we compare the corresponding forms in the verbs KRAD ‘steal’, KOVA ‘forge’ and DÁVA ‘give’ (IM 39):

(2)		IGRÁJ	KRAD	KOVA	DÁVA
	Present	igrá-e	krad-é	kov-é	dáva
	Imperfect	igrá-e-š-e	krad-é-š-e	kov-é-š-e	dáva-š-e

All these forms have in the third singular imperfect a final unstressed *-e* which can reasonably be equated with that of *igrá-e-š-e*. However, in the present tense the picture is different: unlike *igrá-e*, none of these other three verbs has a final unstressed *-e*. In these verbs, therefore, it is not the case that ‘third singular’ is unambiguously realised by unstressed *-e*, independently of whether ‘imperfect’ is present, so it is not the case that the Peripherality Constraint is violated. Moreover, the fact that *igrá-e-š-e* parallels in shape the corresponding forms in the other verbs whereas *igrá-e* does not casts doubt on whether the final *-e* should be analysed as the same affix in these two forms; and, if it is not, then the Peripherality Constraint is not violated in IGRÁJ either.

What I have just said assumes that there is no relevant level of analysis at which *krad-é*, *kov-é* and *dáva* in (2) contain a final suffix *-e* realizing ‘third singular’. Stump would deny that assumption. He presents the Bulgarian verb forms in two versions, one showing them as they appear on the surface (IM 39) and another (IM 40) showing them in a kind of underlying representation from which the surface forms are derived through a variety of morphophonological processes (IM 48–49) based on Scatton’s (1984) analysis of Bulgarian. In this underlying representation, the crucial *-e* suffix appears. Under this analysis, the *-e* remains in the surface forms *igrá-e*, *krad-é* and *kov-é*, receiving stress in the latter two examples by a default stress rule applying to verbs that are analysed as having no lexical stress, and is deleted immediately following the stem vowel *-a* in *dáva*. Why, then, is this *-e* not deleted in *igrá-e* also? Scatton’s and Stump’s answer is that the relevant stem form is underlyingly *igráj-*, not *igrá-*, so the

suffix *-e* does not follow a vowel at the relevant level of analysis. However, it is not that the stem form *igrá-* never occurs; in fact, for aorist forms of IGRÁJ to come out right, such as the third singular *igrá*, it is necessary to assume in that tense *igrá-*, not *igráj*, as underlying.

To assess in detail now this way of analysing Bulgarian verb forms would take us too far afield. However, it seems fair to say that Stump's case against the Peripherality Constraint would be stronger if the main counterexample that he has advanced so far worked clearly on the basis of surface forms, without having to rely on so many morphophonological assumptions. In any case, this discussion is to some degree anachronistic, inasmuch as the Peripherality Constraint has now been superseded by an Ancestry Constraint, which (among other things) imposes tighter constraints on inward sensitivity (Carstairs-McCarthy 2001, 2003). But that does not affect the currently important point. Any evidence in favour of constraints of the kind discussed here on inflection class organisation or on allomorphic conditioning must count against any theory whose central tenets include ones that encourage us to expect the non-existence of such constraints. Unfortunately Paradigm Function Morphology is such a theory.

3.2. *Second problem: stem distribution and morphomic inflection*

Some of the most perceptive and enjoyable sections of *IM*, in my view, are in chapter 6, where Stump discusses certain patterns of stem distribution in Sanskrit. In that language, many consonant-final nouns, adjectives and participles have a special form of the stem, traditionally called 'Strong', which is used in the masculine nominative and accusative singular and dual, the masculine nominative plural, and the neuter nominative and accusative plural. This assortment of morphosyntactic property combinations shows that the Strong stem does not correlate neatly with any morphosyntactic feature value. It also does not correlate with any phonological characteristics of the accompanying affixes. For example, as Stump points out, nominative and accusative plural masculine forms both have the same suffix *-as*, yet only the former has a Strong stem. On the other hand, syntagmatic phonological factors do come into play in the choice between the 'Middle' and 'Weakest' stems that is exhibited in some lexemes. Here (with marginal exceptions) the Middle stem is found before consonant-initial suffixes and the Weakest stem before vowel-initial ones.

Stump uses the term 'stem indexing' for the kind of distinction that Sanskrit exhibits between Strong and Weak (or Strong, Middle and Weakest) stems. He demonstrates convincingly that, at least for some lexemes, the formation of a stem with a given index cannot always be predicted from that of any counterpart with a different index. Indeed, even if one knows all the stem shapes available for a given lexeme, one cannot always predict what index each will have. Thus Stump

puts forward what he calls the Indexing Autonomy Hypothesis: stem indexing is in principle independent of stem formation. This stem behaviour in Sanskrit does not serve either syntactic or phonological needs, yet it is robust and productive. Hence, as Stump says, it is morphomic, in Aronoff's (1994) terms: an instance of 'morphology by itself'.

If one stem were associated consistently with a given morphosyntactic property (say, 'plural' or 'accusative'), the distribution of the various stems could be handled in Stump's framework by a single straightforward rule of exponence, rather like the rule which (unless blocked by the Panini Principle) suffixes *-i* to realise 'locative singular' in nominals (i.e. nouns or adjectives) (rule (7g), *IM* 181). But precisely what is interesting about Sanskrit's stem indexing is that, although the stem distribution pattern is consistent across a wide range of lexemes, the morphosyntactic contexts of the stems are not expressible in terms of natural classes of feature values. In this respect, it is like the stem distribution pattern in Italian verbs discussed by Maiden (1992): in the present tense, a fair number of common verbs exhibit a distinct stem form in the first singular and third plural of the indicative, plus all the subjunctive except the first and second plural. This distinct stem form usually (though not always) terminates in a velar consonant, so we can call it for convenience the velar stem. Let us introduce the term 'special stem' for the Sanskrit Strong stem and the Italian velar stem, and the term 'triggering cells' for the uniform set of cells where, among the lexemes in question, the special stem is used. The information content associated with the choice of the special stem (its 'meaning', in a broad sense) is then not something conventionally semantic, involving the non-linguistic world, nor even something residing in language outside morphology (e.g. 'first singular' or 'masculine nominative or accusative'), but something purely intramorphological and in that sense morphomic: 'This stem is the same as is used in all the triggering cells in the paradigm of this lexeme.' What one looks for in a satisfactory theory of inflection, therefore, is an analysis which predicts the impossibility of a Pseudo-Sanskrit or a Pseudo-Italian where this consistency of distribution is lacking, i.e. where, for multi-stemmed lexemes of the relevant kind, there is no uniform set of triggering cells but a variety of distinct but partially overlapping sets of cells where the special stem is used, so that for many or all lexemes the stem distribution pattern has to be lexically specified.

Unfortunately, Stump's framework does not supply the kind of analysis that we are seeking. Stump does indeed recognise that morphomic factors are at work, but crucially not when it comes to stating the difference in information content between the stem alternants. To justify this claim, it will help if I define two different ways of analysing the relationship between the special stem alternants and the paradigmatic cells where they occur. I will call these the morphosyntactic realisation analysis and the distributional uniformity analysis—cumbersome labels, perhaps, but (I hope) apt:

Morphosyntactic realisation analysis:

Individual cells (or morphosyntactically natural classes of them, such as ‘direct case’ embracing ‘nominative and accusative’, or all the singular cells in a given tense and mood) are associated with the special stem directly by means of realisation rules, similar in form and function to realisation rules which add affixes. For example, one realisation rule of stem selection for Sanskrit will say something like ‘For direct-case plural neuter, choose the Strong stem’; a rule for Italian will say ‘For singular present subjunctive, choose the velar stem’. Through such rules, the information content of the special stem emerges as a disjunction of morphosyntactic property combinations: in Sanskrit ‘direct case masculine singular OR dual, OR nominative masculine plural, OR direct case neuter plural’, and in Italian ‘indicative first singular OR third plural, OR subjunctive singular OR third plural’.

Distributional uniformity analysis:

Individual cells are not associated with the special stem directly. Rather, the set of triggering cells is associated en bloc with the special stem for all multi-stemmed lexemes of the relevant kind. The information content of the special stem in any wordform emerges then as a fact about its own distribution within the paradigm of the lexeme to which the wordform belongs: ‘special stem in all the other triggering cells too’.

At first sight it may seem unclear what the fuss is about here. One way or another, special stems have to be associated with a set of morphosyntactic property combinations. The arbitrariness of this set is inescapable. What does it matter whether the special stem is associated with the members of this set directly one by one, or indirectly and en bloc?

The answer lies in the implications for the possibility that different lexemes of the relevant kind may make different but overlapping choices as to which morphosyntactic property combinations the special stem will be used with. Let us imagine a Pseudo-Sanskrit where multi-stemmed consonant-final nominals exhibit two different stem distribution patterns: all such nominals use the special stem in the masculine nominative and accusative dual, but some use it also in the masculine nominative and accusative singular (let us call that ‘distribution A’), while others use it also in the masculine nominative plural and the neuter nominative and accusative plural (‘distribution B’). Under the morphosyntactic realisation analysis, it would be easy to accommodate this situation. It would be a matter of recognising two distinct but partially overlapping sets of stem-selecting realisation rules, with lexical specification to indicate which rules are

appropriate for any given nominal. Under the distributional uniformity analysis, however, such an option is excluded by the fact that the relationship between stems and morphosyntactic properties is indirect. At the same time, the expression ‘all the other triggering cells’ no longer stands for anything, because there is no longer any such uniform set of cells. If a Pseudo-Sanskrit lexeme has a special stem in the masculine nominative and accusative dual, it will have a special stem somewhere else too—but one cannot tell if the other cells in question will be those of distribution A or distribution B. The distributional uniformity analysis therefore implies that Pseudo-Sanskrit as defined here, and languages that resemble it in appropriate respects, could not exist, whereas the morphosyntactic realisation analysis makes no such empirical prediction.

Now, Stump’s analysis for Sanskrit is a morphosyntactic realisation analysis. He treats stem selection rules as one kind of realisation rule for morphosyntactic properties, alongside rules of exponence (*IM* 172). For stem distribution in the relevant Sanskrit nominals, Stump posits four stem selection rules ((3a–d) at *IM* 179), two of which stipulate morphosyntactic contexts where Strong stems occur (in certain masculine and neuter cells respectively), one of which stipulates a more specific masculine context where Middle stems occur (in the accusative plural), and one of which stipulates the Middle stem as the default that will occur unless otherwise specified, by the Panini Principle.

A relatively minor problem is the fact that Stump’s formulation suggests that actual Sanskrit, with only two stems mentioned in these four rules, is scarcely more complex than a second imaginary Pseudo-Sanskrit with four stems, one for each rule. More importantly, however, there is nothing in Stump’s framework to guarantee the consistency of distribution that I have emphasised. Each of Stump’s stem-selection rules for Sanskrit contains a restriction (an L-index, in his terminology) to the effect that it applies to consonant-stem nominals. But L-indices are also used to restrict rules to particular declension or conjugation classes, as in Bulgarian verbs (*IM* 44–46). Such indices could thus be used to restrict stem-selection rules to arbitrary subclasses of consonant-stem nominals, as in the first Pseudo-Sanskrit described above. So, although Stump successfully disentangles stem formation from stem indexing, the way in which he associates a stem’s index with its distribution seems likely to be empirically too permissive. Only a distributional uniformity analysis exploits appropriately and to the full the morphomic character of stem distribution patterns.

There is much more to be said about stem distribution patterns and possible restrictions on them (see, e.g. Carstairs-McCarthy 2002, Enger 2004). I will limit myself here to mentioning two unresolved issues, indicating why I think that the relevant evidence so far does not support Stump’s position and may support mine. The first issue concerns the domains within which uniformity of stem distribution is observed. In my discussion earlier, I was careful to talk of ‘multi-stemmed lexemes *of the relevant kind*’. What constitutes ‘the relevant kind’? If one interprets this as meaning ‘all multi-stemmed nominals (or verbs, etc., as

the case may be) in the language, without exception', then one need look no further than Sanskrit for counterevidence. Other Sanskrit nouns display stem distribution patterns quite different from the one we have been discussing—or may do so, depending on one's analysis of the boundary between stem and affix. However, these patterns involve certain nouns whose stems ending in vowels, not consonants. This phonological difference seems almost certainly relevant. It will not be surprising if phonological shape turned out to be a factor in distinguishing the domains within which particular stem distribution patterns apply. Correspondingly, it seems unlikely that this kind of difference in behaviour is irreconcilable with distributional uniformity analyses.

The second issue concerns the information content of stems. Under the morphosyntactic realisation analysis for stem distribution, I pointed out that stems emerge with an information content of the form 'property-set A OR property-set B OR ...'. This may remind readers of the No Blur Principle, which in effect rules out inflectional affixes whose information content contains 'inflection class A OR class B OR ...'. The common factor is a disjunction of incompatible or at least potentially incompatible items. That in turn recalls the Exclusive Disjunction Bar (Carstairs-McCarthy 1998:7), according to which (it is suggested) the meaning of an inflectional affix cannot contain a disjunction of incompatible properties, such as 'Past OR Future', 'Subjunctive OR Optative', 'Instrumental OR Dative'. Let us explore the idea that what is important in determining whether a given stem distribution pattern can occur or not is whether it is describable without such a disjunction. One way to avoid such a disjunction is for the pattern to be uniform in all lexemes of the relevant kind, with a single set of triggering cells. However, this is not the only way. Suppose that distribution patterns are nested in the following fashion, where p, q, r and so on represent cells in a paradigm and where some lexemes use the special stem in all six cells enclosed within the brackets labelled 1, others use it in only the four cells enclosed within the brackets labelled 2, and yet others use it in only the two cells enclosed within the brackets labelled 3:

$$(3) \quad p \ q \ [{}_1 r \ s \ [{}_2 t \ u \ [{}_3 v \ w]_3]_2]_1 \ x \ y \ z$$

The information content of a special stem in cell v can now be interpreted as 'special stem also in cell w '. The information content of a special stem in cell u is more substantial: 'special stem also in cells t, v and w '. The information content of a special stem in cell r or s is clearly the most substantial of all: 'special stem in all the cells enclosed within any of the brackets'. Uniform distribution is thus equivalent to a pattern in which only the outermost brackets exist. By contrast, the pattern of partially overlapping distributions for the special stem that I posited for Pseudo-Sanskrit earlier, represented schematically in (4), crucially involves disjunctive information content:

$$(4) \quad p \ q \ [{}_1 r \ s \ ({}_2 t \ u)_1 \ v \ w]_2 \ x \ y \ z$$

Here, the information content of a special stem in cell *t* is ‘special stem in cell *u* and in cells *r* and *s* OR cells *v* and *w*’. Thus, if avoidance of disjunction is an acceptable alternative to uniformity, we should indeed fail to observe a pattern of stem distribution such as is represented in (4), but we should expect to observe patterns of the nested kind represented in (3).

There is some evidence that this is correct. The velar-stem pattern for Italian mentioned earlier is a subset of a pattern in which the set of cells with the special stem includes also the second and third singular indicative. This pattern is observed in the numerous verbs with the stem extension *-isc-* (e.g. *finire* ‘finish’, indicative third singular *finisce*, second plural *finite*, by contrast with *partire* ‘leave’, whose corresponding forms are *parte* and *partite*), as well as in *udire* ‘hear’ and *uscire* ‘go out’ (special stems *od-* and *esc-* respectively). Polish masculine nouns exhibit a similar nested pattern, with some nouns having a special stem in both the locative and the vocative singular, while others have it in the vocative alone (Cameron-Faulkner and Carstairs-McCarthy 2000). In German verbs, the number of ‘special’ stem alternants is not limited to two, but the availability of some special stem or other for any given cell in the verbal paradigm conforms to a multi-stage implicational ‘paradigm structure condition’ (Bittner 1985, Carstairs-McCarthy 1991), equivalent to a complex nested pattern on the lines of (3). So, although much work remains to be done, what has been done so far seems promising enough to warrant preference for a version of distributional uniformity analysis over the morphosyntactic realisation analysis that Stump adopts for multi-stemmed lexemes.

There are reasons why it is perhaps not surprising that Stump chooses a morphosyntactic realisation analysis. In Sanskrit, consonant-stem nominals are accompanied by a single set of inflectional affixes, uncontroversially realising properties of number, gender and case. For Stump to recognise the stems to which these affixes are attached as expressing a kind of information or ‘meaning’ quite distinct from morphosyntactic exponence would compromise his denial of any significant difference between concatenative and non-concatenative inflection. It would also cast doubt on his emphasis on morphosyntax (categories and properties such as ‘tense’ and ‘past’) as the only kind of content that inflectional morphology expresses. Evidence of the kind discussed in this section provides part of my reason for disagreeing with Stump on both these issues.

3.3. Third problem: rule blocks and rule competition in Georgian

Anderson (1982, 1984, 1986, 1992) has provoked debate about two intriguingly complex sets of data, involving verb inflection in Georgian and Potawatomi (Halle and Marantz 1993, Jensen and Stong-Jensen 1984). Stump treats both of these, but I will concentrate on Georgian, discussing his attempt to establish a framework for morphological rule interaction that is more restrictive than Anderson’s. This case study illustrates neatly, I suggest, the flavour of Stump’s

approach to theory-construction in general, with both its strengths and its weaknesses. The strength lies in the rigour with which individual concepts and hypotheses are defined and the ingenuity with which their consequences are explored, including consequences that are by no means immediately obvious. The weakness lies in the way in which proposals made in different parts of the book interact. In this instance, a strong claim advanced in chapter 3 is undermined by an analytical possibility introduced in chapter 2 to deal with an apparently unrelated phenomenon in Bulgarian, and also by Stump's rejection in chapter 1 of any distinction between properties of content and properties of context.

Let us imagine a language (like many Indo-European languages) in which verbs inflect for the morphosyntactic categories of voice, mood, tense, person and number. Is it likely that the realisation of properties from each of these categories will be scattered around verbal wordforms inconsistently? That is, is it likely that 'past' will be realised by an outermost prefix in some forms, by stem vowel change in other forms, and by an inner suffix in yet others, with similar randomness in the realisation of all other properties? Surely not. The morphological realisation of properties is nearly always more consistent than that. Bybee (1985) has suggested that such consistencies follow from the way in which 'relevance', in her sense, interacts with diachronic processes of grammaticalisation. Be that as it may, one can often identify affixal positions in relation to the root such that all properties belonging to a given category are realised in one position, whether or not they are realised elsewhere too. For example, in German verbs, person and number are cumulated and always realised in the last suffixal position, so that the second singular suffix *-st* is consistently last in *(du) lobst* 'you praise', *lobtest* 'praised', *schläfst* 'sleep', *schliefst* 'slept', *schlafest* 'sleep (present subjunctive)' and *schliefe* 'should sleep (past subjunctive)', even though one might argue that second singular is also partially realised in the choice of *ä* [ɛ] over *a* [a] in the stem of *schläfst*, because of the contrast with *a* in forms such as the second plural present *schlaft* '(you plural) sleep'. In a good theory of inflectional morphology, this consistency of positioning should emerge as natural and expected, one feels. Perhaps inflectional rules are organised in rule blocks such that all the rules that account for affixes occupying the same position in relation to the stem (affixes that belong to the same 'position class', in Stump's terminology) are grouped together in the same rule block. Clearly the rules within a rule block apply disjunctively, in the sense that no more than one of them can apply in the formation of any one wordform.

Quite independently of position classes, some rules are mutually incompatible in that they realise rival properties within the same morphosyntactic category (or rival values of the same feature). A verb form cannot realise both 'past tense' and 'future tense' simultaneously, for example. Such rules as these, too, are disjunctive in their application. So what is the relationship between the two motivations for disjunctivity? Although they are logically independent, it would clearly be highly significant if they coincided in their effect: that is, if all the rules in every position-class block turned out to realize properties belonging to just one

category (or one set of cumulated categories, such as person–number or case–number), and all the rules relating to any one category turned out to belong to just one position-class block. Such a finding would allow one to bring the notions ‘position class’ and ‘rule block’ into neat alignment. So is this how things are? The answer one has to give is frustratingly equivocal. In many languages, there is a pretty close correspondence between the two kinds of disjunctivity. For example, in Swahili (discussed by Stump in *IM* chapter 5), one can identify at least three prefixal position classes for verbs, in which are located respectively realisations for negation, person–number–class of the subject, and tense. On the other hand, there are many examples where things are less neat and tidy. In the face of this, what lesser degree of constraint can one impose on rule block organisation?

Stump suggests an answer (*IM* 20–25) that is intriguing and that looks at first sight to have tight empirical consequences. He proposes a Paninian Determinism Hypothesis, according to which, for any two rules in a rule block, one or other of two conditions must be met: either (a) there is no input to which both can apply, so that their order of application does not matter (and they can perhaps be thought of as applying simultaneously), or (b) one rule is more specific than the other, so that their order of application is determined by the well-known Panini Principle (the rule that is narrower in scope taking precedence). The empirical consequences are not as substantial as they at first seem, however, for reasons that I will explain.

Georgian has a complex verb system involving person agreement with both the subject and the object. In Table 1 are set out the present indicative forms of the verb *xedav* ‘see’, indicating how person agreement is marked (except where subject and object are the same person, requiring special reflexive object forms that are always grammatically third person). As will be seen, there is a prefix position and a suffix position. This suggests that it would make sense to analyse Georgian in terms of a prefix rule block and a suffix rule block. That is what Anderson has done. However, it is not the case that prefixes all mark agreement with subjects and suffixes all mark agreement with objects, or vice versa; rather, markers of both kinds of agreement appear in both positions. For example, in Anderson’s prefix rule block, there is a rule that prefixes *g-* to realize ‘second singular object’, and one that prefixes *v-* to realize ‘first person subject (singular or plural)’. We can represent these rules for our purposes as follows:

(5) 2nd singular object → *g-*

(6) 1st subject → *v-*

But what happens in a verb form meaning ‘I see you’, to which both rules would seem applicable? They cannot both apply, since they are competing for one prefixal position. In fact it is rule (5) which applies, yielding a form *gxedav*, with no overt marking of first singular subject (despite which, as inspection of Table 1 will confirm, *gxedav* turns out to be unambiguous). But there seems to be no general

Table 1. Subject and object agreement in the present tense of Georgian XEDAV 'see'

	Singular objects		
	1 sg	2 sg	3 sg
Subject:			
1 sg	—	g-xedav	v-xedav
2 sg	m-xedav	—	xedav
3 sg	m-xedav-s	g-xedav-s	xedav-s
1 pl	—	g-xedav-t	v-xedav-t
2 pl	m-xedav-t	—	xedav-t
3 pl	m-xedav-en	g-xedav-en	xedav-en
	Plural objects		
	1 pl	2 pl	3 pl
Subject:			
1 sg	—	g-xedav-t	v-xedav
2 sg	gv-xedav	—	xedav
3 sg	gv-xedav-s	g-xedav-t	xedav-s
1 pl	—	g-xedav-t	v-xedav-t
2 pl	gv-xedav-t	—	xedav-t
3 pl	gv-xedav-en	g-xedav-en	xedav-en

principle, relating either to the rules' substance or their form, that predicts this. Anderson therefore appeals to a principle of 'stipulated disjunctivity': when two rules within a rule block compete to apply to one form, and when neither is more specific than the other so as to take precedence by virtue of the Panini Principle, their order can be stipulated in the grammar. It is simply a brute fact about Georgian that these rules are ordered in this way, so that 'I see you' is expressed as *gxedav* rather than *vxedav*.

Stipulated rule ordering is clearly a powerful device. It would be desirable if it could be dispensed with. Carmack (1997) argues that it can be dispensed with, at least in the Georgian case, by appeal to what he calls 'analogical blocking'. In some tenses for some verbs, as in *da-g-malvi-var* 'I hid from you', the first singular subject is represented by a suffix *-var*. (A preverb *da-* precedes the person-number prefix *g-*.) If in such a form the prefix *v-* were to appear rather than *g-*, so as to yield **da-v-malvi-var*, the first person would be realised twice (by both *v-* and *-var*) and the second person object would not be realised at all. It therefore does not seem surprising that, on the basis of principles elaborated by Carmack, *g-* beats *v-* in this form. And it is by analogy with such forms that *g-* blocks *v-* also in forms such as *g-xedav* 'I see you' (Carmack argues).

Page length restrictions imposed by Cambridge University Press prevented Stump from discussing Carmack's explanation in *IM*. In any case, he prefers an explanation based on his Paninian Determinism Hypothesis. But how can the Paninian Determinism Hypothesis be reconciled with the fact that both of rules (5) and (6) can apply to an input meaning 'I see you', and neither is more specific than the other? Stump's answer hinges on the distinction between what he calls 'unexpanded' and 'expanded' modes for the application of realization rules. The definition of this distinction that Stump offers (*IM* 72–73) is complex and technical, but it boils down to this: a rule *R* applying in expanded mode to a lexeme of class *C* in language *L* is a rule schema comprising all rules that realise not only the feature values *V* that are specified in *R* but also all other feature values (from those available for *C* in *L*) that are compatible with *V*. Thus the expanded version of rule (5) is a rule schema whose constituent rules will also spell out the following combinations:

- (7) 2nd singular object, first subject → *g*-
- (8) 2nd singular object, 3rd singular subject → *g*-
- (9) 2nd singular object, 3rd plural subject → *g*-

(Recall that this paradigm does not handle reflexive forms, so the expanded version of (5) contains no rules for second person subjects.) Now, by the Paninian Determinism Hypothesis, rule (7), yielding *g*- in the prefix position, overrides rule (6), which would have yielded *v*-. The effect of Anderson's stipulated disjunctive ordering is therefore achieved, simply by applying rule (5) in the expanded mode.

Both more specific and more general questions are likely to occur to readers at this point. A more specific question relates to the (so to speak) superfluous rules (8) and (9), which are conjured into existence by the expanded application of rule (5). Won't they cause trouble by overriding in unwanted fashion the rules for realising third person subjects? It so happens that they will not, because third person subject agreement features are spelled out in the suffix rule block rather than the prefix rule block, and Stump's definition of Paninian determinism does not allow a rule in one block to override a rule in another block. But a more fundamental question still remains: is Stump's approach really more restrictive than Anderson's? At first sight, we are merely replacing a stipulated rule ordering with a stipulated mode of rule application. Stump's answer is that his approach is indeed more restrictive than Anderson's, because, within any rule block, two competing rules that are indexed for the same inflection class (i.e. that potentially apply to lexemes within the same inflection class) cannot both apply in expanded mode (*IM* 72–75). Stump's explanation of why this is so is rather condensed, so I will illustrate his account with a hypothetical example of my own. This illustration will raise doubts about whether it is really desirable for the theory of realisation

rules to be as restrictive as Stump claims. In any case, ironically, two other aspects of Stump's framework undermine this restrictiveness by supplying loopholes by means of which Andersonian freedom of stipulation can be mimicked.

In principle, Anderson's framework allows for stipulated ordering to apply to an unlimited succession of rules: rule A may be stipulated to apply before rule B, rule B before rule C, and so on. At first sight, the same effect might be achievable in Stump's framework by allowing more than one rule in a rule block to apply in expanded mode. But in fact this would not work. To see why, consider the following hypothetical pattern of inflectional behaviour. Let us imagine a lexeme class in some language L which inflects for categories F, G and H, such that F has two values (or contains two properties) α and β , G has two values γ and δ , and H has two values ε and ζ . Following Stump, I will use the notation 'F: α ' to mean 'the value α of the category F'. Assume that these categories are semantically and syntactically independent in the sense that no combination of values for them yields a morphosyntactic contradiction. Now suppose that these features compete for realisation in such a way that the three realisation rules (10)–(12) must, in Anderson's terms, be stipulated to apply in this order:

(10) F: α \rightarrow p

(11) G: γ \rightarrow q

(12) H: ε \rightarrow r

(The letters p , q and r stand for phonologically distinct realisations, but their phonological content does not matter for present purposes.) According to Stump, the apparent effect of (10) overriding (11) (so that a wordform with the morphosyntactic specification {F: α , G: γ } would be marked p , not q) must be due not to stipulated ordering but to the application of (10) in expanded mode, as in (13):

(13) Expansion of (10):

(i) {F: α , G: γ , H: ε } \rightarrow p

(ii) {F: α , G: γ , H: ζ } \rightarrow p

(iii) {F: α , G: δ , H: ε } \rightarrow p

(iv) {F: α , G: δ , H: ζ } \rightarrow p

(13i) and (13ii) are more specific than (11), so, to realise a set of morphologically feature values containing {F: α , G: γ }, Paninian determinism will ensure that one of them applies rather than (11), yielding the realisation p rather than q .

Likewise, the apparent effect of (11) overriding (12) (so that a wordform with the morphosyntactic specification {F: β , G: γ , H: ε } would be marked q , not r) must be due not to stipulated ordering but to the application of (11) in expanded mode, as in (14):

- (14) Expansion of (11):
 (i) $\{F:\alpha, G:\gamma, H:\varepsilon\} \rightarrow p$
 (ii) $\{F:\alpha, G:\gamma, H:\zeta\} \rightarrow q$
 (iii) $\{F:\beta, G:\gamma, H:\varepsilon\} \rightarrow q$
 (iv) $\{F:\beta, G:\gamma, H:\zeta\} \rightarrow q$

But rule (14i) is incompatible with rule (13i). The Paninian Determinism Hypothesis therefore predicts that no rule block can contain both rules. The same applies to rules (14ii) and (13ii). And because the categories and values in our illustration are purely schematic and empty of content, what has been shown is that such a contradiction is bound to arise whenever, within a rule block, more than one competing rule indexed for the same inflection class applies in expanded mode. Therefore the Paninian Determinism Hypothesis requires that no more than one such rule per block should apply in the expanded mode, and hence limits to two the number of rules per block (and applicable to the same inflection class) which may look as if they need to apply in a stipulated order.

If this result is correct, it emerges as an unexpected consequence of the Paninian Determinism Hypothesis, and therefore constitutes significant evidence in favour of that Hypothesis. But is this result correct? To investigate that, we will need to supply content for our schematic categories and their values. Let us assume that their content is as follows, applying to verbs:

(15)	F	FINITE	G	TENSE	H	PERSON
	α	no	γ	past	ε	1st
	β	yes	δ	present	ζ	2nd

The schematic disjunctively ordered rules at (10)–(12) now emerge as (16)–(18):

- (16) FINITE:no $\rightarrow p$
 (17) TENSE:past $\rightarrow q$
 (18) PERSON:1st $\rightarrow r$

The effect of this disjunction is to say that only on finite verbs is realisation of the past tense (by q) possible, and only on finite verbs in the present tense is realisation of the first person (by r) possible. If Stump is right in outlawing disjunctions that go beyond two rules, the situation just described should be impossible or at least implausible. Yet it is by no means implausible. It is in fact very close to the actual state of affairs in Russian, and can be brought even closer if we insert two further rules, disjunctively ordered after (17) just as (18) is:

- (19) PERSON:2nd $\rightarrow s$
 (20) PERSON:3rd $\rightarrow t$

In Russian, just as in this hypothetical language, person agreement is found only in the present tense, not the past, and there is no distinction of tense in one of the non-finite forms, namely the infinitive. In Russian, too, an analysis ascribing all these rules to one rule block seems perfectly sensible: they all add a suffix that immediately follows the theme vowel, if any. Yet this is an analysis that Stump's framework seems to disallow. Perhaps the extra room for manoeuvre provided by Anderson's framework is an advantage after all.

Does Stump's framework then implicitly claim that Russian is an impossible language? Not quite, because Stump's discussion of the verbal inflection of Bulgarian in chapter 2 opens a loophole. Bulgarian verbal inflection is not the same as Russian: the tense system is richer, with person agreement in all finite forms. However, as in Russian, not all features are freely combinable. For example, only finite forms can be specified for mood, only participles and indicative forms can be specified for tense, and only participles can be specified for gender agreement (*IM* 41–42). Stump therefore proposes that the morphology of Bulgarian contains a set of morphosyntactic 'property co-occurrence restrictions' that discriminate between well-formed sets of properties (i.e. well-formed combinations of values for mood, tense, etc.) and ill-formed sets. I suspect that Stump would invoke precisely such property co-occurrence restrictions to account for the Russian-like state of affairs just mentioned. The restrictions would look like this, in Stump's notation (*IM* 42):

- (21) a. For any permissible α , σ is an extension of $\{\text{TNS}:\alpha\}$ iff σ is an extension of $\{\text{FINITE}:\text{yes}\}$.
- b. For any permissible β and γ , σ is an extension of $\{\text{PER}:\beta, \text{NUM}:\gamma\}$ iff σ is an extension of $\{\text{TNS}:\text{present}\}$.

So, given that the Russian-like data at (15)–(21) merely instantiate a general schema for multi-stage stipulated disjunctivity in rule ordering, what emerges is that Stump's property co-occurrence restrictions can be used as a loophole to mimic almost any rule ordering (within a single rule block) that is of the kind purportedly excluded by the Paninian Determinism Hypothesis. (I say 'Russian-like data' rather than 'Russian data' because actual Russian does distinguish tense in some non-finite forms, namely participles; but that does not affect my argument. I say 'almost any ordering' rather than 'any ordering' because of the unlikely possibility that the property co-occurrence restrictions are restricted to one inflection class, so that they are not available to mimic blocking for lexemes in another inflection class.)

Property co-occurrence restrictions are not the only loophole that Stump's own theory provides to evade the effects of expanded-mode application and the Paninian Determinism Hypothesis. Consider rule (17), which realises 'past' in a hypothetical language where tense distinctions are restricted to finite verb

forms, that is (in terms of the notation in (15)) to verb forms with the property ‘yes’ in the category FINITE. This property ‘yes’ thus provides the context for rule (17) to apply. Yet, as noted in section 3, a central claim of PFM is that there is no motivation for a distinction between properties of content and properties of context. If that is so, the q that realises ‘TNS:past’ in (17) is equally a realisation of ‘FINITE:yes’. In fact, the formulation of all of rules (17)–(20) is strictly incomplete. If we insert the contextual properties among the properties of content, (17)–(20) become (22)–(25):

(22) FINITE:yes, TENSE:past $\rightarrow q$

(23) FINITE:yes, TENSE:present, PERSON:1st $\rightarrow r$

(24) FINITE:yes, TENSE:present, PERSON:2nd $\rightarrow s$

(25) FINITE:yes, TENSE:present, PERSON:3rd $\rightarrow t$

Because of the incompatibilities in their values for TENSE and PERSON, these rules along with (16) do not compete to apply to the same input, so no ordering problem arises. But, again, (16), (22) and (23) merely instantiate a general schema for multi-stage stipulated disjunctivity in rule ordering. Therefore, because of Stump’s decision not to distinguish between content and context, empirical consequences of the Paninian Determinism Hypothesis are again subverted. Recall that this Hypothesis was claimed to prevent what looks like two-stage stipulated ordering for competing rules such as are represented schematically in (10)–(12):

(10) $F:\alpha \rightarrow p$

(11) $G:\gamma \rightarrow q$

(12) $H:\varepsilon \rightarrow r$

If we insert contextual properties among properties of content in (11) and (12) too, they become (26) and (27):

(26) $F:\beta, G:\gamma \rightarrow q$

(27) $F:\beta, G:\delta, H:\varepsilon \rightarrow r$

But (26) and (27) no longer compete either with each other or with (10), in the sense that none of them can apply to the same input. The Paninian Determinism Hypothesis can therefore no longer constrain their application.

I said in section 1 that teasing out the empirical implications of Stump's claims is often difficult. The discussion in this section illustrates that. What Stump presents as a clear empirical difference between his framework and Anderson's—a difference that allegedly counts in Stump's favour, because his framework both fits the facts and is more restrictive—disappears on closer inspection. But this closer inspection involves working out the intricate and unforeseen ways in which some of Stump's hypotheses interact.

4. CONCLUSION

Stump's *Inflectional Morphology* is commendable for two reasons: the wealth of data adduced from a wide range of languages, and the sternly rigorous formulation of the interlinked definitions and hypotheses. (This rigour does not make for easy reading, but Stump would no doubt say that it is not a researcher's job to pretend that complex issues are simple.) The case he makes for the importance of the paradigmatic dimension in inflection is strong, and on that matter I am firmly on his side, in opposition to what has at least until recently been the dominant view in Distributed Morphology. But recognising the importance of the paradigmatic dimension does not entail the kinds of analytical restrictiveness that I have criticised.

One may object to my conclusion on the following lines: 'It's all very well to criticise Stump's theoretical framework, but what theoretical framework have you got to offer instead?' Nothing so all-embracing, certainly. But that is a weak objection, because making progress in any area of inquiry (linguistic morphology, for instance) does not depend on having a grand theoretical framework covering everything in it. Grand theories have been popular in linguistics for decades (another legacy of the Chomskyan revolution). Whether that has been beneficial in syntax and phonology is an issue that I will not comment on here. In morphology, however, there is good evidence that solid progress can be made by tackling clear-cut issues independently of one another, not ignoring any connections between them or with other aspects of language, but not being in too much of a hurry to bring them all under one overarching theoretical umbrella. In most disciplines, such a moderate, middle-of-the-road approach would be uncontroversial. If among linguists it raises some eyebrows, that may be just be a sign that our training in recent decades has encouraged us to be overambitious.

NOTE

* For comments on earlier drafts I thank Steve Anderson, Peter Arkadiev, Mark Aronoff, James Blevins, Joan Bresnan, Hans-Olav Enger, Jen Hay, Martin Maiden and Rolf Noyer. None of them should be assumed to agree with anything I say. I particularly thank Greg Stump for an

extensive critique of an earlier draft. On numerous issues he put forward counterarguments or presented new data, for which I am grateful even if my evaluation of them often differs from his. Faults that remain are solely my responsibility.

REFERENCES

- Anderson, Stephen R. (1982). Where's morphology? *Linguistic Inquiry* 13, 571–612.
- Anderson, Stephen R. (1984). On representations in morphology: Case, agreement and inversion in Georgian. *Natural Language and Linguistic Theory* 2, 157–218.
- Anderson, Stephen R. (1986). Disjunctive ordering in inflectional morphology. *Natural Language and Linguistic Theory* 4, 1–31.
- Anderson, Stephen R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Aronoff, Mark (1976). *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Aronoff, Mark (1994). *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press.
- Bauer, Laurie (1997). Derivational paradigms. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1996*. Dordrecht: Kluwer, 243–256.
- Bittner, Andreas (1985). Implikative Hierarchien in der Morphologie: Das “Stark-Schwach-Kontinuum” der neuhochdeutschen Verben. *Acta Linguistica Academiae Scientiarum Hungaricae* 35, 31–42.
- Bobaljik, Jonathan David (2002). Syncretism without paradigms: Remarks on Williams 1981, 1994. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2001*. Dordrecht: Kluwer, 53–85.
- Booij, Geert (1996). Inherent versus contextual inflection and the split morphology hypothesis. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1995*. Dordrecht: Kluwer, 1–16.
- Bybee, Joan L. (1985). *Morphology: A Study of the Relation Between Meaning and Form*. Amsterdam: Benjamins.
- Cameron-Faulkner, Thea and Andrew Carstairs-McCarthy (2000). Stem alternants as morphological signata: Evidence from blur avoidance in Polish nouns. *Natural Language and Linguistic Theory* 18, 813–835.
- Carmack, Stanford (1997). Blocking in Georgian verb morphology. *Language* 73, 314–338.
- Carstairs, Andrew (1983). Paradigm economy. *Journal of Linguistics* 19, 115–128.
- Carstairs, Andrew (1987). *Allomorphy in Inflexion*. London: Croom Helm.
- Carstairs, Andrew (1988a). Some implications of phonologically conditioned suppletion. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1988*. Dordrecht: Foris, 67–94. Reprinted (2001) In Charles Kreidler (ed.), *Phonology: Critical Concepts* (Vol. 5: *The Interface with Morphology and Syntax*). London: Routledge, 111–139.
- Carstairs, Andrew (1988b). Nonconcatenative inflection and paradigm economy. In Michael Hammond and Michael Noonan (eds.), *Theoretical Morphology: Approaches in Modern Linguistics*. San Diego: Academic Press, 71–77.
- Carstairs-McCarthy, Andrew (1991). Inflection classes: Two questions with one answer. In Frans Plank (ed.), *Paradigms: The Economy of Inflection*. Berlin: Mouton de Gruyter, 213–253.

- Carstairs-McCarthy, Andrew (1994). Inflection classes, gender and the Principle of Contrast. *Language* 70, 737–788.
- Carstairs-McCarthy, Andrew (1998). How lexical semantics constrains inflectional allomorphy. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1997*. Dordrecht: Kluwer, 1–24.
- Carstairs-McCarthy, Andrew (2001). Grammatically conditioned allomorphy, paradigm structure and the Ancestry Constraint. *Transactions of the Philological Society* 99, 223–245.
- Carstairs-McCarthy, Andrew (2002). How stems and affixes interact: Stem alternants as morphological signata. In S. Bendjaballah, W. U. Dressler, O. E. Pfeiffer and M. D. Voekova (eds.), *Morphology 2000*. Amsterdam: Benjamins, 49–57.
- Carstairs-McCarthy, Andrew (2003). Directionality and locality in allomorphy: A response to Adger, Béjar and Harbour. *Transactions of the Philological Society* 101, 117–124.
- Chomsky, Noam (1957). *Syntactic Structures*. The Hague: Mouton.
- Clark, Eve V. (1987). The Principle of Contrast: A constraint on language acquisition. In Brian MacWhinney (ed.), *Mechanisms of Language Acquisition*. Hillsdale: Erlbaum, 1–33.
- Clark, Eve V. (1993). *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.
- Corbett, Greville G. and Norman M. Fraser (1993). Network Morphology: A DATR account of Russian Nominal inflection. *Journal of Linguistics* 29, 113–142.
- Dresher, Bezalel Elan (1981). Abstractness and explanation in phonology. In Norbert Hornstein and David Lightfoot (eds.), *Explanation in Linguistics: The Logical Problem of Language Acquisition*. London: Longman, 76–115.
- Dressler, Wolfgang U. (1985). Typological aspects of Natural Morphology. *Wiener Linguistische Gazette* 35(6), 3–26. Also in *Acta Linguistica Academiae Scientiarum Hungaricae* 35, 51–70.
- Dressler, Wolfgang U., Willi Mayerthaler, Oswald Panagl and Wolfgang U. Wurzel (1987). *Leitmotifs in Natural Morphology*. Amsterdam: Benjamins.
- Elgersma, Diana and Paul Houseman (1999). Optimality Theory and Natural Morphology: An analysis of German plural formation. *Folia Linguistica* 33, 333–353.
- Enger, Hans-Olav (2004). A possible constraint on non-affixal inflection. *Lingua* 114, 59–75.
- Halle, Morris and Alec Marantz (1993). Distributed Morphology and the pieces of inflection. In Kenneth Hale and Samuel Jay Keyser (eds.), *The View from Building 20*. Cambridge, MA: MIT Press, 111–176.
- Hockett, Charles F. (1954). Two models of grammatical description. *Word* 10, 210–231. Reprinted in Martin Joos (ed.), *Readings in Linguistics I* (4th edn., 1966). Chicago: University of Chicago Press, 386–399.
- Hooper, Joan B. (1976). *An Introduction to Natural Generative Phonology*. New York: Academic Press.
- Jackendoff, Ray (1975). Morphological and semantic regularities in the lexicon. *Language* 51, 639–671.
- Jensen, John T. and Margaret Stong-Jensen (1984). Morphology is in the lexicon! *Linguistic Inquiry* 15, 474–498.
- Kilani-Schoch, Marianne (1988). *Introduction à la morphologie naturelle*. Bern: Lang.
- Lass, Roger (1990). How to do things with junk: Exaptation in language evolution. *Journal of Linguistics* 26, 79–102.

- Lieber, Rochelle (1992). *Deconstructing Morphology: Word Formation in Syntactic Theory*. Chicago: University of Chicago Press.
- Luís, Ana and Andrew Spencer (2005). A paradigm function account of 'mesoclisis' in European Portuguese. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2004*. Dordrecht: Kluwer, 177–228.
- Maiden, Martin (1992). Irregularity as a determinant of morphological change. *Journal of Linguistics* 28, 285–312.
- Maiden, Martin (2005). Morphological autonomy and diachrony. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2004*. Dordrecht: Kluwer, 137–175.
- Matthews, P. H. (1972). *Inflectional Morphology*. Cambridge: Cambridge University Press.
- Mayerthaler, Willi (1981). *Morphologische Natürlichkeit*. Wiesbaden: Athenaion.
- Neset, Tore and Hand-Olav Enger (2002). Morphological splits—iconicity and optimality. In S. Bendjaballah, W. U. Dressler, O. E. Pfeiffer and M. D. Voeikova (eds.), *Morphology 2000*. Amsterdam: Benjamins, 271–282.
- Noyer, Rolf (1992). Features, positions and affixes in autonomous morphological structure. MIT PhD dissertation. Cambridge, MA: MIT Working Papers in Linguistics.
- Robins, R. H. (1959). In defence of WP. *Transactions of the Philological Society* 116–144.
- Scatton, Ernest A. (1984). *A Reference Grammar of Modern Bulgarian*. Columbus, Ohio: Slavica.
- Selkirk, Elisabeth O. (1982). *The Syntax of Words*. Cambridge, MA: MIT Press.
- Spencer, Andrew (2003). Putting some order into morphology: Reflections on Rice (2000) and Stump (2001). *Journal of Linguistics* 39, 621–646.
- Steele, Susan (1995). Towards a theory of morphological information. *Language* 71, 260–309.
- Stump, Gregory T. (1997). Template morphology and inflectional morphology. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1996*. Dordrecht: Kluwer, 217–241.
- Stump, Gregory T. (2001). *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
- Thomas-Flinders, Tracy (ed.) (1981). *Inflectional Morphology: Introduction to the Extended Word-and-Paradigm Theory*. Los Angeles: UCLA.
- Wurzel, Wolfgang Ullrich (1984). *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie-Verlag.
- Zwicky, Arnold M. (1985). How to describe inflection. In Mary Niepokuj, Mary Van Clay, Vassiliki Nikiforidou and Deborah Feder (eds.), *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*. Berkeley: California, 372–386.

*Department of Linguistics
School of Classics and Linguistics
University of Canterbury
Private Bag 4800
Christchurch
New Zealand*

andrew.carstairs-mccarthy@canterbury.ac.nz

Some criticisms of Carstairs-McCarthy's conclusions

GREGORY T. STUMP

In his paper "Affixes, stems and allomorphic conditioning in Paradigm Function Morphology", Andrew Carstairs-McCarthy (hereafter C-M) reacts to those parts of *Inflectional Morphology* (*IM*) that relate to his own research interests; here, I show that the central assertions in his paper rest on empirically indefensible assumptions and on faulty argumentation. The main points in C-M's paper are that, contrary to the assumptions of *IM*: (a) concatenative and non-concatenative inflection play distinct roles in the architecture of a language's morphology; (b) an affix can be related to a set σ of morphosyntactic properties in either of two ways: it can be an exponent of σ (i.e. σ can specify its content), or it may require some accompanying piece of morphology to realise σ (i.e. σ can specify its context); (c) the association of a morphomic stem with particular cells in a lexeme's paradigm is not effected by multiple rules of exponence; and (d) a theory incorporating property co-occurrence restrictions is no more restrictive than one incorporating disjunctive rule ordering.

C-M's argument in support of (a) is that the No Blur Principle (Cameron-Faulkner and C-M 2000) entails a fundamental distinction between concatenative and non-concatenative inflection; his argument in support of (b) is that the Peripherality Constraint (Carstairs 1987) entails a fundamental distinction between content properties and context properties. Thus, he argues that the assumptions of *IM* are incompatible with the No Blur Principle and the Peripherality Constraint. But as I show in sections 1 and 2, the mere fact of this incompatibility is not a convincing argument against the assumptions in *IM*, since neither the No Blur Principle nor the Peripherality Constraint can be validly maintained. In support of conclusion (c), C-M argues that the cells in a lexeme's paradigm with which a morphomic stem is associated are stipulated *en bloc*; as I show in section 3, this claim is empirically unsupported. In support of conclusion (d), C-M argues that by recurring to property co-occurrence restrictions, virtually any instance of disjunctive rule ordering can be mimicked; as I show in section 4, this claim is false.

Though my arguments here are intended to point out the errors in C-M's paper, I hope that they will be more generally seen as a reminder that in the study of human language, the need to subject one's assumptions to a high standard of empirical verifiability is unrelenting.

1. THE NO BLUR PRINCIPLE IS NOT VALID

In *IM*, I assert that

although concatenative and non-concatenative inflection differ in their phonological expression, there is no convincing basis for assuming that

they perform different functions or occupy different positions in the architecture of a language's morphology; there is, in other words, no empirical obstacle to the assumption in [1]

- [1] There is no theoretically significant difference between concatenative and non-concatenative inflection. (p. 9)

C-M rejects this assumption on the grounds that it is incompatible with his No Blur Principle (NBP). This principle entails that among the rival affixes for any inflectional cell, at most one affix may fail to be a class-identifier, in which case that one affix is the class-default for that cell (Cameron-Faulkner and C-M 2000:816). To see why the NBP presupposes a fundamental boundary between concatenative and non-concatenative inflection, suppose that the NBP were reformulated as referring not specifically to rival affixes, but to rival inflectional markings of any sort:

- (2) Hypothetical reformulation of the NBP: Among the rival inflectional markings for any inflectional cell, at most one marking may fail to be a class-identifier (in which case that one marking is the class-default for the cell).

This hypothetical reformulation is clearly falsified by a range of evidence. Consider, for example, the declension of the five Russian nouns in Table 1¹; although these all belong to Declension II, they belong to the five distinct stress classes in (3).²

- (3) a. Stress class Bi: stem stress–nom. pl.; ending stress–elsewhere
 b. Stress class Bii: stem stress–acc. sg. and nom. pl.; ending stress–elsewhere
 c. Stress class Ci: stem stress–sg. and nom. pl.; ending stress–elsewhere
 d. Stress class D: stem stress–pl.; ending stress–sg.
 e. Stress class Di: stem stress–pl. and acc. sg.; ending stress–elsewhere

The Russian paradigms in Table 1 disconfirm the hypothetical principle in (2), for although stem stress and ending stress are rival inflectional markings in each of the case/number cells in these paradigms, neither is a class-identifier; in the accusative singular cell, for instance, stem stress is an expression of membership in Classes Bii, Ci and Di, and end stress is an expression of membership in Classes Bi and D. If the NBP is to be reconciled with these facts, it must be weakened so as to refer specifically to rival affixes rather than to rival inflectional markings in general. Thus, C-M concludes, the NBP is the locus of a theoretically significant difference between concatenative and non-concatenative morphology.

Table 1. Paradigm of five Declension II nouns in Russian

Stress class	SKOVORODA 'frying pan'		BORODA 'beard'		DOL'A 'portion'		DIRA 'hole'		DUŠA 'soul'	
	Bi		Bii		Ci		D		Di	
	Sg	Pl	Sg	Pl	Sg	Pl	Sg	Pl	Sg	Pl
Nom	skovorodá	skovorodi	borodá	bórodi	dól'a	dól'i	dirá	díri	dušá	dúši
Acc	skovorodú	skovorodi	bórodu	bórodi	dól'u	dól'i	dirú	díri	dúšu	dúši
Gen	skovorodí	skovoród	borodí	boród	dól'i	dól'ej	dirí	dír	duší	dúš
Prep	skovorodé	skovorodáx	borodé	borodáx	dól'e	dól'ax	diré	dírax	dušé	dúšax
Dat	skovorodé	skovorodám	borodé	borodám	dól'e	dól'am	diré	díram	dušé	dúšam
Instr	skovorodój	skovorodám'i	borodój	borodám'i	dól'oj	dól'am'i	dirój	díram'i	dušój	dúšam'i

But just as (2) is disconfirmed by a range of empirical evidence, so, it turns out, is the NBP itself; for instance, it is counterexemplified by at least three distinct phenomena in Sanskrit: the third-person plural active agreement suffixes *-(a)n* and *-ur*; the dual direct-case suffixes *-ī* and *-au*; and the locative singular suffixes *-i* and *-ām*. Consider each of these in turn.

In Vedic (the earliest attested form of Sanskrit), the suffixes *-(a)n* and *-ur* both realise the morphosyntactic property set ‘third person plural active’. In the conditional and subjunctive moods, *-(a)n* invariably appears to the exclusion of *-ur*; in the optative and precative moods and the non-subjunctive forms of the perfect tense, *-ur* invariably appears to the exclusion of *-(a)n*. In indicative forms of the imperfect and aorist tenses, however, both *-(a)n* and *-ur* are found. Consider first the imperfect indicative. In the imperfect paradigms of third-conjugation verbs, *-ur* appears uniformly; in those of verbs belonging to the first conjugation or any of the fourth through the tenth conjugations, *-(a)n* appears uniformly. In the imperfect paradigms of second-conjugation verbs, both *-(a)n* and *-ur* are found: *-ur* appears (apparently as an optional alternant of *-(a)n*) in the paradigms of all such verbs whose roots end in *ā* and in certain others as well; *-(a)n* alone appears in the imperfect paradigms of the remaining second-conjugation verbs. These facts are summarised in Table 2³; note that throughout, *-(a)n* loses its vowel after a stem ending in *a*.

Turn now to the aorist indicative. In the aorist indicative paradigms of verb whose aorist stems end in short *a* (= verbs belonging to the thematic, reduplicating, and *sa*-aorist conjugations), *-(a)n* appears uniformly; in the remaining sigmatic conjugations (= the *s-*, *iṣ-*, and *siṣ-*aorist conjugations), *-ur* appears uniformly. In the aorist indicative paradigms of verbs belonging to the root-aorist conjugation, both *-(a)n* and *-ur* appear: *-ur* appears in the paradigms of all such verbs whose roots end in *ā* and of certain others as well; *-(a)n* appears in the aorist indicative paradigms of the remaining root-aorist verbs. These facts are summarised in Table 3.

According to the NBP, only one of the two suffixes *-ur* and *-(a)n* can be a default; the other must be a class-identifier; but neither is a class-identifier, since they both cut across a range of conjugation classes. The two suffixes therefore counterexemplify the NBP.

At first blush, three ways of rescuing the NBP might appear to be available: one would be to argue that *-ur* and *-(a)n* are not actually in competition, but are in fact restricted to complementary phonological environments; another would be to argue that the two suffixes are not actually in competition because one of the two suffixes has a special stem alternant as a part of its “*signatum*”; and the last would be to argue that the NBP remains valid if one relativises it to smaller subparadigms. Consider each of these possibilities in turn.

The claim that *-ur* and *-(a)n* are restricted to complementary phonological environments might appear to be supported by the fact that in the present, the perfect, and the aorist, there is a correlation between the choice of suffix and

Table 2. The distribution of the Vedic third-person plural active agreement suffixes *-(a)n* and *-ur* in the present and imperfect tenses (Macdonell 1910: §§412, 425, 428, 436, 443, 446, 459, 462, 468, 474, 480, Whitney 1889: §§550, 621)

Tense/mood	Conjugation	-(a)n	-ur	Examples	
Present optative	all			tarey-ur 'they might pass'	
	all	x	x	dadha-n 'they should put'	
Present subjunctive	Thematic:				
	1st	x		abhava-n 'they were'	
	4th	x		avaya-n 'they wove'	
	6th	x		aviśa-n 'they entered'	
	10th	x		acintaya-n 'they thought'	
	Athematic:				
	2nd	roots in ā	(x)	x	ap-ur 'they protected' [root pā], ayā-n ~ ay-ur 'they went' [root yā]
		certain lexically specified roots	(x)	x	atviṣ-ur 'they were stirred up', avid-ur 'they knew', as-an ~ as-ur 'they threw', caḅ-ur 'they saw', duh-ur 'they milked'
		other roots	x		adhi-an 'they smeared', anāv-an 'they praised', abruv-an 'they said', avy-an 'they enjoyed', āv-an 'they went' [root i], ās-an 'they were', -ghn-an 'they smote'
	Imperfect indicative	3rd		x	ājuhav-ur 'they sacrificed', adadh-ur 'they put', amamad-ur 'they were exhilarated'
5th		x		akṛiv-an 'they made'	
7th		x		abhind-an 'they split'	
8th		x		avanv-an 'they won'	
9th		x		akṛiṅ-an 'they bought'	

Table 3. The distribution of the Vedic third-person plural active agreement suffixes *-(a)n* and *-ur* in the perfect and aorist tenses (Macdonell 1910: §§412, 485, 487, 489, 499, 502, 504, 508, 514, 522, 529, 534, 536, Whitney 1889: §§550, 621)

Tense/mood	Conjugation	-(a)n	-ur	Examples	
Perfect indicative	all		x	jagm-ur 'they went'	
	all		x	jagamy-ur 'they might go'	
	all	x		tatana-n 'they should stretch'	
Aorist indicative	Asigmatic:	roots in ā	x	ag-ur 'they went' [root ga], ad-ur 'they gave' [root da], adh-ur 'they put' [root dhā], ap-ur 'they drank' [root pā], asth-ur 'they stood' [root sthā]	
			certain lexically specified roots	x	akram-ur 'they strode' [root kram], nṛt-ur 'they danced' [root nṛt], ayam-ur 'they reached' [root yam], dabh-ur 'they harmed' [root dabh], mand-ur 'they were exhilarated' [root mand]
				other roots	x
	Thematic Reduplicating		x	asica-n 'they poured'	
		Sigmatic:	x	ajjana-n 'they gave birth'	
			s	x	abhāṣ-ur 'they feared'
	iṣ		x	apāviṣ-ur 'they cleansed'	
	Aorist optative	Sigmatic:		x	ayāṣ-ur 'they went'
			sa	x	adukṣa-n 'they milked'
				x	āśy-ur 'they might attain'
	Aorist subjunctive		x	yama-n 'they should reach'	

the phonology of the verb stem: *-(a)n* is used if the stem ends in a short *a* (as e.g. with the aorist indicative stems *asica-*, *aḥjana-* and *adhukṣa-*), and *-ur* otherwise (as with the aorist indicative stems *akram-*, *abhaiṣ-*, *apāviṣ-* and *ayāsiṣ-*). But this correlation does not hold in the imperfect indicative, where *-(a)n* appears widely both after short *a* and after a consonant. Moreover, the correlation is far from perfect in the aorist indicative, where *-(a)n* appears widely after consonant-final stems in the root aorist; indeed, in the forms *nṛt-ur* 'they danced' and *avṛt-an* 'they turned' the suffixes appear in environments that are phonologically and prosodically identical in all relevant respects. The choice between *-ur* and *-(a)n* is not phonologically conditioned.

A second means of rescuing the NBP would be to try to argue that *-ur* and *-(a)n* are not actually in competition because one of them has a special stem alternant as a part of its "signatum"; this is the maneuver that Cameron-Faulkner and C-M (2000) exploit to reconcile the NBP with some otherwise problematic evidence from Polish. Whether one regards this as a valid solution for the Polish data, it is clear that it has no application here. Consider, for instance, the second-conjugation verbs *vid* 'know' and *dih* 'smear' in Table 2: both verbs have a strong stem (*ved-*, *deh-*) and a weak stem (*vid-*, *dih-*), and in the imperfect, *-ur* joins with the weak stem of one and *-(a)n* with the weak stem of the other. Neither suffix is invariably associated with a special stem alternant with which the other suffix is incompatible; that is, the rivalry between the two suffixes cannot be dismissed by claiming that one of them has a special stem alternant as a part of its "signatum".

The last hope for rescuing the NBP would be to argue that it holds with respect to smaller subparadigms ("slabs", in C-M's terminology) in the inflection of a Vedic verb. In particular, one might claim that *-ur* is the default third-person plural active suffix in a verb's optative and perfect indicative slabs, and that *-(a)n* is the default third-person plural active suffix in all other slabs. This means that *-ur* must be a class-identifier in the imperfect and the aorist indicative. But it isn't quite a class-identifier in the imperfect, since it appears in more than one of the relevant conjugation classes (namely the second and third conjugations); moreover, it competes directly with *-(a)n* in the imperfect of the second conjugation. Nor can one say that *-ur* is a class-identifier in the aorist indicative, since it appears in four of the seven relevant conjugation classes (namely the root, *ṣ*, *iṣ* and *siṣ* aorist conjugations), in one of which it competes directly with *-(a)n*. In short, any hope of rescuing the NBP has now run out.

The Sanskrit dual direct-case suffixes *-ī* and *-au* also counterexemplify the NBP. The distribution of these suffixes is not phonologically conditioned; for instance, they appear in the same phonological context in the paradigm of *BALIN* 'powerful', whose dual direct-case forms are *balinau* (masculine/feminine) and *balinī* (neuter). The suffix *-au* regularly appears in masculine and feminine forms in all declensions except the *i-* and *u-*stem declensions (whose masculine and feminine dual direct-case forms are suffixless forms with a lengthened stem vowel)

Table 4. The distribution of the Sanskrit dual direct-case suffixes *-ī* and *-au* (Whitney 1889: §§339, 364, 424, Vasu 1962, Vol. I, p. 123f)

Declension	Gender	Suffixes		Suffixless form with lengthened stem vowel	Examples
		<i>-ī</i>	<i>-au</i>		
i-stem	masc			x	alī ‘two bees’
	fem			x	matī ‘two minds’
	neut	x			vāriṇī ‘two waters’
u-stem	masc			x	paśū ‘two cattle’
	fem			x	dhenū ‘two cows’
	neut	x			vastunī ‘two things’
derived ā-stem	fem	x			senā-ī (→ sene) ‘two armies’
	masc		x		gopau ‘two cowherds’
other declension/ gender combinations	masc		x		rājānau ‘two kings’
	fem		x		nadyau ‘two rivers’
	neut	x			nāmnī ‘two names’

and the feminine ā-stem declension; the suffix *-ī* regularly appears in feminine forms in the ā-stem declension as well as in all neuter forms. (The examples in Table 4 illustrate.) Since neither suffix is specific as to gender, both must be assumed to realise the property set {CASE:direct NUMBER:dual}; since both affixes cut across a range of declension classes, neither is a class-identifier; and neither coincides regularly with the use of a special stem.

The Sanskrit singular locative-case suffixes *-i* and *-ām* further counterexemplify the NBP. The distribution of these suffixes is not phonologically conditioned; indeed, they alternate dialectally if not freely in some paradigms, e.g. that of *भू* ‘earth’ (locative singular *bhūvi* ~ *bhuvām*). The suffix *-ām* regularly appears in the ā-stem and derived *-ī*- and *-ū*-stem declensions (regardless of gender—cf., e.g. the locative singular *senānyām* of the masculine noun *SENĀNĪ* ‘army general’); it also appears as an optional alternant in the feminine *-ī*- and *-ū*-stem declensions and in the radical *-ī*- and *-ū*-stem declensions. The suffix *-i* appears in all declensions other than (i) the ā-stem declension, (ii) the masculine and feminine *-ī*- and *-ū*-stem declensions and (iii) the derived *-ī*- and *-ū*-stem declensions. The masculine and feminine *-ī*- and *-ū*-stem declensions instead exhibit a suffixless form in which *au* replaces the stem vowel; among feminine *-ī*- and *-ū*-stems, this suffixless form alternates with the *-ām*-suffixed form. Among radical *-ī*- and *-ū*-stems, *-i*-suffixed locatives alternate with *-ām*-suffixed locatives. (The examples in Table 5 illustrate.) Since neither *-i* nor *-ām* is specific as to gender, both must be assumed to realise the property set {CASE:locative NUMBER:singular}; since both affixes cut across a range of declension classes, neither is a class-identifier; and neither coincides regularly with the use of a special stem.

Table 5. The distribution of the Sanskrit locative singular suffixes *-ī* and *-ām* (Whitney 1889: §§339, 341, 351, 364, 424, Vasu 1962, vol. I, p. 135ff)

Declension	Gender	<i>-ī</i>	<i>-ām</i>	Suffixless form with au in place of stem vowel	Examples
derived ā-stem	fem		x		senāyām 'in the army'
	masc	x			gopi 'at the cowherd'
derived ī-stem	fem		x		nadyām 'in the river'
	masc	x	x		senānyām 'at the general' vātapramī 'at the antelope' vadhvām 'at the woman'
derived ū-stem	fem		x		aticamvām 'at one victorious over armies'
	masc	x	x		khalapvi 'at the sweeper'
i-stem	fem		x		matyām, matau 'in the mind'
	masc			x	alau 'at the bee'
	neut	x		x	vāriṇī 'in the water'
u-stem	fem		x		dhenvām, dhenau 'at the cow'
	masc			x	paśau 'at the beast'
	neut	x			madhuni 'in the honey'
radical ī-stem	fem	x	x		dhiyi, dhiyām 'in the thought'
radical ū-stem	fem	x	x		bhuvī, bhuvām 'in the earth'
other declension/gender combinations		x			rājñi 'at the king'

Readers familiar with other heavily inflected languages will have no trouble uncovering comparable counterevidence to the NBP. In view of such evidence, it is clear that the NBP cannot disconfirm assumption [1] because the NBP is itself invalid.

2. THE PERIPHERALITY CONSTRAINT IS NOT VALID (OR IS IRRELEVANT)

In *IM*, I observe that lexical theories of inflectional morphology (in which affixes have the status of full-fledged lexical items)

make it possible to associate an affix with a morphosyntactic property in two different ways: a given property may, on the one hand, serve as part of an affix's content; on the other hand, it may serve as part of an affix's subcategorization restriction, limiting the range of contexts into which the affix may be inserted. [...] In inferential-realizational theories, by contrast, an affix's morphosyntactic properties are not artificially sorted into properties of content and properties of context [... Such theories] are compatible with assumption [4].

[4] Exponence is the only association between inflectional markings and morphosyntactic properties. (p. 10f)

C-M rejects this assumption on the grounds that it is incompatible with the Peripherality Constraint (PC), which depends upon a fundamental distinction between an affix's content properties and its context properties.

According to the PC, the realisation of a property may be sensitive inwards (to a property whose exponent is situated more centrally in the word-form) but not outwards (to an individual property whose exponent is situated more peripherally); cf. C-M (1992:213). As it is formulated, this principle does not obviously necessitate a distinction between content properties and context properties, but the particular notion of 'sensitivity' on which the PC is based is that of 'pure sensitivity' (Carstairs 1987:150ff). Under that interpretation, the PC entails that if (i) affix *x* has property P as its content and is restricted to the context of property Q, and (ii) affix *y* has Q as its content and appears in both the presence and the absence of P (making it Q's "principal exponent"), then *y* cannot be peripheral to *x* in any word in which they both appear.⁴ In C-M's view, the PC is the locus of a theoretically significant difference between content properties and context properties, and as such, is incompatible with assumption [4]. But like the NBP, the PC turns out to be disconfirmed by a range of empirical evidence; consider some cases in point.

Table 6. Partial paradigms of two Eastern Mari verbs (Sebeok and Ingemann 1961:49f)

Conjugation	ŠOGAŠ 'stand'		TOLAŠ 'come'	
	<i>em</i> -class		<i>am</i> -class	
	General indicative	Past personal	General indicative	Past personal
1sg	šog-e-m	šog-əš-əm	tol-a-m	tol'-əm
2sg	šog-e-t	šog-əš-əč	tol-a-t	tol'-əč
3sg	šog-a	šog-əš	tol-eš	tol'-o
1pl	šog-e-na	šog-əš-na	tol-ə-na	tol-na
2pl	šog-e-da	šog-əš-ta	tol-ə-da	tol-da
3pl	šog-a-t	šog-əš-t	tol-ə-t	tol'-əč

In Swahili, a negative verb form's polarity is most regularly expressed by its leftmost prefix. Ordinarily, this is the prefix *ha-* preceding the verb form's subject-agreement prefix, as in *ha-tu-ta-taka* [NEG-1PL-FUT-want] 'we will not want'; thus, on C-M's assumptions, a negative verb form's leftmost prefix is its principal exponent of polarity. Now, in a verb form inflected for tense, tense is expressed by a prefix following the subject-agreement prefix, as in the example just cited. Yet, in negative past-tense forms, the tense prefix *ku-* also expresses negative polarity, as in *ha-tu-ku-taka* 'we did not want'; compare the affirmative from *tu-li-taka* [1PL-PAST-want] 'we wanted'. Since the principal exponent of negative polarity is peripheral to the tense prefix in forms such as *hatukutaka*, the sensitivity of the tense prefix to negative polarity in such forms is contrary to the PC.

In Eastern Mari, a finite verb form's subject-agreement properties are most regularly expressed by its inflectional termination, which, on C-M's assumptions, must be identified as that verb form's principal exponent of subject agreement: 1sg *-m*, 2sg *-t/-č*, 3sg (none), 1pl *-na*, 2pl *-da*, 3pl *-t/-č*. In most finite verb forms, the agreement termination is preceded by a tense suffix. In some instances, this tense suffix also expresses a subject-agreement property: thus, in the inflection of ŠOGAŠ 'stand' in Table 6, the present-tense suffix *-a* expresses third person, and in the inflection of TOLAŠ 'come', the present-tense suffix *-ə* expresses plural number. Since the principal exponent of subject agreement is peripheral to the tense suffix, the sensitivity of the tense suffix to properties of subject agreement is contrary to the PC.

In Pech (Chibchan; Honduras), a finite verb form's principal exponent of tense is a suffix, e.g. *-wá* (present/immediate past), *-(u)ši* (recent past), *-(r)í?* (earlier past), *-h* (immediate future), or *-pí* (future). The position of this tense suffix is peripheral to that of the verb's subject-agreement suffix; yet, as the forms in Table 7 show, a Pech verb's agreement suffix is sometimes sensitive to tense. In particular, the usual first-person singular suffix *-a* (in underlined boldface

Table 7. Partial paradigm of the Pech verb *ašah* 'sit' (Holt 1999:49ff)

	Present/immediate past	Recent past	Earlier past	Immediate future	Future
1sg	<i>ašah-a</i> -wá	<i>ašah-à</i> -u-ši ¹	<i>ašah-a</i> -rŋ	<i>ašah-á</i> -h	<i>ašah</i> -pá-há?
2sg	<i>ašah-y</i> -á	<i>ašah-i</i> -ši	<i>ašah-u</i> -rŋ	<i>ašah-ú</i> -h	<i>ašah-u</i> -pí-há?
3sg	<i>aših</i> -wá	<i>aših</i> -u-ši	<i>aših</i> -ŋ	<i>aših-i</i> -h	<i>aših</i> -pí-há?
1du excl	<i>ašah-pär-wá</i>	<i>ašah-pär-ši</i>	<i>ašah-pär-ŋ</i>	<i>ašah-par-á-h</i>	<i>ašah-a?</i> -pí-há?
1pl	<i>ašah-bàr-wá</i>	<i>ašah-bàr-ši</i>	<i>ašah-bar-ŋ</i>	<i>ašah-bar-á-h</i>	<i>ašah-bàr-pí-há?</i>
2pl	<i>ašah-y</i> -a -wí	<i>ašah-i</i> -šu -wa	<i>ašah-u</i> -rŋ-wa	<i>ašah-u</i> -wá	<i>ašah-u</i> -pí-wá
3pl	<i>aših-ir</i> -wá	<i>aših-ir</i> -ši	<i>aših-ir</i> -ŋ	<i>aših-ir</i> -á-h	<i>aših-ir</i> -pí-há?

¹ By a principle of sandhi, *ašah-à*-u-ši is ultimately realized as *ašahōši*.

in Table 7) is omitted in the future tense; the usual second-person agreement suffix -u is supplanted by -i in the present and recent past; and the usual first-person dual exclusive suffix -par is supplanted by -a? in future-tense forms. Since a verb's principal exponent of tense is peripheral to that of its principal exponent of subject agreement, the sensitivity of certain subject-agreement suffixes to properties of tense is contrary to the PC.

Perhaps it was such evidence that moved C-M (1992:210ff) to provide a loophole for the PC by decreeing that it pertains only to layered inflection, not to templatic inflection; so as to take advantage of this loophole, one might try to argue that verb inflection is templatic in Swahili, Eastern Mari and Pech, and is therefore irrelevant to evaluating the validity of the PC. But if one examines the criteria that have been used to distinguish layered morphology from templatic morphology, one finds that by those criteria, inflection is inherently templatic (Stump 1997). Some may find this conclusion alarming because of the connotations of the unfortunate "template" metaphor. But given the possibility of conceiving of "templates" as paradigm functions rather than as positive output constraints, this conclusion becomes perfectly reasonable, since there is strong independent motivation for the postulation of paradigm functions (Stump 2001, 2005). Once the inherently templatic nature of inflection is acknowledged, the claim that the PC does not pertain to templatic inflection amounts to the claim that it doesn't pertain to inflection, period—in which case the PC is simply irrelevant to the evaluation of assumption [4].

3. MORPHOMIC STEMS ARE NOT ALWAYS ASSOCIATED WITH CELLS *en bloc*

In his discussion, C-M distinguishes two alternative ways of analysing the distribution of a morphomically indexed stem within the paradigm in which it appears. On one hand, there is the morphosyntactic realisation analysis, according to which "[i]ndividual cells (or morphosyntactically natural classes of them, such as 'direct case' embracing 'nominative and accusative', or all the singular cells in a given tense and mood) are associated with the special stem directly by means of realisation rules"; the alternative is the distributional uniformity analysis, according to which "[i]ndividual cells are not associated with the special stem directly. Rather, the set of triggering cells is associated *en bloc* with the special stem for all multi-stemmed lexemes of the relevant kind". (C-M is mum on the non-trivial question of how this association actually takes place, but the essential point is that across all of the relevant paradigms, a certain set of cells is associated once and for all with the use of a special stem.) C-M argues that the distributional uniformity analysis is superior, on the grounds that it is more restrictive: unlike the morphosyntactic realisation analysis, it excludes the possibility that a stem bearing a particular morphomic index might appear in one set of cells in

one paradigm but that an identically indexed stem might appear in a partially overlapping set of cells in some other paradigm.

C-M's discussion focusses on a pseudo-Sanskrit language which he invents for the purpose of this argument; but let us return to the language whose analysis is the basis for C-M's objection. As it turns out, the Sanskrit evidence is incompatible with the assumptions of the distributional uniformity analysis. Consider, for example, the five adjectival lexemes in Table 8: each of these lexemes exhibits an alternation between what have traditionally been characterised as strong and weak stems. Each of the present-participial lexemes *ADANT* 'eating', *BHAVANT* 'being' and *JUHVANT* 'sacrificing' has a strong stem in the so-called *Guṇa* grade (in which a short *a* appears before the stem-final *nt* cluster); these are the respective stems *adant-*, *bhavant-* and *juhvant-*. Each of the other two lexemes in Table 8 has a strong stem appearing sometimes in the *Guṇa* grade—as *caḥṛvāms-* and *ekamūrdhan-*, respectively—and sometimes in the so-called *Vṛddhi* grade (in which a long *ā* appears before the stem-final consonantism)—as *caḥṛvāms-* and *ekamūrdhān-*. Table 8 indicates the particular cells in which the strong stem appears in each lexeme's paradigm. As scrutiny of this evidence reveals, there is no way to specify *en bloc* the cells in which a strong stem appears in a way which is valid across all of the relevant paradigms. The distributional uniformity analysis fails here for three reasons. First, it fails to account for the fact that the cells occupied by the strong stem in the relevant paradigms are subject to variation; second, it fails to account for the *Guṇa* ~ *Vṛddhi* alternations appearing in some (but not all) paradigms; and third, it fails to capture important generalisations across lexemes, such as the fact that the strong stem always appears in the *Guṇa* grade in the vocative singular, the locative singular, and the direct cases of the neuter dual (Stump 2001:193f, 286, fn. 14). The morphosyntactic realisation analysis, by contrast, has no difficulty accounting for these facts: the distribution of the strong stem forms in Table 8 can be accounted for by applying the stem-selection rules in (5) and the stem-indexing rules in (6) as in Table 9.

(5) Stem-selection rules

- | | | |
|----|---|---------------------------------------|
| a. | Strong stem in masc. direct cases | [cf. Stump (2001), rule (3a), p. 179] |
| b. | Middle stem in masc. acc. pl. | [cf. rule (3b), p. 179] |
| c. | Strong stem in neut. pl. direct cases | [cf. rule (3c), p. 179] |
| d. | <i>Guṇa</i> -grade stem in the neut. du. direct cases | [cf. footnote 14, p. 286] |
| e. | <i>Guṇa</i> -grade stem in the masc. voc. sg. | [cf. rule (19), p. 194] |
| f. | <i>Guṇa</i> -grade stem in the neut. voc. sg. | |
| g. | <i>Guṇa</i> -grade stem in the loc. sg. | [cf. footnote 14, p. 286] |

Table 8. Five variant patterns of strong stem distribution in Sanskrit nominal paradigms (Whitney 1889: §§424, 435, 447, 453, 461)

	Masculine						Neuter			
	N sg	V sg	A sg	NVA du	NV pl	L sg	V sg	NVA du	NVA pl	L sg
ADANT 'eating'	Guṇa	Guṇa	Guṇa	Guṇa	Guṇa					
BHAVANT 'being'	Guṇa	Guṇa	Guṇa	Guṇa	Guṇa			Guṇa	Guṇa	
JUHVAṆT 'sacrificing'	Vrddhi	Guṇa	Vrddhi	Vrddhi	Vrddhi				Vrddhi	(Guṇa)
CAKRVAS 'one who did'	Vrddhi	Guṇa	Vrddhi	Vrddhi	Vrddhi			(Guṇa)	(Guṇa)	(Guṇa)
EKAṀURDHAN 'facing the same direction'	Vrddhi	Guṇa	Vrddhi	Vrddhi	Vrddhi			(Guṇa)	Vrddhi	(Guṇa)

Note. N = nominative; V = vocative; A = accusative; L = locative; (Guṇa) = free variation between Guṇa-grade stem and weak stem.

Table 9. A morphosyntactic realisation analysis of the distribution of the strong stem forms in Table 8

	Stem-selection rules							Stem-indexing rules	
	(6a)	(6b)	(6c)	(6d)	(6e)	(6f)	(6g)	(7a)	(7b)
ADANT ‘eating’	x	x	x		x			x	
BHAVANT ‘being’	x	x	x	x	x			x	
JUHVANT ‘sacrificing’			(x)					x	
CAKṚVAS ‘one who did’	x	x	x		x				x
EKAMŪRDHAN ‘facing the same direction’	x	x	x	(x)	x	(x)	(x)		x

(6) Stem-indexing rules

- a. Strong stem = Guṇa-grade stem [cf. rule (16a), p. 190]
- b. Strong stem = Vṛddhi-grade stem [cf. rule (16d), p. 190]

The claim here is not, of course, that one must always resort to the morphosyntactic realisation analysis; surely there are some cases in which the distributional uniformity analysis will suffice (provided that the intuitive notion of *en bloc* associations can be given some precise content). But it is wrong to imagine that the distributional uniformity analysis will always afford a credible alternative to the morphosyntactic realisation analysis; it certainly does not in Sanskrit.

4. PROPERTY CO-OCCURRENCE RESTRICTIONS ARE NOT EQUIVALENT TO DISJUNCTIVE RULE ORDERINGS

C-M asserts that

Stump’s property cooccurrence restrictions can be used as a loophole to mimic almost any rule ordering (within a single rule block) that is of the kind purportedly excluded by the Paninian Determinism Hypothesis [PDH]. [...] I say ‘almost any ordering’ rather than ‘any ordering’ because of the unlikely possibility that the property cooccurrence restrictions are restricted to one inflection class, so that they are not available to mimic blocking for lexemes in another inflection class.

Here is proof that this loophole assertion is false. Suppose that in some language, adjectives inflect for 3 cases × 3 numbers × 3 genders, and that in Block I, every case/number/gender combination has a different exponent, as in (7a); this means

that in this language, there are no restrictions whatever on the co-occurrence of an adjective's morphosyntactic properties. Suppose, too, that Block II contains the three rules in (7b), and that these rules must be mutually exclusive in their application. One way of guaranteeing the mutual exclusion of these rules' application is to treat them as being disjunctively ordered; an alternative, according to C-M, is to exploit the use of property co-occurrence restrictions. But that alternative is not available here, since the rules in Block I exclude the existence of any pertinent property co-occurrence restrictions in this language. Here, then, is an instance of three disjunctively ordered rules belonging to a single block (=Block II) whose ordering cannot be mimicked by property co-occurrence restrictions.

(7) A hypothetical analysis involving disjunctively ordered rules that cannot be reconciled with the PDH by resorting to property co-occurrence restrictions

a. Block I (rules do not compete)

$X_v, \{\text{nom sg fem}\} \rightarrow Xa$	$X_v, \{\text{acc sg fem}\} \rightarrow Xj$	$X_v, \{\text{gen sg fem}\} \rightarrow Xs$
$X_v, \{\text{nom sg masc}\} \rightarrow Xb$	$X_v, \{\text{acc sg masc}\} \rightarrow Xk$	$X_v, \{\text{gen sg masc}\} \rightarrow Xt$
$X_v, \{\text{nom sg neut}\} \rightarrow Xc$	$X_v, \{\text{acc sg neut}\} \rightarrow Xl$	$X_v, \{\text{gen sg neut}\} \rightarrow Xu$
$X_v, \{\text{nom du fem}\} \rightarrow Xd$	$X_v, \{\text{acc du fem}\} \rightarrow Xm$	$X_v, \{\text{gen du fem}\} \rightarrow Xv$
$X_v, \{\text{nom du masc}\} \rightarrow Xe$	$X_v, \{\text{acc du masc}\} \rightarrow Xn$	$X_v, \{\text{gen du masc}\} \rightarrow Xw$
$X_v, \{\text{nom du neut}\} \rightarrow Xf$	$X_v, \{\text{acc du neut}\} \rightarrow Xo$	$X_v, \{\text{gen du neut}\} \rightarrow Xx$
$X_v, \{\text{nom pl fem}\} \rightarrow Xg$	$X_v, \{\text{acc pl fem}\} \rightarrow Xp$	$X_v, \{\text{gen pl fem}\} \rightarrow Xy$
$X_v, \{\text{nom pl masc}\} \rightarrow Xh$	$X_v, \{\text{acc pl masc}\} \rightarrow Xq$	$X_v, \{\text{gen pl masc}\} \rightarrow Xz$
$X_v, \{\text{nom pl neut}\} \rightarrow Xi$	$X_v, \{\text{acc pl neut}\} \rightarrow Xr$	$X_v, \{\text{gen pl neut}\} \rightarrow Xü$

b. Block II (rules are disjunctively ordered)

$X_v, \{\text{acc}\} \rightarrow X\ddot{e}$	$X_v, \{\text{sg}\} \rightarrow X\ddot{i}$	$X_v, \{\text{fem}\} \rightarrow X\ddot{o}$
---	--	---

Counterexamples to C-M's loophole claim need not even involve a second rule block. Suppose that some language has only a single block of rules for the inflection of its adjectives, and that these fall into two inflection classes, A and B. The rule block has 27 rules for adjectives belonging to Class A, namely the rules in (7a), and these supply a distinct exponent for each case/number/gender combination; this means that in this language, there are no restrictions whatever on the co-occurrence of an adjective's morphosyntactic properties (on the natural assumption, apparently endorsed by C-M in the passage above, that property co-occurrence restrictions are never limited to members of a particular inflection class). Then suppose that the lone block of rules has three rules for the inflection of adjectives belonging to Class B, namely the three rules listed in (7b). It is clear that property co-occurrence restrictions cannot be exploited to cause the application of these three rules to mimic a disjunctive ordering restriction, since there are no property co-occurrence restrictions to which adjectives are subject in this language.

Thus, it is not true that property co-occurrence restrictions are always available to mimic the effects of disjunctive rule ordering. In the published version of his paper, however, C-M does not claim that they are always available, only that they are “almost” always available. But what does “almost” mean here? C-M cannot mean that the purported loophole is open in almost all logically possible instances of disjunctive rule ordering, since there are countless imaginable variants of the counterexamples sketched above. He also cannot mean that the purported loophole is open in almost all empirically attested systems of inflectional rules; this would be an irresponsible claim unless one could show evidence of having investigated a provably representative sample of the thousands of systems at issue. In actuality, C-M has shown very little about the tradeoff between disjunctive rule ordering and property co-occurrence restrictions: only that it is possible to construct a system of disjunctively ordered rules whose effects can be mimicked by property co-occurrence restrictions. But demonstrating that possibility does not prove much, since as I have shown, one can just as easily construct a system of disjunctively ordered rules whose effects cannot be mimicked in this way. This latter possibility proves there is a real empirical difference between Paradigm Function Morphology (PFM) and frameworks that countenance disjunctive rule ordering, and that fact cannot be altered simply by denying that any difference exists.

Before leaving the loophole claim, I must respond to C-M’s suggestion that property co-occurrence restrictions are essential to reconciling PFM with the facts of Russian verb morphology:

Does Stump’s framework then implicitly claim that Russian is an impossible or at least an implausible language? Not quite, because Stump’s discussion of the verbal inflection of Bulgarian in chapter 2 opens a loophole [that of property co-occurrence restrictions].

This statement too is false. The reason that my framework does not imply that Russian is an impossible or implausible language has nothing to do with property co-occurrence restrictions, but with basic assumptions about exponence. A fundamental assumption is that if an inflectional marking *m* only appears in words that express the morphosyntactic property *p*, then *m* is an exponent of *p*. In view of that assumption, each of the terminations in the present-tense paradigm of the Russian verb *СТОЯТ* ‘stand’ (Table 10) must be seen as an exponent of tense and finiteness as well as of person/number agreement; similarly, each of the terminations in the past-tense paradigm of *СТОЯТ* must be seen as an exponent of tense and finiteness as well as of gender/number agreement. Thus, the rules of exponence relevant for the partial paradigm in Table 10 are as in (8).⁵ The analysis in (8) is neither impossible nor implausible, and in no way requires the postulation of any property co-occurrence restrictions.⁶

Table 10. Partial paradigm of the Russian verb *STOJAT'* 'stand'

	1st	2nd	3rd
Nonpast			
Singular	stoj-u	stoi-š'	stoi-t
Plural	stoi-m	stoi-te	stoj-at
	Masc	Fem	Neut
Past			
Singular	stoj-a-l	stoj-a-l-a	stoj-a-l-o
Plural	stoj-a-l-i		

Infinitive: *stoj-a-t'*.

(8) **Block I** Where X is a past-tense stem:

X, {finite past} → *Xl* X, {infinitive} → *Xr'*

Block II Where X is a non-past-tense stem:

X, {finite nonpast 1 sg} → *Xu* X, {finite nonpast 2 pl} → *Xte* X, {finite past fem sg} → *Xa*
 X, {finite nonpast 2 sg} → *Xš'* X, {finite nonpast 1 pl} → *Xm* X, {finite past neut sg} → *Xo*
 X, {finite nonpast 3 sg} → *Xt* X, {finite nonpast 3 pl} → *Xat* X, {finite past pl} → *Xi*

C-M concludes his discussion of disjunctive rule ordering with a claim that the PDH is subverted by the fact that content properties and context properties are not distinguished in PFM; in particular, he claims that a rule whose interaction with other rules is inconsistent with the PDH can invariably be brought into conformity with it by including context properties in the property set that the rule realises. This argument, however, rests on an imperfect understanding of the import of the PDH. According to the PDH, the optimal definition of a language's grammar is one in which competition between morphological rules is always resolved by Pāṇini's principle. That is, the PDH entails not merely that it is possible to analyse morphological rule competition purely in terms of Pāṇini's principle, but that this will always be the most economical way of analysing such competition.

This is not a vacuous claim. Consider again the disjunctively ordered rules in (7b). According to these rules, an adjectival lexeme having the schematic paradigm in Table 11 exhibits the accusative suffix *-ë* in the (light-shaded) cells 4–6, 13–15, and 22–24; the singular suffix *-ï* in the (dark-shaded) cells 1–3 and 7–9; and the feminine suffix *-ö* in the (boxed) cells 11, 17, 20 and 26. Given their formulation in (7b), the desired interaction among these rules is incompatible with the PDH; in order to bring them into conformity with the PDH, it is necessary

Table 11. Cells in the paradigms defined by the rules in (7)

	Nominative			Accusative			Genitive		
	Masc	Fem	Neut	Masc	Fem	Neut	Masc	Fem	Neut
Singular	1	2	3	4	5	6	7	8	9
Dual	10	11	12	13	14	15	16	17	18
Plural	19	20	21	22	23	24	25	26	27

to treat the *-ë* rule as applying in expanded mode (i.e as realising all well-formed extensions of {acc}) and to split the *-ö* rule into two rules, one realising the property set {fem du} and the other realising the property set {fem pl}. Thus, although one can bring the rules in (7b) into conformity with the PDH, the resulting analysis is less economical than that afforded by the disjunctive ordering in (7b).

As this example shows, the validity of the PDH is an empirical question. One can easily invent morphologies such as (7) in whose definition one would have to resort to disjunctive rule ordering in order to capture generalisations that would otherwise remain unaccounted for; other morphologies can easily be invented which require still other sorts of devices, such as the Feature Ranking Principle (Noyer 1992), and so on. The essential claim of the PDH is that the morphologies of real languages are not like those inventions—that in the definition of real morphologies, there are no significant generalisations about the resolution of rule competition whose formulation depends on devices other than Pāṇini's principle.

NOTES

¹ In Table 1, I use the system of transcription used by Brown et al. (1996), in which the representation of unstressed vowels abstracts away from the effects of automatic processes of unstressed vowel reduction.

² I assume here the analysis of Russian stress presented by Brown et al. (1996).

Note that SKOVORODA, BORODA and DOL'A all exhibit stem stress in the accusative plural because, as inanimate nouns, their accusative plural is syncretised with their nominative plural.

³ The injunctive mood is omitted from Tables 2 and 3 because it generally patterns after the indicative.

⁴ There is, however, a loophole: *y* may be peripheral to *x* if *x* is restricted to the context of every property belonging to the same category as *Q*. (See Carstairs (1987:165ff) for discussion.)

⁵ For brevity's sake, I follow C-M in ignoring participles and gerunds, whose characteristics in no way affect the general point being made here.

⁶ One could, however, use property co-occurrence restrictions to abbreviate the rules in (8); for instance, a restriction that required any complete well-formed extension of the property

set $\{\alpha\}$ (for any tense α) to be an extension of $\{\text{finite}\}$ would make it possible to eliminate the specification 'finite' from the rules in (8).

REFERENCES

- Brown, Dunstan, Greville Corbett, Norman Fraser, Andrew Hippisley and Alan Timberlake (1996). Russian noun stress and Network Morphology. *Linguistics* 34, 53–107.
- Cameron-Faulkner, Thea and Andrew Carstairs-McCarthy (2000). Stem alternants as morphological signata: Evidence from blur avoidance in Polish nouns. *Natural Language and Linguistic Theory* 18, 813–835.
- Carstairs, Andrew (1987). *Allomorphy in Inflexion*. London: Croom Helm.
- Carstairs-McCarthy, Andrew (1992). *Current Morphology*. London: Routledge.
- Holt, Dennis (1999). *Pech (Paya)*. Munich: Lincom Europa.
- Macdonell, A. A. (1910). *Vedic Grammar* [*Grundriss der Indo-Arischen Philologie und Altertumskunde*, Band I, Heft 4]. Strassburg: Karl J. Trübner.
- Noyer, Robert Rolf (1992). *Features, Positions and Affixes in Autonomous Morphological Structure*. Doctoral dissertation, MIT.
- Sebeok, Thomas and Frances J. Ingemann (1961). *An Eastern Cheremis Manual: Phonology, Grammar, Texts and Glossary*. Bloomington: Indiana University publications [Uralic and Altaic series, v. 5].
- Stump, Gregory T. (1997). Template morphology and inflectional morphology. In G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1996*. Dordrecht: Kluwer, 217–241.
- Stump, Gregory T. (2001). *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
- Stump, Gregory T. (2005). Rules about paradigms. In C. Orhan Orgun and Peter Sells (eds.), *Morphology and the Web of Grammar: Essays in Memory of Steven G. Lapointe*. Stanford: CSLI Publications, 49–82.
- Vasu, Śrīśa Chandra (ed. and tr.) (1962). *The Siddhānta Kaumudī of Bhaṭṭoji Dīkṣita* [1995 reprint]. Delhi: Motilal Banarsidass.
- Whitney, W. D. (1889). *Sanskrit Grammar*, 2nd edn. Cambridge, MA: Harvard University Press.

University of Kentucky

gstump@uky.edu

Book reviews

Paul Boucher (ed.), *Many Morphologies*. Sommerville: Cascadilla Press, 2002, xvi + 267 p. ISBN 1-57473-025-8

Reviewed by Peter Arkadiey

The book reviewed here is a collection of nine papers dealing with various aspects of current morphological research. The ambitious title *Many Morphologies* suggests a wide coverage of topics from diverse theoretical perspectives. However, although many different conceptions are indeed present in the articles comprising the book, one cannot regard it as an adequate survey of the field. I believe, though, that this volume's goal was not to represent all major (formal) theories of morphology, but instead to show a variety of (not always purely morphological) approaches to morphological problems. Indeed, most of the papers are to a large extent 'inter-modular' and interdisciplinary, dealing with relationships between morphology and syntax, semantics, and lexicon.

The book begins with an introduction (pp. vii–xv) by Paul Boucher and Marc Plénat, where they give a short characteristic of the volume as a whole, followed by a very useful 'guide' to the book, which consists of informative summaries of the articles. Although there is no division of articles by topic, a natural grouping suggests itself: 'theoretical' papers, 'morphology and lexical semantics', morphological patterns from 'exotic' languages, and 'computational morphology'. I will deal with these groups in the reversed order.

Nabil Harthout, Fiammetta Namer and Georgette Dal's article 'An experimental constructional database: The MorTAL project' (pp. 178–209) presents a detailed description and comparison of two systems of morphological databases for French, DéCor and DériF, whose aim is to extract morphological information from annotated corpora. The second paper dealing with computational morphology is Béatrice Daille, Cécile Fabre, and Pascale Sébillot's general outline of the field, 'Applications of computational morphology' (pp. 210–234). They survey different kinds of morphological information used in natural language processing, as well as a large variety of applications using this information, and give brief descriptions of certain systems, primarily of those designed in France. Although both articles are informative and instructive they give an impression that almost all what is done in current 'theoretical' approaches to morphological phenomena is to a large extent orthogonal to problems which arise in the field of natural language processing. It is not so simple to decide whose fault this situation is.

Christian Bassac and Pierrette Bouillon's paper 'Middle transitive alternations in English: A generative lexicon approach' (pp. 29–47) presents a nice and convincing account of various semantic and syntactic restrictions on middle formation in English, such as its argument structure and aspectual properties, in the

framework of *Generative Lexicon* (Pustejovsky 1995), which aims at representing both lexical and constructional aspects of syntax and semantics in a unified way.

In her paper 'Unaccusativity mismatches and unaccusativity diagnostics from derivational morphology' (pp. 48–81) Bozena Cetnarowska addresses the problem of using morphological derivatives as diagnostic tests for unaccusativity/unergativity. Carefully investigating data from English and Polish, she reaches the following conclusion: the occurrence of a derivative proves that its base is an unaccusative or unergative verb, whereas the non-occurrence is non-conclusive in this respect. Also, Cetnarowska shows that existence or non-existence of a certain derived form are usually subject to fine-grained semantic and pragmatic constraints.

Susan Steele in her article 'Many plurals: Inflection, informational additivity, and morphological processes' (pp. 82–108) presents an account of plural formation in the Uto-Aztecan language Luiseño in the framework of the so called 'information-based morphology' (cf. Steele 1995), whose main theoretical postulate is that all morphological operations add information (where information may be phonological, semantic, or syntactic). Conceptual ideas of Steele's proposal are of certain interest, but the overall impression of the article is unsatisfactory. The author develops a very sophisticated system of rules and principles in order to account for the facts which, I believe, if not rather trivial, certainly do not demand such a complex description and so many stipulations. The main problem which Steele addresses is the fact that it is words and not stems which serve as bases for plural formation in Luiseño. If to account for a property of the overwhelming majority of the world's languages one needs formal devices so sophisticated as Steele's, then, in my opinion, one, instead of proving the superiority of one's theory, shows that its basic assumptions require thorough revision.

Much of the same is true of Jacqueline Lecarne's paper 'Gender "polarity": Theoretical aspects of Somali nominal morphology' (109–141), which presents a Distributed Morphology account of the gender 'reversal' common in the Cushitic languages, when some nouns change gender (from masculine to feminine and vice versa) in the plural. Lecarne gives a detailed description of the facts and evaluates some previous proposals, and then gives her own account, whose main idea is that plural morphemes in Somali belong to a special functional category intermediate between N^0 and DP. This conception allows the author to neatly describe all the necessary facts, but considered from the theoretical perspective, it leaves an impression (honestly speaking, similar to that left by other DM proposals) of a purely ad hoc solution.

Anna Maria Di Sciullo's article 'The asymmetry of morphology' (pp. 1–28) aims at showing that in morphology, as well as in syntax, asymmetrical relations between items play major part. She argues that although both syntactic and morphological processes and rules are sensitive to the same relations (i.e., Spec-Head and Head-Complement asymmetries), and although both components require

similar representations (binary branching trees) and operations (SHIFT, which derives complex categories from more elementary ones, and LINK which is analogous to chain formation), morphology and syntax are nevertheless not the same. The crucial difference, as Di Sciullo shows, lies in that morphological heads, e.g. nominalising affixes, are sensitive to the asymmetrical argument structure of their bases, whereas syntactic heads are not. Di Sciullo analyses English suffixes *-er*, *-able*, *-ify*, *-ize*, and *-ee* and shows that their application depends on the argument structure of the base. She also deals with English compounding, and shows how her framework can account for special behaviour of compounds containing *wh*-expressions (such as *everywhere*). I cannot but appreciate the way Di Sciullo presents her framework and argues for it. It's main advantage lies, I think, in that it avoids ad hoc stipulations and counter-intuitive solutions.

Luigi Burzio in his article 'Surface-to-surface morphology: When your representations turn into constraints' (pp. 142–177) develops an Optimality-theoretic implementation of insights by Bybee (1985), who proposed that morphology should be regarded as a network of relations between surface forms. The main principle of Burzio's conception is Gradient Attraction (GA), which states that 'the overall structure of a word *w* (in both its phonological and semantic components) is influenced by that of other words in the lexicon to which *w* is independently similar'. GA itself is explained as a result of summation of entailments generated by lexical items, predicted by the Representational Entailments Hypothesis (REH), which says that a mental representation of a lexical item with a structure AB generates the entailments $A \rightarrow B$ and $B \rightarrow A$. The more entailments a given lexical item violates, the less similar it is to its 'relatives'. Burzio shows that GA and REH effects account for various phenomena, such as stress patterns in English and morphophonological alternations in English, Polish, Italian and French.

In his article 'A common basis for syntax and morphology: Tri-Level lexical insertion' (pp. 235–262), Joseph Emonds pursues the goal of completely reducing morphology to syntax, the view advocated in particular by Lieber (1992). Emonds states that morphology cannot be 'entirely explained in terms of current theories of phrasal syntax', and that there is need to develop a theory of word-internal syntax, which, however, would not differ substantially from the syntax *per se*. He addresses two problems which he believes to be 'the most relevant for morphology': 'productive processes of syntactic compounding' and 'conditions which relate the permanently stored lexicon to syntax'. The main tenet of his approach is the Domain Size Restriction, which precludes syntactic phrases to occur within words. Emonds claims that principles of syntax always treat bound morphology and syntactic compounds in the same way, therefore there is no need for any 'autonomous' morphology (as proposed, e.g., by Anderson (1992) or Stump (2001)). The evidence Emonds invokes for his claim is of several kinds. First, he shows that in English and French rules determining the position of heads in phrases, compounds and derivatives cut across the boundary between 'phrasal'

and 'word-internal' domains. Second, he shows that English derivational affixes such as *-able*, *-age*, *-en*, *-er* etc. are 'lexical categories which lack properties of simple words', and behave in all respects like 'normal' words except that they cannot go all by themselves. With regards to semantics of bound morphological items Emonds proposes to distinguish between what he calls *cognitive syntactic features* such as HUMAN or AGENT, and *purely semantic features* of higher specificity (such as those borne by items like *-holic* in *workaholic*). Emonds claims that bound morphological affixes cannot bear purely semantic features, differing in this respect from pieces of compounds. Thus the difference between bound morphology and syntactic compounding is reduced to semantics and has little to do with 'syntax proper'. Further, Emonds observes the variable behavior of morphological items and proposes that they may be inserted on three different levels of derivation: items bearing purely semantic features are inserted at the first level with 'ordinary' lexical items; those whose cognitive syntactic features contribute to interpretation, are inserted 'during syntactic derivation leading to LF', while those which do not bear any (non-predictable) semantic features, 'are inserted in the part of a derivation inaccessible to LF' and thus contribute to PF only. Emonds illustrates this by analyzing the variable behavior of English nominalizing affixes.

Emonds' approach has some weak points which undermine it considerably. First, I wonder how Emonds is going to tackle with such 'classical' morphological problems as non-affixal inflection (which is quite widespread in English, although he even does not mention it), extended exponence, surface-to-surface relationships investigated in Burzio's article etc. Second, the claim that only 'cognitive syntactic features' may serve as content for morphological affixes, seems to me untenable; see Talmy (1985) for an abundance of counterexamples. At last but not least, I cannot think of the natural way Emonds' theory can naturally account for different *syntactic* properties of morphology and syntax, such as limited vs. unlimited recursion of embedding or differences in argument structure sensitivity explored by Di Sciullo. Thus, I think that Emonds' arguments for his theory are not convincing and the evidence he brings about is not sufficient to prove it.

Now, what is one to say about the value of the book as a whole? It is, undoubtedly, a useful book; all articles, though differing profoundly in topic, data, framework, and persuasiveness of the argumentation, present interesting ideas worthy of discussion. However, I must confess that I am somewhat disappointed, because I waited more from a book with such a title and such authors. One of the main weak points is, I think, a certain degree of self-containment: sometimes it seems that authors are simply not aware of the problems and data discussed by their fellow-contributors, and this ignorance seriously affects the weight and value of their own argumentation.

Closing my review, I have to say that the book is well designed, edited, and printed, but there is a number of misprints, some of them in crucial points.

REFERENCES

- Anderson, Stephen R (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Bybee, Joan (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.
- Lieber, Rochelle (1992). *Deconstructing Morphology*. Chicago: University of Chicago Press.
- Pustejovsky, James (1995). *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Steele, Suzan (1995). Towards a theory of morphological information. *Language* 71, 260–309.
- Stump, Gregory T. (2001). *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
- Talmy, Leonard (1985). “Lexicalization patterns: semantic structure in lexical form”. In Timothy Shopen (ed.), *Language typology and syntactic description. Vol. III. Grammatical categories and the lexicon*. Cambridge: Cambridge University Press, 57–149.

Danielle Corbin, Pierre Corbin, and Martine Temple (eds.), *Lexique 16 (2004) La formation des mots: horizons actuels*. Villeneuve d’Ascq: Presses Universitaires du Septentrion. EUR 23.00. ISBN 2-85939-844-9

Reviewed by Claudio Iacobini

This is the last volume of the *Lexique* series conceived by Danielle Corbin (hereinafter D.C.). Its publication—originally scheduled in 1999—was delayed due to the illness and death of D.C. in August 2000. It is mainly thanks to the painstaking editorial work of Pierre Corbin that this publication has now appeared.

There are two main reasons for paying special attention to *Lexique 16*: it can be read both for the high scientific value of the individual contributions, and as a *state of the art* of morphological studies in France.

This second reason of interest is made evident by Pierre Corbin’s introductory paper. It mainly consists in a sort of intellectual biography of D.C. presenting her role as a promoter of morphological studies in France. At the same time readers are acquainted with the developments of morphology in France over the last twenty years; the seven other papers of this issue are critically presented within the context of current research lines of the small but active and organised group of French morphologists.

Although the work of D.C. has not been given due consideration in English-speaking countries (perhaps owing to the fact that most publications have appeared in the French language; amongst the few exceptions is an article by D.C., Corbin 1989, in the *Yearbook of Morphology 1989* translated from French into English by Geert Booij) it has attracted the attention of many European scholars and was crucial to the development of morphological studies in France from the late 1980s onwards (some have compared it to the influential morphological work

of Arsène Darmesteter in the last quarter of the 19th century—see Darmesteter 1894).

One of the most tangible results of the impact of D.C.'s morphological theory as well as her teaching and organizational skills was the publication—in 1991—of *Lexique 10 La formation des mots: structures et interprétations* (Corbin 1991). It is against this background that the last issue of *Lexique* can be best appreciated. A confirmation of this view can be found in the title of the opening article to the volume: 'Introduction: *Lexique 16*, treize ans après *Lexique 10*' (Introduction: *Lexique 16*, thirteen years after *Lexique 10*).

The two volumes reflect two different stages—which can be defined as those of 'youth' and 'maturity'—in the development of the French morphological studies inspired by D.C.

Lexique 10 is an act of (re)-foundation for French morphology. It is centred on a single topic—suffixal derivation—which is one of the core aspects of what D.C. defined as *morphologie constructionnelle* (*constructional morphology*—an opportune term designating the area of morphology dealing with lexeme construction i.e. word formation). The various papers are empirical applications of D.C.'s model and are meant to be a contribution to the project aiming at a detailed and theoretically consistent description of French constructional morphology.

Lexique 16, on the contrary, shows that morphology in France currently enjoys the status of an independent research field, with a specific object and its own methodology, and can rely on substantial results both on the theoretical and the empirical side. At this (mature) stage the research turns its attention to the boundaries of *morphologie constructionnelle* and its interaction with phonology, semantics and syntax; some of the basic notions—such as that of lexeme—are discussed and reconsidered, and the interactions between constructional morphology and applied domains, such as natural language processing, are explored.

Although *Lexique 16* comes with a greater diversity of topics than *Lexique 10*, all topics share a number of characteristic points which can be summarised as follows: (a) the lexicon is not a list of irregularities, on the contrary the *lexique construit* (i.e. complex words) is governed by rules that it is up to morphological theory to explain; (b) a search for an explicit and consistent stipulation of the principles regulating the construction of lexical units; (c) association between word structure and meaning; (d) multi-level approach to the lexicon (and to word-formation rules); (e) close interaction between theory and analysis of linguistic phenomena (taking into account a large amount of empirical data).

Besides serving as an introduction, the article by Pierre Corbin (pp. 9–52) presents the history of morphology in France as it revolves around the figure of D.C. Such an interesting and detailed review reaches into the future and describes the main research projects under way as well as the conferences and meetings which are organized in France both to ensure better coordination as a consequence of the decentralisation of research institutions and to help the

increasing exchanges with foreign scholars. Corbin's contribution comes to an end after providing a full list of D.C.'s publications.

The volume also includes an article by D.C. (pp. 53–66), in which a synthesis of the central tenets of her model of morphology are illustrated in order to provide both the scientific background and the aim of two closely connected ambitious lexicographical projects: the *Dictionnaire des affixes dérivationnels du français* and the *Dictionnaire dérivationnel du français*. The first consists in a lexicographical presentation of the rules and processes that govern the construction of lexical units; the second is the structural and semantic description of French morphologically complex words. Unfortunately it is most unlikely that such projects will ever be completed, now that D.C.'s organisational skills and scientific knowledge are no longer available.

D. Amiot (pp. 67–83) examines the status of some elements, which can be used, in French, both as prepositions and as initial elements of word formation. The convincing conclusion is that these elements form a continuum with at one end clear instances of prepositions, at the other end real prefixes, and in between some elements that lean towards one or the other of the two poles. It is a kind of analysis that is currently agreed upon by a number of scholars, and also the set of criteria employed to disentangle homophonous elements are fairly well established. Amiot's contribution, besides the clear descriptive results, stands out for providing an explanation of the endocentricity of prefixal derivatives and of the exocentricity of prepositional compounds. According to Amiot, this distinction depends on the different argument structure of prepositions and prefixes: the first relate two elements whereas the latter refer to only one argument (the base).

The long article (pp. 85–123) by F. Kerleroux 'Sur quels objets portent les opérations morphologiques de construction?' examines the kind of information a word formation rule requires about its base to operate properly. The data considered are clipping and suffixal derivation. According to Kerleroux the lexeme, in the way it is currently defined (see Matthews 1974 and Aronoff 1994), doesn't provide as much information as is needed for morphological construction rules (whereas inflectional rules have enough information). For example, the intrinsic ambiguity of deverbal nouns suffixed with *-tion* (between processes and result interpretation) must be solved in order to justify clipping processes. The need for more semantic information provided by the lexeme is shown by the fact that only nouns without argument structure can be possible candidates for clipping—see examples (1a) and (1b)—whereas process nouns cannot; see examples (2a) and (2b):

- (1a) *La manifestation des étudiants a duré cinq heures*
- (1b) *La manif des étudiants a duré cinq heures*
- (2a) *La manifestation de la vérité aura pris cinquante ans*
- (2b) **La manif de la vérité aura pris cinquante ans*

Kerleroux's position about the need for more semantic information provided by the lexeme in order to be a suitable base for derivational processes is interesting, but at the same time it poses serious problems regarding the identity of the lexeme once it is provided with richer semantic information.

Clipping is the main issue of D. Delaplace's article (pp. 125–158). A clear review of all the principal theoretical proposals about processes of word shortening serves as a base for a detailed analysis of clipping from learned compounds and from suffixed deverbal nouns. Delaplace demonstrates that clipping is governed in both cases not only by constraints relating to prosodic phonology, but is conditioned also by the morphological structures of complex words. The balance between phonological and morphological factors is a matter for future research.

The article by M. Plénat and M. Roché (pp. 159–198) also deals with the interaction between morphology and phonology. The authors introduce a new term, *suffixation décalée* (shifted suffixation) instead of the current *interfix* and *interfixation*. They criticise *interfix* because such a denomination recalls a morphemic element, whereas interfixes are usually lacking in semantic relevance. The authors' conclusion is that the rather complex phonological conditions (both prosodic and segmental) that govern interfixation in Occitan are not the only sector responsible for the shape of derivatives: morphological factors also play a role, since the interfixes are often derived from the reduced forms of suffixes and their choice is not entirely predictable from the phonological context.

G. Dal, N. Hathout and F. Namer (pp. 199–229) describe a project called "MorTAL" devoted to the creation of a database for Natural Language Processing. The illustration of this project shows that applied disciplines and more theoretical ones can mutually benefit from each other. In particular computer programs for Data Mining and Information Retrieval can be implemented, taking advantage of the morphological decomposition of constructed lexemes.

Starting from a review of the interesting book by B. Fradin (2003), G. Dal (pp. 231–263) tackles the fundamental problem of which categorial dimensions lexeme construction rules are sensitive to. Reminding us that no word formation rule selects as bases all the members of a category (noun, adjective and verb) and that some affixes can select bases of more than one category, Dal emphasises the fact that bare category information is insufficient to describe derivational rules. The base of Dal's discussion is Croft's (1991) theory of category, according to which categories can be defined by the combination of semantic dimensions and pragmatic functions. Unlike Croft, Dal argues that a (more fine-grained) semantic dimension suffices in accounting for lexeme construction processes. The promising proposal (here only sketched) also seems able to account for the preferences an affix shows when it can be attached to bases of more than one category.

In conclusion, although morphological studies in France will no longer enjoy D.C.'s contribution, by now they can rely on a group of active researchers with common objectives, a sound internal organisation and increasing relationships

abroad. The vitality of French morphology is also shown by the fact that—besides *Lexique 16*—two leading French journals have recently published special issues on morphology (Dal 2003, Fradin and Kerleroux 2003). Coming soon is D.C.'s posthumous work (Corbin in press) consisting in a revision of her theory of constructional morphology, as well as a book dedicated to her (Fradin in press).

REFERENCES

- Aronoff, M. (1984). *Morphology by Itself. Stems and Inflectional Classes*. Cambridge: MIT Press.
- Corbin, D. (1987). *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Niemeyer; republished Villeneuve d'Ascq: Presses Universitaires de Lille, 1991.
- Corbin, D. (1989). Form, structure and meaning of constructed words in an associative and stratified lexical component. In G. Booij and J. van Marle (eds.), *Yearbook of Morphology 2*. Dordrecht: Foris, 31–54.
- Corbin, D. (ed.) (1991). *Lexique 10 La formation des mots: structures et interprétations*. Villeneuve d'Ascq: Presses Universitaires du Septentrion.
- Corbin, D. (to appear). *Le lexique construit. Méthodologie d'analyse*. Villeneuve d'Ascq: Presses Universitaires du Septentrion.
- Dal, G. (ed.) (2003). La productivité morphologique en questions et en expérimentations. *Langue française*, 140.
- Darmesteter, A. (1894). *Traité de la formation des mots composés dans la langue française comparée aux autres langues romanes et au latin*. 2nd edn. with a preface by G. Paris. Paris: Barillon; republished Paris: Champion, 1967; 1st edn. Paris: Vieweg, 1874.
- Fradin, B. (2003). *Nouvelles approches en morphologie*. Paris: Presses Universitaires de France.
- Fradin, B. (ed.) (in press). *La raison morphologique. Hommage à Danielle Corbin*. Amsterdam: John Benjamins.
- Fradin, B., and F. Kerleroux (eds.) (2003). Quoi de neuf en morphologie?. *Langages*, 152.
- Matthews, P. H. (1974). *Morphology: An Introduction to the Theory of Word-Structure*. Cambridge: Cambridge University Press; 2nd edn. *Morphology*, 1991.

Laurie Bauer, *A Glossary of Morphology*. Edinburgh: Edinburgh University Press, 2004. ISBN 0 7486 1853 8, 124 p. Price GBP 7.99

Reviewed by Geert Booij

This glossary is a useful instrument for courses in morphology. Bauer is well known for his didactic qualities in the realm of morphology, and this little book confirms this once more. It provides clear definitions of current terms in morphology. In addition, it provides a list of fundamental works on morphology, and an annotated 'Select bibliography of books on morphology'. Thus, this book will help students to find their way in the morphological literature.

Piet van Sterkenburg (ed.), *Linguistics Today—Facing a Greater Challenge*. Amsterdam: Philadelphia, 2004, 367 p. ISBN 90 2723223 7. US D 192.

Reviewed by Geert Booij

Every five years the Permanent International Committee of Linguists (CIPL) organises a world congress for linguists. In July 2003, this congress was held in Prague. A number of experts were invited to give their views on one or more of the following topics: typology, endangered languages, linguistic fieldwork, and language and the mind. The papers of these experts are published in this volume, edited by the secretary-general of CIPL, Prof. Piet van Sterkenburg of Leiden University.

Some of them are singled out here, since they are of direct relevance to morphologists. Stephen Anderson's article 'Toward a less "syntactic morphology" and a more "morphological syntax"' defends the position that morphology has 'its own character, quite distinct from the more constructional nature of syntax. Word structure involves the realisation of meaningful properties, not simply the concatenation of meaningful elements' (p. 43). Lyle Campbell, in 'Historical linguistics, the state of the art' presents his view of the current state of historical linguistics. In 'Coherent fieldwork' Daniel L. Everett gives advice as to how to do linguistic fieldwork, and how to store the data, and guarantee their accessibility. Kees Versteegh discusses 'The future of creolistics', and points out how important this discipline is for the understanding of linguistic variation and change.

An additional useful feature of these proceedings is that it also contains a CD-ROM with a large number of papers presented in the thematic sections of this congress.

Book information

Geert Booij, Angela Ralli, Sergio Scalise, and Salvatore Sgroi (eds.), *Morphology and Linguistic Typology*. Proceedings of the 4th Mediterranean Morphology Meeting, Catania, 21–23 September 2004. Published as *Siculorum Gymnasium* 1 (2004). To be ordered from Direzione *Siculorum Gymnasium*, Faculty of Letters and Philosophy, Piazza Dante 32, I-95100 Catania, Italy.

In addition to a selection of papers in the *Yearbook of Morphology 2004*, a number of papers presented at the 4th MMM meeting have been made available in the form specified above. These proceedings contain the following articles: Wolfgang Dressler ‘Morphological Typology and First Language Acquisition: Some Mutual Challenges’; Franz Rainer, ‘Typology, Diachrony, and Universals of Semantic Change in Word Formation. A Romanist’s look at the Polysemy of Agent Nouns’; Greville G. Corbett, Dunstan Brown, Marina Chumakina, and Andrew Hippisley, ‘Resources for Suppletion: A Typological Database and a Bibliography’; Stephen Anderson, ‘Morphological Universals and diachrony’; Antonio Fábregas, ‘Universals and Grammatical Categories: A Distributed Morphology Analysis of Spanish Colour Terms’; Ivan Derzhanski ‘On Diminutive Plurals and Plural Diminutives’; Jan Don, ‘Roots, Deverbal Nouns and Denominal Verbs’; Berthold Crysmann, ‘Hausa Final Vowel Shortening: Phrasal Allomorphy or Inflectional Category?’; Marian Klamer, ‘Explaining Some Structural and Semantic Asymmetries in Morphological Typology’; Nicola Grandi and Fabio Montermini, ‘Prefix-Suffix Neutrality In Evaluative Morphology’; Livio Gaeta, ‘Word Formation and Typology: Which Language Universals?’; Tore Nessel, ‘Rule Counting vs. Rule Ordering: Universal Principles of Rule Interaction in Gender Assignment’; Andrew Koontz-Garboden and Beth Levin, ‘The Morphological Typology of Change of State Event Encoding’; François Nemo, ‘Morphemes and Lexemes versus “Morphemes or Lexemes”?’; Irit Meir, ‘Typology and Boundaries: The Acquisition of a New Morphological Boundary by Modern Hebrew’; Irina Nikolaeva, ‘Modifier-Head Person Concord’; Christian Rathmann and Gaurav Mathur, ‘Unexpressed Features of Verb Agreement in Signed Languages’; Paolo Aquaviva, ‘The Morphosemantics of Transnumeral Nouns’; Darya Kavitskaya, ‘Loan Words and Declension Classes in Czech’; Alice C. Harris, ‘The Challenge of Typologically Unusual Structures’.