On Looking into Words (and Beyond)

Structures, Relations, Analyses

Edited by

Claire Bowern, Laurence Horn, and Raffaella Zanuttini

Empirically Oriented Theoretical Morphology and Syntax ??

language science press

Empirically Oriented Theoretical Morphology and Syntax

In this series:

1. Lichte, Timm. Syntax und Valenz: Zur Modellierung kohärenter und elliptischer Strukturen mit Baumadjunktionsgrammatiken

ISSN: 2366-3529

On Looking into Words (and Beyond)

Structures, Relations, Analyses

Edited by

Claire Bowern, Laurence Horn, and Raffaella Zanuttini

language science press

Claire Bowern, Laurence Horn, and Raffaella Zanuttini (ed.). 2017. *On Looking into Words (and Beyond): Structures, Relations, Analyses* (Empirically Oriented Theoretical Morphology and Syntax ??). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog

ISBN: no digital ISBN
no hardcover ISBN
no softcover ISBN
ISSN: 2366-3529
no DOI

Cover and concept of design: Ulrike Harbort
Fonts: Linux Libertine, Arimo, DejaVu Sans Mono
Typesetting software: XeLaTeX

Language Science Press
Unter den Linden 6
10099 Berlin, Germany
langsci-press.org

Storage and cataloguing done by FU Berlin

Contents

Preface **v**

1 Compensatory lengthening and structure preservation revisited yet again
Darya Kavitskaya **1**

2 Nominal verbs and transitive nouns: Vindicating lexicalism
Paul Kiparsky **17**

3 A-morphous iconicity
Ryan Lepic, Carol Padden **51**

4 Romansh allomorphy (Again!)
Martin Maiden **77**

5 Iconicity chains in sign languages
Donna Jo Napoli **97**

Indexes **115**

Preface

On Looking into Words is a wide-ranging volume spanning current research into word structure and morphology, with a focus on historical linguistics and linguistic theory. The papers are offered as a tribute to Stephen R. Anderson, the Dorothy R. Diebold Professor of Linguistics at Yale, who retired at the end of the 2016–2017 academic year. The contributors are friends, colleagues, and former students of Professor Anderson, all important contributors to linguistics in their own right. As is typical for a Festschrift, the contributions span a variety of topics relating to the interests of the honorand. In this case, the central contributions that Anderson has made to so many areas of linguistics and cognitive science, drawing on synchronic and diachronic phenomena in diverse linguistic systems, are represented through the papers in the volume.

The 23 papers that constitute this volume are unified by their discussion of the interplay between synchrony and diachrony, theory and empirical results, and the role of diachronic evidence in understanding the nature of language. Central concerns of the volume include morphological gaps, learnability, increases and declines in productivity, and the interaction of different components of the grammar. Although the volume is not divided into labeled parts, the papers deal with a range of linked synchronic and diachronic topics in phonology, morphology, and syntax (in particular, cliticization), and their implications for linguistic theory.

Several of the papers take as their general topic synchrony and diachrony, tackling the relationship between synchronic alternations and diachronic change. Thráinsson, for example, uses evidence from Modern Icelandic and Modern Faroese to argue for different trajectories in historical change in *u*-umlaut, and differing degrees of synchronic productivity of the umlaut rules. Maiden argues for the use of comparative and historical evidence in deciding between possible synchronic interpretations of phenomena, a point also made by de Chene for Japanese.

Some papers focus specifically on morphology. Maiden revisits paradigmatic phenomena in Romansch and investigates whether allomorphy in verb stems is phonologically conditioned or is evidence for paradigm autonomy. In this, he continues a long-standing debate with Anderson about the status of paradigms and their role in shaping change. The theme of residue and productivity is also taken up in Lepic & Padden's contribution. Yang's paper also uses data from Romansch (as well as English) as a test case for a proposal for identifying where rules are likely to be productive and how morphological gaps appear.

Hyman & Inkelas present a different type of morphological problem. Lusoga, a Bantu language spoken in Uganda, has multiple exponence of causative, applicative, subjunctive and perfective suffixes, which the authors argue to pose challenges for theories of

morphology that minimize redundancy and that treat derivation as distinct (and ordered before) inflection. Lusoga thus provides evidence for a previously unidentified pathway to multiple exponence. Stump also considers multiple exponence in Limbu, a Tibeto-Burman language of Nepal.

Several papers concentrate on phonology at the level of the word (and below). Kavitskaya provides a generalization about a class of apparent exceptions to rules of compensatory lengthening. Blevins looks at the typology of TR cluster resolution strategies and the role of language contact. Kaisse draws on evidence from Macedonian to investigate why stress assignment is word-bounded, despite the existence of higher-level prosodic interactions –- that is, she discusses the reasons why lexical stress is specifically lexical. Round's contribution takes a different approach to learnability, discussing Yidiny (Pama-Nyungan) word-final deletion as a case of exceptional phonological patterns.

Steve Anderson is perhaps best known for his work in phonology and morphology, but his research also profoundly informs our understanding of syntax, an interaction reflected in several of the contributions to the volume. Hale investigates the syntax that underlies the distribution of clitics that fall under "Wackernagel's Law". Hale looks at the areas where there are multiple demands on a Wackernagel position, and how those conflicts are repaired. Aissen similarly considers exceptional clitic placement, but in Tsotsil (Mayan). While Hale argues for syntactic constraints on exceptionality based on Greek, Aissen focuses on the role of phonology in clitic placement. Chung's paper examines the phonology, morphology and syntax of Chamorro causative marking, historically analyzed as a derivational prefix that appears outside inflection; she argues instead that the causative is a prosodically deficient verb. This has implications for morphological structure and dissolves an apparent counterexample to the morphological generalization that inflection is outside derivation.

The syntactic contributions also include Hendrick's argument for English of as a phrasal affix rather than a syntactic head in certain contexts and Horvath & Siloni's use of evidence from idioms to argue against theories such as Construction Grammar that treat knowledge of language as a network of stored constructions. Both Hendrick's and de Chene's papers look at the relationship of morphology to syntax, though in different languages and domains. Finally, Deo's article traces the morphosyntax of ergative alignment loss in Indo-Aryan.

Some papers investigate properties of languages and language change in the aggregate, rather than being tied to individual languages or language families. That is, they directly concern the nature of linguistic theory. Newmeyer, for example, examines the theoretical status of parameters, while Timberlake tackles the place of performance and usage in theories of transmission. Spencer's paper concerns the place of morphology within the general theory of grammar. Klein's paper, though couched in a different framework, also engages with the question of how we know what we know, and how our theories provide insight to the phenomenon of language. Newmeyer, Timberlake, and Klein focus on syntax, Spencer and Stump (discussed above) on morphology, while Goldstein's domain is articulatory phonology and its application to birdsong. Finally, both Spencer and Aronoff discuss Steve Anderson's and Peter Matthews's contributions

to morphological theory.

The volume offers excellent cross-linguistic depth. Languages discussed in the contributions include Ancient and Modern Greek (Hale), Macedonian (Kaisse), Sanskrit (Kiparsky, Deo, Hale), Indo-Aryan (Deo), Romansch (Maiden, Yang), Icelandic and Faroese (Thráinsson), Welsh (Hammond), American Sign Language (Lepic & Padden), Limbu (Stump), Lusoga (Hyman & Inkelas), Yidiny (Round), Dyirbal (Kiparsky), and Japanese (de Chene), and Mayan (Aissen) as well as English.

Chapter 1

Where, if anywhere, are parameters?: A critical historical overview of parametric theory

Frederick J. Newmeyer
University of Washington, University of British Columbia, and Simon Fraser University

Since the late 1970s, crosslinguistic variation has generally been handled by means of UG-specified parameters. On the positive side, thinking of variation in terms of parameterized principles unleashed an unprecedented amount of work in comparative syntax, leading to the discovery of heretofore unknown morphosyntactic phenomena and crosslinguistic generalizations pertaining to them. On the negative side, however, both macroparameters and microparameters have proven themselves to be empirically inadequate and conceptually nonminimalist. Alternatives to parameters are grounded approaches, epigenetic approaches, and reductionist approaches, the last two of which seem both empirically and conceptually quite promising.

1 Introduction

The existence of crosslinguistic variation has always been problematic for syntacticians. If there is a universal grammar, one might ask, then why aren't all languages exactly the same? In the earliest work in generative syntax, characterizing the space in which languages could differ, whether at the surface or at a deep level, was not a priority. At the time, surface differences between languages and dialects were generally attributed to language-particular rules or filters.

In the late 1970s, however, a strategy was developed that allowed the simultaneous development of a rich theory of Universal Grammar (UG) along with a detailed account of the limits of crosslinguistic morphosyntactic variation. In this view, syntactic complexity results from the interaction of grammatical subsystems, each characterizable in terms of its own set of general principles. The central goal of syntactic theory now became to identify such systems and to characterize the degree to which they might vary (be 'parameterized') from language to language. **chomsky1995** describes succinctly

how such variation might be accounted for in what, by the early 1980s, was called the 'principles-and-parameters' (P&P) approach.

> Within the P&P approach the problems of typology and language variation arise in somewhat different form than before. Language differences and typology should be reducible to choice of values of parameters. A major research problem is to determine just what these options are, and in what components of language they are to be found. (**chomsky1995**)

The first mention of parameters, I believe, was in **chomsky1976**

> Even if conditions are language- or rule-particular, there are limits to the possible diversity of grammar. Thus, such conditions can be regarded as parameters that have to be fixed (for the language, or for the particular rules, in the worst case), in language learning … It has often been supposed that conditions on applications of rules must be quite general, even universal, to be significant, but that need not be the case if establishing a 'parameteric' condition permits us to reduce substantially the class of possible rules. (**chomsky1976**)

An interesting question is why Chomsky at this point would propose parameters, since there is nothing in his 1976 paper that suggests that they need to be incorporated into the theory. A possible answer is that in the same year an MIT dissertation appeared (**kim1976**) that showed that Korean obeys a form of the Tensed-S-Condition, even though Korean does not distinguish formally between tensed and non-tensed clauses. That fact might have planted the seed for the idea of parameterized principles. At around the same time, an 'external' inspiration for parameters was provided by the work of Jacques Monod and François Jacob (**monod1972**; **jacob1977**). Their idea was that slight differences in timing and arrangement of regulatory mechanisms that activate genes could result in enormous differences. **berwick2011n** has claimed that 'Jacob's model in turn provided part of the inspiration for the Principles and Parameters (P&P) approach to language …'

Whatever the direct inspiration for parameterized principles might have been, their adoption triggered an unprecedent explosion of work in comparative syntax. One unquestionably positive consequence of the P&P approach to linguistic theory was to spur investigation of a wide variety of languages, particularly those with structures markedly different from some of the more familiar Western ones. The explanation for this is straightforward. In earlier transformational grammar (oversimplifying somewhat), one worked on the grammar of English, the grammar of Thai, the grammar of Cherokee, and so on, and attempted to extract universal properties of grammars from the principles one found in common among these constructed grammars. But now the essential unity of all grammars, within the limits of parametric variation, was taken as a starting point. One could not even begin to address the grammar of some language without asking the question of how principles of Case, binding, bounding, and so on are parameterized in that language. And that in turn demanded that one have a rough feel for the degree of parameterization possible for the principle. As Chomsky noted, to delimit the domain

of core grammar, we 'rely heavily on grammar-internal considerations and comparative evidence, that is, on the possibilities for constructing a reasonable theory of UG and considering its explanatory power in a variety of language types, with an eye open to the eventual possibility of adducing evidence of other kinds' (**chomsky1981**).

The core idea of the P&P approach is that both the principles of UG and the possible parameter settings are part of our genetic endowment:

> [W]hat we 'know innately' are the principles of the various subsystems of S_0 [= the initial state of the language faculty – FJN] and the manner of their interaction, and the parameters associated with these principles. What we learn are the values of these parameters and the elements of the periphery (along with the lexicon, to which similar considerations apply). The language that we then know is a system of principles with parameters fixed, along with a periphery of marked exceptions. (**chomsky1986**)

The original idea was that there are a small number of parameters and small number of settings. This idea allowed two birds to be killed with one stone. Parametric theory could explain the rapidity of acquisition, given the poor input, and explain the crosslinguistic distribution of grammatical elements. As Norbert Hornstein noted:

> The second reason in favor of parameter setting models has been their ability to provide (at least in principle) an answer to Plato's Problem [the fact that we know so much about language based on so little direct evidence – FJN]. The idea is that construing language acquisition as parameter setting eases the problem faced by the child, for setting parameter values is easier than learning the myriad possible rules of one's native language. In other words, the PLD [= Primary Linguistic Data – FJN] can be mined for parameter values more easily than it can be for rules. (**hornstein2009**)

The need to base a theory of parametric variation on the investigation of a wide variety of languages resulted in what Bernard Comrie, always a major critic of the generative approach, referred to approvingly as 'one of the most interesting recent developments in linguistic typology … the entry of generative syntax into the field' (**comrie1988**). Comparative studies of the distribution of null-subjects, binding domains, configurationality, and so on became routine by the 1980s and provided a generative interpretation of the kind of crosslinguistic typological studies that were initiated by the work of Joseph Greenberg. In this regard, it is instructive to observe Chomsky's changing rhetorical evaluation of Greenbergian typological work. His first reference to Greenberg was somewhat dismissive, noting that 'Insofar as attention is restricted to surface structures, the most that can be expected is the discovery of statistical tendencies, such as those presented by **greenberg1963n**' (**chomsky1965**). In 1981, Chomsky offered what was perhaps his first favorable reference to this line of research:

> Universals of the sort explored by Joseph Greenberg and others have obvious relevance to determining just which properties of the lexicon have to be learned

> in this manner in particular grammars – and to put it in other terms just how much has to be learned as grammar develops in the course of language acquisition. (**chomsky1981**)

By 1982 he was writing that 'Greenbergian universals … are ultimately going to be very rich. … They have all the difficulties that people know, they are "surfacy," they are statistical, and so on and so forth, but nevertheless they are very suggestive' (**chomsky1982**). And in 1986, they are 'important, … yielding many generalizations that require explanation …' (**chomsky1986**).

In this paper, I do not question the fertility of the research program that was launched by the P&P approach. What I do is to provide a critical review of the various approaches that have been taken to parameters since the late 1970s, discussing their *conceptual* strengths and weaknesses. Given space limitations, my overview will in places be unavoidably somewhat superficial. The paper is organized as follows. Sections 2 through 5 outline various approaches that have been taken with respect to parameters: UG-principle-based, microparametric, macroparametric, and interface-based respectvely. Section 6 outlines some of the major conceptual and empirical problems with the classical view of parameters and §7 discusses alternatives to the classical approach. Section 8 is a brief conclusion.

2 Parameterized UG principles

All of the subsystems of principles in the Government-Binding Theory were assumed to be parameterized. Consider a few concrete examples:

(1) Examples of parameterized UG principles:
   a. BINDING (**lasnik1991**). Principle C is parameterized to allow for sentences of the form *John$_i$ thinks that John$_i$ is smart* in languages like Thai and Vietnamese.
   b. GOVERNMENT (**manzini1987**). The notion 'Governing Category' is defined differently in different languages.
   c. BOUNDING (**rizzi1982**). In English, NP and S are bounding nodes for Subjacency, in Italian NP and S'.
   d. X-BAR (**stowell1981**). In English, heads precede their complements; in Japanese heads follow their complements.
   e. CASE and THETA-THEORY (**travis1989**). Some languages assign Case and/or Theta-roles to the left, some to the right.

Fewer and fewer parameterized UG principles have been proposed in recent years for the simple reason that there are fewer and fewer widely accepted UG principles. The thrust of the Minimalist Program (MP) has been to reduce the narrow syntactic component and to reinterpret broad universal principles as economy effects of efficient computation. Economy principles are generally assumed not to be parameterized:

> There is simply no way for principles of efficient computation to be parameterized […], it strikes me as implausible to entertain the possibility that a principle like 'Shortest Move' could be active in some languages, but not in others. Put differently, […] there can be no parameters within the statements of the general principles that shape natural language syntax. (**boeckx2011**)

On the same page Boeckx proposes the 'Strong Uniformity Thesis': Principles of narrow syntax are not subject to parameterization; nor are they affected by lexical parameters.

It should be noted that the very idea of looking for principles of UG has fallen into disrepute in recent years. For example, Chomsky has attributed to them what can only be described as negative qualities:

> [T]ake the LCA (Linear Correspondence Axiom) [**kayne1994**]. If that theory is true, then the phrase structure is just more complicated. Suppose you find out that government is really an operative property. Then the theory is more complicated. If ECP really works, well, too bad; language is more like the spine [i.e., poorly designed – FJN] than like a snowflake [i.e., optimally designed]. (**chomsky2002**)

So if in theory there are very few UG principles and no parameters associated with them, then the question is where and how to capture systematic crosslinguistic variation. Given the organization of grammars in a P&P-type model, the simplest assumption to make is that one group of languages contains a particular feature attraction mechanism that another group lacks, thus allowing the presence or absence of this mechanism to divide languages into typological classes. Some early examples can be illustrated by whether or not a feature setting determines whether V moves to I in a particular language (**emonds1978**; **pollock1989**), whether V moves to C (to derive V2 order) (den Besten 1977/1983), and whether N incorporates into V (**baker1988**).

A major debate within parametric theory has centered on the host of the attracting feature. In 'microparametric' approaches, the locus of variation lies in individual functional heads. 'Macroparametric' approaches are not so restricted. They will be discussed in §3 and §4 respectively.

3 The Borer-Chomsky Conjecture and microparametric approaches

Hagit Borer, in *Parametric Syntax* (**borer1984**), made two proposals, which she may or may not have regarded as variants of each other. One is that parameters are restricted to the idiosyncratic properties of lexical items, the other that they are restricted to the inflectional system. Borer wrote:

> …interlanguage variation would be restricted to the idiosyncratic properties of lexical items. These idiosyncrasies, which are clearly learned, would then interact with general principles of UG in a particular way. (**borer1984**)

By way of example, she discussed a rule that inserts a preposition in Lebanese Arabic – a rule that does not exist in Hebrew:

(2) ø ------> *la* / [$_{PP}$ ... NP]

Along the same lines, **manzini1987** pointed to language-particular anaphors that have to be associated with parameters: *cakicasin* and *caki* in Korean; *sig* and *hann* in Icelandic.

Now every language has thousands of lexical items, but nobody ever entertained the possibility that every lexical item might be a locus for parametric variation. Borer's proposal that only inflectional elements provide the locus for parametric variation was designed to forestall this possibility. In the same book she wrote:

> It is worth concluding this chapter by reiterating the conceptual advantage that reduced all interlanguage variation to the properties of the inflectional system. The inventory of inflectional rules and of grammatical formatives in any given language is idiosyncratic and learned on the basis of input data. (**borer1984**)

The restriction of parameters to the inflectional system is a somewhat different proposal than their restriction to lexical items. After all, not all lexical items are part of the inflectional system and not all inflections are lexical. However, Borer took 'inflectional' in a pretty broad sense, namely, to encompass Case and agreement relations, theta-role assignment, and so on. She recognized an immediate problem here: Inflection-based parameters could not handle some of the best known cases of crosslinguistic variation such as differences in head-order and extraction possibilities.

In any event, the hypothesis that the locus of parametric variation is restricted to exclude major lexical categories came to be known as the 'Borer-Chomsky Conjecture'.

borer1984 appeared before the distinction between lexical and functional categories had been elaborated. Once this distinction had become well accepted, it seemed natural to associate parameters with functional heads, rather than with inflectional items. This idea was first proposed as The Functional Parameterization Hypothesis (FPH) in **fukui1988** In this view, only functional elements in the lexicon (that is, elements such as Complementizer, Agreement, Tense, etc.) are subject to parametric variation.[1]

It is important to stress that FPH is not a simple extension of the idea that parameters are inflection-located. There have been countless functional categories proposed that have nothing to do with inflection, no matter how broadly this concept is interpreted. So adverbs, topic, focus, and so on are typically thought to be housed in functional categories, even though they are not in many languages 'inflectional'.

Associating parameters with functional heads has been claimed to have both methodological and theoretical advantages. Methodologically, it allows 'experiments' to be constructed comparing two closely-related variants, thereby pinpointing the possible degree of variation. The ideal situation then would be to compare speech varieties that differ from each other only in terms of (most ideally) one or, failing that, only a few variables. Richard Kayne remarks:

[1] Fukui himself exempted ordering restrictions from this hypothesis.

> If it were possible to experiment on languages, a syntactician would construct an experiment of the following type: take a language, alter a single one of its observable syntactic properties, examine the result and see what, if any, other property has changed as a consequence. If some property has changed, conclude that it and the property that was altered are linked to one another by some abstract parameter. Although such experiments cannot be performed, I think that by examining pairs (and larger sets) of ever more closely related languages, one can begin to approximate the results of such an experiment. To the extent that one can find languages that are syntactically extremely similar to one another, yet clearly distinguishable and readily examinable, one can hope to reach a point such that the number of observable differences is so small that one can virtually see one property covarying with another. (**kayne2000**)

In other words, in Kayne's view, this 'microparametric variation' (as he called it) is the best testing ground for the hypothesis that syntactic variation can be reduced to a finite set of parameters.

Along more theoretical lines, it has been claimed that functional-category-situated microparameters impose a strong limit on what can vary, crosslinguistic differences now being reduced to differences in features, thereby restricting learning to the lexicon (**kayne2000**; **roberts2010**; **thornton2013**).[2] Indeed, Chomsky has often asserted that microparameters are necessary in order to solve Plato's Problem:

> Apart from lexicon, [the set of possible human languages] is a finite set, surprisingly; in fact, a one-membered set if parameters are in fact reducible to lexical properties [associated with functional categories – FJN] … How else could Plato's problem be resolved? (**chomsky1991**)

4 Macroparameters

Not all minimalists have embraced the Borer-Chomsky Conjecture and consequent turn to microparameters. Mark Baker, in particular, while not denying that there are micro-level points of variation between languages, has defended what he call 'macroparameters' (**baker1996**), that is, parametric differences that cannot be localized in simple differences in attracting features of individual functional heads. He gives as examples, among others, the Head Directionality Parameter (i.e. VO vs. OV), where functional categories play no obvious role, the Polysynthesis Parameter, which in his account refers to the lexical category 'Verb', and an agreement parameter (**baker2008**) distinguishing Niger-Congo languages from Indo-European languages, which, in opposition to a strong interpretation of the Borer-Chomsky Conjecture, applies to the full range of functional categories. Another example of a macroparameter is the compounding parameter of

[2] As an anonymous reviewer points out, this claim is highly dependent on the nature of the features and the role that they play in the system.

snyder2001 which divides languages into those that allow formation of endocentric compounds during the syntactic derivation and those that do not. The NP/DP macroparameter of **boskovic2011** distinguishes 'NP languages', which lack articles, permit left-branch extraction and scrambling, but disallow NEG-raising, from 'DP languages', which can have articles, disallow left-branch extraction and scrambling, but allow NEG-raising. And **huang2007** points to many features that distinguish Chinese-type languages from English-type languages, including a generalized classifier system, no plural morphology, extensive use of light verbs, no agreement, tense, or case morphology, no overt *wh*-movement, and radical pro-drop.

Baker and other advocates of macroparameters share the conviction long held by advocates of holistic typology that languages can be partitioned into macro-scope broad classes, typically (or, at least, ideally) where the setting of one feature entails a cascade of shared typological properties. As Baker puts it, 'the macroparametric view is that there are at least a few simple (not composite) parameters that define typologically distinct sorts of languages' (**baker2008**).

5 Parameters as being stated at the interfaces

Under the perspective that parameters are stated at the interfaces, lexical items are subject to a process of generalized late insertion of semantic, formal, and morphophonological features after the syntax, which is where all variation would take place. Or, as another possibility, the parametric differences would derive from the way in which such features are interpreted by the interfaces or by processes that manipulate the features on the path from spell out to the interfaces. There has been some debate as to whether there is parametric variation at the Conceptual-Intentional (C-I) interface. Angel Gallego remarks:

> … it would be odd for semantic features to be a source of variation, which leaves us with formal and morphophonological features as more likely suspects. … Considered together, the observations by **chomsky2001b** and **kayne2005**; **kayne2008** appear to place variation in the morphophonological manifestation of closed classes (i.e. functional categories, which contain unvalued features).' (**gallego2011**)

However, for **ramchand2008** the narrow syntax provides a 'basic skeleton' to C-I, but languages vary in terms of how much their lexical items explicitly encode about the reference of variables like T, Asp, and D.

6 Conceptual and empirical problems with parameters

Before moving on to nonparametric approaches to variation, it would be useful to highlight some of the main problems with the classic view of parameters as being innately-provided grammatical constructs (for an earlier discussion, see **newmeyer2005**).

6.1 No macroparameter has come close to working

The promise of parameters in general and macroparameters in particular is that from the interaction of a small number of simply-stated parameters, the vast complexity of human language morphosyntax might be derived. As Martin Haspelmath put it:

> According to the principles and parameters vision, it should be possible at some point to describe the syntax of a language by simply specifying the settings of all syntactic parameters of Universal Grammar. We would no longer have any need for thick books with titles like *The Syntax of Haida* (cf. **enrico2003** 2003's 1300-page work), and instead we would have a simple two-column table with the parameters in the first column and the positive or negative settings in the second column. (**haspelmath2008**)

Needless to say, nothing remotely like that has been achieved. The problem is that 'few of the implicational statements at the heart of the traditional Principles-and-Parameters approach have stood the test of time' (**boeckx2011**). The clustering effects are simply not very robust. The two most-studied macroparameters, I believe, are the Null Subject (Pro-drop) and the Subjacency parameters, neither of which is much evoked in recent work. As for the former: 'History has not been kind to to the Pro-drop Parameter as originally stated' (**baker2008**). And Luigi Rizzi notes that 'In retrospect, [subjacency effects] turned out to be a rather peripheral kind of variation. Judgments are complex, graded, affected by many factors, difficult to compare across languages, and in fact this kind of variation is not easily amenable to the general format of parameters …' (**rizzi2014**).

6.2 'Microparameter' is just another word for 'language-particular rule'

Let's say that we observe two Italian dialects, one with a *do*-support-like structure and one without. We could posit a microparametric difference between the dialects, perhaps hypothesizing that one contains an attracting feature that leads to *do*-support and one that does not. But how would such an hypothesis differ in substance from saying that one dialect has a rule of *do*-support that the other one lacks? Indeed, Norbert Hornstein has stressed that 'microparameter' is just another words for 'rule':[3]

> Last of all, if parameters are stated in the lexicon (the current view), then parametric differences reduce to whether a given language contains a certain lexical item or not. As the lexicon is quite open ended, even concerning functional items as a glance at current cartographic work makes clear, the range of variation between grammars/languages is also open ended. In this regard it is not different from a rule-based approach in that both countenance the possibility that there is no bound on the possible differences between languages. (**hornstein2009**)

[3] See **rizzi2014** for a defense of the idea that microparameters are not merely rules under a different name and **boeckx2014** for a reply to Rizzi.

Michal Starke has made a similar observation:

> Thirty years ago, if some element moved in one language but not in another, a movement rule would be added to one language but not to the other. Today, a feature 'I want to move' ('EPP', 'strength', etc.) is added to the elements of one language but not of the other. In both cases (and in all attempts between them), variation is expressed by stipulating it. Instead of a theory, we have brute force markers. (**starke2014**)

6.3 There would have to be hundreds, if not thousands, of parameters

Tying parameters to functional categories was a strong conjecture in the 1980s, since there were so few generally recognized functional categories at the time. There were so few, in fact, that it was easy to believe that only a small number of parameters would be needed. Pinker (1994: 112), for example, speculated that there are just 'a few mental switches'. **lightfoot1999** suggested that there are about 30 to 40 parameters. For **adger2003n** 'There are only a few parameters'. **roberts2005** increased the presumed total to between 50 and 100. **fodor2001** was certainly correct when she observed that 'it is standardly assumed that there are fewer parameters than there are possible rules in a rule-based framework; otherwise, it would be less obvious that the amount of learning to be done is reduced in a parametric framework'. At this point in time, many hundreds of parameters have been proposed. Gianollo, Guardiano, and Longobardi (2008) propose 47 parameters for DP alone on the basis of 24 languages, only five of which are non-Indo-European, and in total representing only 3 families. **longobardi2011** up the total to 63 binary parameters in DP. As Cedric Boeckx has stressed: 'It is not at all clear that the exponential growth of parameters that syntacticians are willing to entertain is so much better a situation for the learner than a model without parameters at all' (**boeckx2014**).

One way to circumvent this problem would be to posit nonparametric differences among languages, thereby maintaining the possibility of a small number of parameters. Let us examine this idea now.

6.4 Nonparametric differences among languages undercut the entire parametric program

Are all morphosyntactic differences among languages due to differences in parameter setting? Generally that has been assumed not to be the case. Charles Yang was expressing mainstream opinion when he wrote that '… it seems highly unlikely that all possibilities of language variation are innately specified …' (**yang2011n**). From the beginning of the parameteric program it has been assumed that some features are extraparametric. Outside of (parametrically relevant) core grammar are:

> … borrowings, historical residues, inventions, and so on, which we can hardly expect to – and indeed would not want to – incorporate within a principled theory of UG. … How do we delimit the domain of core grammar as distinct from marked

> periphery? … [We] rely heavily on grammar-internal considerations and comparative evidence, that is, on the possibilities for constructing a reasonable theory of UG and considering its explanatory potential in a variety of language types … (Chomsky 1981: 8-9)

In other words, some language-particular features are products of extraparametric language-particular rules. Consider, for example, the treatment of Hixkaryana in **baker2001** based on an earlier proposal in **kayne1994** This language for the most part manifests OVS word order:

(3) Hixkaryana**derbyshire1985**
Kanawa yano toto
canoe took person
'The man took the canoe.'

One's first thought might be that what is needed is a parameter allowing for OVS order. But in fact Baker rejects the idea that a special word order parameter is involved here. Rather, he argues that Hixkaryana is (parametrically) SOV and allows the fronting of VP by a movement rule:

(4) S[OV] → [OV]S

In other words, in this account word order is determined *both* by a parameter and a language-specific rule.

It is quite implausible that every syntactic difference between languages and dialects results from a difference in parameter settings. Consider the fact that there are several dozen systematic morphosyntactic differences between the Norfolk dialect and standard British English (**trudgill2003**), most of which appear to be analytically independent. If each were to be handled by a difference in parameter setting, then, extrapolating to all of the syntactic distinctions in the world's languages, there would have to be thousands – if not millions – of parameters. That is obviously an unacceptable conclusion from an evolutionary standpoint, given that the set of parameters and their possible settings is, by hypothesis, innate. Furthermore, many processes that can hardly be described as 'marginal' have been assumed to apply in PF syntax (where the standard view, I believe, is that parameters are not at work), including extraposition and scrambling (**chomsky1995**); object shift (**holmberg1999**; **erteschik-shir2005**); head movements (**boeckx2001**); the movement deriving V2 order (**chomsky2001b**); linearization (i.e. VO vs. OV) (**chomsky1995**; **takano1996**; **fukui1998**; **uriagereka1999**); and even *Wh*-movement (**erteschik-shir2005**).

I think that it is fair to say that, after 35 years of investigation, nobody has a clear idea about which syntactic differences should be considered parametric and which should not be.[4] But one thing seems clear: If learners need to learn rules anyway, very little is gained by positing parameters.

[4] But see **smith2009n** for an interesting discussion of criteria for distinguishing parametric and nonparametric differences.

6.5 Parametric theory is arguably inherently unminimalist

There are a number of ways that the assumptions of the Minimalist Program have entailed a rethinking of parameters and the division of labor among the various components for the handling of variation. In one well-known formulation, 'FLN [= the faculty of language in the narrow sense – FJN] comprises only the core computational mechanisms of recursion as they appear in narrow syntax and the mapping to the interfaces' (**hauser2002**). Where might parameters fit into such a scenario? In one view, '… if minimalists are right, there cannot be any parameterized principles, and the notion of parametric variation must be rethought.' (**boeckx2011**). That is, given that the main thrust of the minimalist program is the reduction to the greatest extent possible of the elements of UG, there would seem to be no place for innately-specified parameters.

Despite the above, a great deal of work within the general envelope of the MP is still devoted to fleshing out parameters, whether micro or macro. For example, **yang2011n** writes that 'a finite space of parameters or constraints is still our best bet on the logical problem of language acquisition'. Note also that in many approaches, 'the mapping to the interfaces' encompasses a wide variety of operations. To give one example, 'UG makes available a set **F** of features (linguistic properties) and operations C_{HL} … that access F to generate expressions' (**chomsky2000**). In addition to features and the relevant operations on them, minimalists have attributed to the narrow syntax principles governing agreement, labelling, transfer, probes, goals, deletion, and economy principles such as Last Resort, Relativized Minimality (or Minimize Chain Links), and Anti-Locality. None of these fall out from recursion *per se*, but rather represent conditions that underlie it or that need to be imposed on it. To that we can add the entire set of mechanisms pertaining to phases, including what nodes count for phasehood and the various conditions that need to be imposed on their functioning, like the Phase Impenetrability Condition. And then there is the categorial inventory (lexical and functional), as well as the formal features they manifest. The question, still unresolved, is whether any of these principles, conditions, and substantive universals could be parameterized, in violation of the Strong Uniformity Thesis, but not of weaker proposals. If so, that would seem to allow for parametric variation to be manifested in the journey towards the interfaces.

7 Alternatives to the classic Principles-and-Parameters model

chomsky2005 refers to 'the three factors in language design', namely, genetic endowment, experience, and principles not specific to the faculty of language. The last-named 'third factor explanations', which include principles of data analysis and efficient computation, among other things, provide a potential alternative to the nonminimalist proliferation of parameters and their settings.[5]

[5] In what follows, I consider classical functional explanations of grammatical structure to be of the third factor type. It is not clear whether Chomsky shares that view.

The following subsections discuss alternatives to the classic P&P model, all appealing to one degree or another to third factor explanations. They are grounded approaches (§7.1), epigenetic (or emergentist) approaches (§7.2), and reductionist approaches (§7.3).

7.1 Grounded approaches

A grounded approach is one in which some principle of UG is grounded in – that is, ultimately derived from – some third factor principle. Along these lines, a long tradition points to a particular constraint, often an island constraint, and posits that it is a grammaticalized processing principle. One of the first publications to argue for grounded constraints was **fodor1978** where two island constraints are posited, one of which is the Nested Dependency Constraint (NDC):

(5) The Nested Dependency Constraint: If there are two or more filler-gap dependencies in the same sentence, their scopes may not intersect if either disjoint or nested dependencies are compatible with the well-formedness conditions of the language.

As Fodor noted, the processing-based origins of this constraint seem quite straightforward.

Another example is the Final Over Final Constraint (FOFC), proposed originally in **holmberg2000**

(6) Final Over Final Constraint: If α is a head-initial phrase and β is a phrase immediately dominating α, then β must be head-initial. If α is a head-final phrase, and β is a phrase immediately dominating α, then β can be head-initial or head-final.

As one consequence of the FOFC, there are COMP-TP languages that are verb-final, but there are no TP-COMP languages that are verb-initial. Holmberg and his colleagues interpret this constraint as the following UG principle:

(7) A theoretical reinterpretation of the FOFC: If a phase-head PH has an EPP-feature, then all the heads in its complement domain from which it is nondistinct in categorial features must have an EPP-feature. (**biberauer2008**)

walkden2009 points out that FOFC effects are accounted for by the processing theory developed in **hawkins2004** and hence suggests that (7) is a good example of a grounded UG principle.[6]

Note that neither Fodor nor Walkden have reduced the number of UG constraints; they have merely attributed the origins of these constraints to what in Chomsky's account would be deemeed a third factor. Naturally, the question arises as to whether these principles would need to be parameterized. The answer is 'apparently so', since the NDC does not govern Swedish grammar (**engdahl1985**) and the FOFC is not at work in Chinese (**chan2013**). In other words, grounded approaches, whatever intrinsic interest they might possess, do not prima facie reduce the number of UG principles and parameters.

[6] **mobbs2014** builds practically all of Hawkins's parsing theory into UG.

7.2 Epigenetic approaches

Let us turn now to 'epigenetic' or 'emergentist' approaches to variation, where parameters are not provided by an innate UG. Rather, parametric effects arise in the course of the acquisition process through the interaction of certain third factor learning biases and experience. UG creates the space for parametric variation by leaving certain features underspecified. There are several proposals along these lines, among which are **gianollo2008**; **boeckx2011** and **biberauer2014** (preceded by many papers by the same four authors). For reasons of space, I focus exclusively on **biberauer2014** In their way of looking at things, the child is conservative in the complexity of the formal features that it assumes are needed (what they call 'feature economy') and liberal in its preference for particular features to extend beyond the input (what they call 'input generalization' and which is a form of the superset bias). The idea is that these principles drive acquisition and thus render parameters unnecessary, while deriving the same effects. Consider first their word order hierarchy, represented in Figure 1:

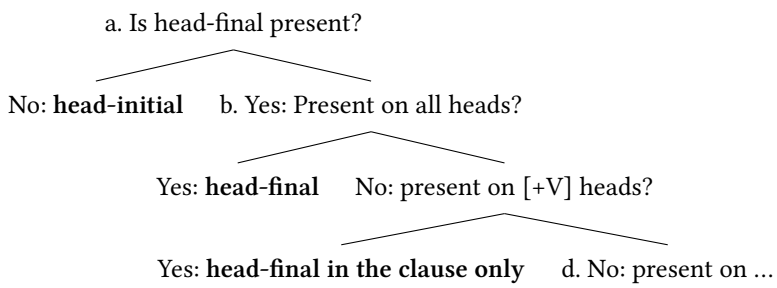

Figure 1: The Word Order Hierarchy of **biberauer2014**

To illustrate with a made up example, let's say that a language is consistently head-initial except in NP, where the noun follows its complements. However, there is a definable class of nouns in this language that do precede their complements and a few nouns in this language behave idiosyncratically in terms of the positioning of their specifiers and complements (much like the English word *enough*, which is one of the few degree modifiers that follows the adjective). In their theory, the child will go through the following stages of acquisition, zeroing in step-by-step on the adult grammar. First it will assume that all phrases are head-initial, even noun phrases. Second, it will assume that all NPs are head-final. Third, it will learn the systematic class of exceptions to the latter generalization, and finally, it will learn the purely idiosyncratic exceptions.

The other hierarchies proposed in Biberauer et al. are more complex and depend on many assumptions about the feature content of particular categories. Consider for example their null argument hierarchy and the questions posed by the child in determining the status of such arguments in its grammar (Figure 2).

There are many issues that one might raise about this acquisition scenario, the most important of which being whether children proceed from general to the particular, cor-

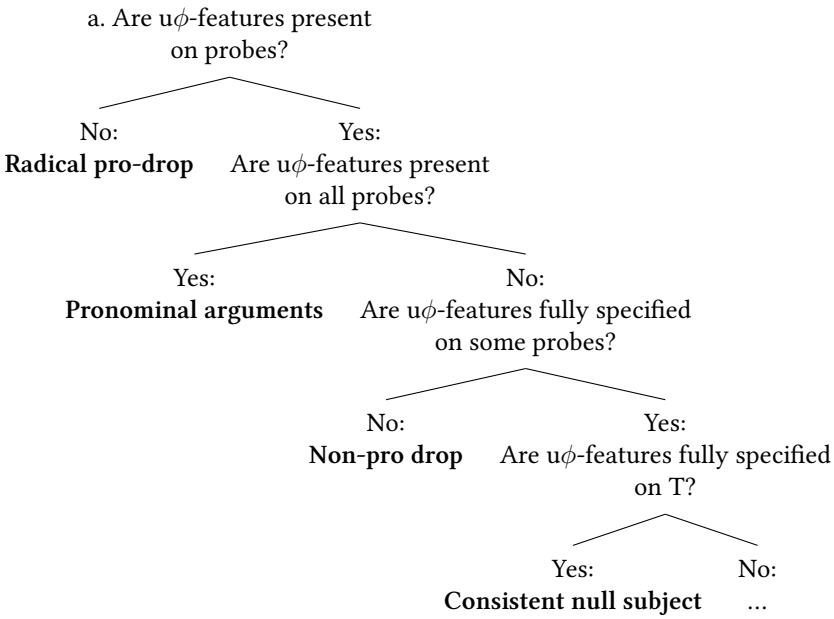

Figure 2: The Null Argument Hierarchy of **biberauer2014**

recting themselves as they go, and gradually zero in on the correct grammar. Indeed, many acquisitionists argue for precisely the reverse set of steps, in which children learn idiosyncratic items before broad generalizations. Theresa Biberauer herself acknowledges (p.c., November 21, 2014) that these steps can be overridden in certain circumstances. For example, the early stages might correspond to a pre-production stage or possibly acquirers pass through certain stages very quickly if counterevidence to an earlier hypothesis is readily available. And finally, as they note, frequency effects can distort the smooth transition through the hierarchies.

One might also ask how much the child has to know *in advance* of the progression through the hierarchy. For example, given their scenario, the child has to know to ask 'Are unmarked phi-features fully specified on some probes?' That implies a lot of grammatical knowledge. Where, one might ask, does this knowledge come from and how does the child match this knowledge with the input?[7]

Despite these unresolved questions, the Biberauer et al. approach presents a view that preserves the major insights of parametric theory without positing UG-based parameters. As such, it needs to be taken very seriously.

[7] An anonymous referee asks: 'To be honest, I don't see how the approach sketched here is different from parameter setting. Perhaps it's my own ignorance of Biberauer et al.'s proposal, but if the hierarchy of questions that the learner must address is innate, how does this differ from parameters that are innately specified?' As I understand their proposal, the hierarchy of questions falls out from general learning principles, though I am hazy on the details of precisely how.

7.3 Reductionist approaches and the need for language-particular rules

Reductionist approaches differ from epigenetic ones by reducing still further the role played by an innate UG in determining crosslinguistic variation. For example, returning to the FOFC, **trotzke2013** provide evidence that the best motivated account is to remove it entirely from the grammar, since, in their view, it can be explained in its entirety by systematic properties of performance systems. They also deconstruct the Head-Complement parameter in a similar fashion:

> [T]he physics of speech, that is, the nature of the articulatory and perceptual apparatus requires one of the two logical orders, since pronouncing or perceiving the head and the complement simultaneously is impossible. Thus, the head-complement parameter, according to this approach, is a third-factor effect. (**trotzke2013**)

Which option is chosen, of course, has to be built into the grammar of individual languages, presumably via its statement as a language-particular rule.

To take another example, **kayne1994** provided an elaborate UG-based parametric explanation of why rightward movement is so restricted in language after language. But **ackema2002** argue that the apparent ungrammaticality of certain 'right-displaced' syntactic structures should not be accounted for by syntax proper (that is, by the theory of competence), but rather by the theory of performance. In a nutshell, such structures are difficult to process. A necessary consequence of their approach is that it is necessary to appeal to language-particular rules to account for the fact that languages differ from each other in the degree to which displacement to the right is permitted.

The microparametric approach to variation is well designed to capture the fact that even closely related speech varieties can vary from each other in many details. The question is why one would want to appeal to microparameters when the traditional term 'rule' seems totally appropriate (see §6.2). The resistance to the idea of reviving the idea of (language-particular) rules is unsettling to some, perhaps because the idea of 'rules' brings back the ghosts of pre-generative structuralism, where it was believed by some that 'languages could differ from each other without limit' (**joos1957**), and the spectre of early transformational grammar, where grammars were essentially long lists of rules. But to call a language-particular statement a 'rule' is not to imply that *anything* can be a rule. Possible rules are still constrained by UG. That of course raises the question of what is in UG. An obvious candidate is the Merge operation or something analogous, but surely there must be a lot more than that. For example, it is hard to see how the broad architecture of the grammar could be learned inductively. Consider the fact that syntactic operations have no access to the segmental phonology: There is no language in which displacement – Internal Merge, if you will – targets only those elements with front vowels. It seems probable that this state of affairs derives from UG.

However, if the general thrust of the work of John A. Hawkins is correct (see **hawkins1994**; **hawkins2004**; **hawkins2014**), the major constraints on the nature of rules derive from the exigencies of language processing. No language has a rule that lowers a filler exactly two clauses deep, leaving a gap in initial position. Such a rule, while theoretically

possible, is so improbable (for processing reasons) that it will never occur. Norbert Hornstein's approach to variation, succinctly stated in the following passage, also stresses that it is not necessary to appeal to UG to explain why certain logically possible properties of grammars do not occur:

> There is no upper bound on the ways that languages might differ though there are still some things that grammars cannot do. A possible analogy for this conception of grammar is the variety of geometrical figures that can be drawn using a straight edge and compass. There is no upper bound on the number of possible different figures. However, there are many figures that cannot be drawn (e.g. there will be no triangles with 20 degree angles). Similarly, languages may contain arbitrarily many different kinds of rules depending on the PLD [= primary linguistic data – FJN] they are trying to fit. However, none will involve binding relations in which antecedents are c-commanded by their anaphoric dependents or where questions are formed by lowering a *Wh*-element to a lower CP. Note that this view is not incompatible with languages differing from one another in various ways. (**hornstein2009**)

In my view, the idea that a grammar is composed of language-particular rules constrained by both UG principles and third factor principles is an appealing vision that stands to inform research on crosslinguistic variation in the years to come.

8 Conclusion

Since the late 1970s, crosslinguistic variation has generally been handled by means of UG-specified parameters. On the positive side, thinking of variation in terms of parameterized principles unleashed an unprecedented amount of work in comparative syntax, leading to the discovery of heretofore unknown morphosyntactic phenomena and crosslinguistic generalizations pertaining to them. On the negative side, however, both macroparameters and microparameters have proven themselves to be empirically inadequate and conceptually nonminimalist. Alternatives to parameters are grounded approaches, epigenetic approaches, and reductionist approaches, the last two of which seem both empirically and conceptually quite promising.

Acknowledgements

Earlier versions of this paper were presented at the Division of Labor Conference in Tübingen in January 2015 and at a University of Illinois Colloquium in February 2015. I would like to thank Laurence Horn, Robert Borsley, and an anonymous referee for their comments on the pre-final manuscript.

Chapter 2

Phonological exceptionality is localized to phonological elements: the argument from learnability and Yidiny word-final deletion

Erich R. Round
University of Queensland

anderson2008 emphasizes that the space of possible grammars must be constrained by limits not only on what is cognitively representable, but on what is learnable. Focusing on word final deletion in Yidiny (**dixon1977a**), I show that the learning of exceptional phonological patterns is improved if we assume that Prince & Tesar's (**princetesar2004**) Biased Constraint Demotion (BCD) with Constraint Cloning (**pater2009r**) is subject to a Morphological Coherence Principle (MCP), which operationalizes morphological analytic bias (**moreton2008r**) during phonological learning. The existence of the MCP allows the initial state of Con to be simplified, and thus shifts explanatory weight away from the representation of the grammar per se, and towards the learning device.

I then argue that the theory of exceptionality must be phonological and diacritic. Specifically, I show that co-indexation between lexical forms and lexically indexed constraints must be via indices not on morphs but on individual phonological elements. Relative to indices on phonological elements, indices on morphs add computational cost for no benefit during constraint evaluation and learning; and a theory without indices on phonological elements is empirically insufficient. On the other hand, approaches which represent exceptionality by purely phonological means (e.g. **zoll1996**) are ill-suited to efficient learning. Concerns that a phonologically-indexed analysis would overgenerate (**gouskova2012**) are unfounded under realistic assumptions about the learner.

1 Exceptionality

What is the nature of representations which are passed from the morphology to the phonology? **anderson1992** demonstrates that the processes that create those representations can be elaborate and complex. Operations that act upon morphological forms, to realize units of morphologically-relevant meaning, involve not only the concatenation of

formatives, but also selection among alternatives and non-concatenative modifications to intermediate representations (see also **anderson2015**; **anderson2016**; **andersontoappearar**). However, what of the final result, which comprises some number of morphs that must then be interpreted phonologically? A constant concern of generative phonology since its inception has been to account adequately for patterned phonological exceptionality, the phenomenon in which segments in a restricted class of morphs exhibit phonologically distinctive behavior as triggers, targets or blockers of alternations, or as participants in exceptional featural, phonotactic or prosodic surface structures. For example, in Yidiny (**dixon1977a**; **dixon1977b**) vowels delete word-finally, if that deletion would prevent the word from surfacing with an unfooted syllable. This is seen in the root *gaɟara-* 'possum' in (1a) and the suffix *-ɲa* ACCUSATIVE in (1b), where feet are marked by parentheses. However, in a restricted set of morphs the final vowel behaves exceptionally, resisting deletion, as in the root *guɟara-* 'broom' (2a) and the suffix *-na* PURPOSIVE (2b).

(1) a. 'possum.ABS' /gaɟara/ → (ga ɟa:r)
b. 'father-ACC' /bimbi-ɲa/ → (bim bi:ɲ)

(2) a. 'broom.ABS' /guɟara/ → (gu ɟa:) ra
b. 'go-PURP' /gali-na/ → (ga li:) na

In order for the phonology to treat morph-specific, exceptional segments appropriately, it must receive from the morphology some kind of discriminating information which it can act upon. For much of the generative period it has been argued that this information is associated with morphs as a whole, and not with their individual phonological elements. Here I present an argument for the contrary view. The contribution, then, is to clarify the nature of one important aspect of the interaction between the morphological and phonological components of grammar. The principle line of evidence is learnability, namely the learnability of an optimality-theoretic grammar for phonological exceptionality. **anderson2008** has emphasized that the space of possible human grammars must be constrained not only by limits on what is cognitively representable, but also on what is learnable. The crux of the argument here relies not on specifics, but ultimately on general properties of learnable grammars, and thus I would hope should remain valid even as specific theories undergo refinement as they move closer to answering Anderson's (**anderson2008**) challenge.[1]

The chapter falls into two broad parts. In §2–§5 I discuss the processes and principles required to learn exceptionality. This leads to the positing of a Morphological

[1] A reviewer asks whether the machinery presented here is necessary if one assumes an exemplar-based model of phonology. I assume that learners do store rich, exemplar-like representations of linguistic experiences. However, natural language morphology in general has enough combinatorial complexity that reliance upon retrieved episodes will not be sufficient to reproduce the full range of creative behavior that humans display. Consequently some generative machinery is necessary, which performs not merely simple analogies and concatenations, but which can reproduce with precision the complex patterns generated by a realizational morphology such as Anderson's (**anderson1992**), and by a formal phonological grammar such as entertained here.

Coherence Principle in §6, which operationalizes a morphological bias that ensures successful learning for certain cases. In §7–§9 I am concerned with the underlying theory of these processes and principles. I evaluate two broad approaches to phonological exceptionality: PHONOLOGICAL approaches, which represent exceptionality as a property of individual segments (**bloomfield1939**; **kiparsky1982c**; **inkelas1995**; **zoll1996**), and MORPHOLOGICAL approaches which represent it as a property of morphs (**chomsky1964**; **chomskyhalle1968**; **zonneveld1978**; **pater2000r**). The result is an argument in favor of a DIACRITIC PHONOLOGICAL approach. On this account, exceptionality is represented at the level of individual phonological elements, not morphs; however the means of marking it is by diacritics which are visible to the phonology but not manipulable by it, in contradistinction to the CONCRETE PHONOLOGICAL approach, where the crucial representations are themselves phonological elements. As I show, the function of these '☒-indices' is essentially identical to 'M-indices' which would mark morphs, only there is no assumption that all exponents of a morph *m* be indexed identically. As we shall see, freedom from that assumption is both coherent theoretically and desirable, computationally and empirically. The discussion is illustrated throughout by the facts of word final deletion in Yidiny, to which we turn now in §2.

2 Word-final deletion in Yidiny

2.1 The phenomenon

Yidiny (**dixon1977a**) belongs to the Yidinyic subgroup of the Pama-Nyungan language family. Traditionally it was spoken in the rainforest region southwest of Cairns, in Northeastern Australia. Most examples below are from Dixon's (**dixon1977a**) detailed descriptive grammar; examples marked † are from Dixon's (**dixon1991**) dictionary and texts. An inventory of underlying segments is in Table 1.

Table 1: Yidiny underlying segments, after **dixon1977a**

	Labial	Apical	Laminal	Dorsal
Stop	b	d	ɟ	g
Nasal	m	n	ɲ	ŋ
Lateral, trill		l, r		
Approximant	w	ɻ	y	
Vowels	i, a, u, iː, aː, uː			

Syllable shapes are tightly constrained. Onsets are obligatory and simple. Codas permit only sonorants other than /w/. Codas in word-final position are simple; word-internal codas also permit disegmental, continuant–nasal sequences. Morphologically, the language is almost entirely suffixing and largely agglutinative. Roots are minimally

disyllabic and suffixes are maximally disyllabic (**dixon1977a**). An online appendix[2] discusses the morphological constituency of verbal inflection.

Of Yidiny's phonological alternations, those to receive the greatest attention have been stress placement, vowel length and to a lesser extent, word-final deletion (**dixon1977a**; **dixon1977b**; **hayes1982r**; **hayes1985**; **kager1993**; **crowhurst1995**; **halle1995**; **hall2001**; **pruitt2010**; **hyde2012**; **bowern** *inter alia*). Yidiny's stress and length alternations in particular have featured in significant theoretical works on meter and prosody over the past four decades, and both are nontrivial topics in themselves. Word-final deletion, however, can be studied largely independently of them for reasons that follow.

Although stress placement in Yidiny has proven contentious (**pruitt2010**; **bowern**), word-final deletion is not sensitive to stress *per se*, but rather only to the position of foot boundaries. These have been uncontroversial since their analysis by **hayes1982r** feet in Yidiny are disyllabic and left-aligned within the phonological word.

Many words with word-final deletion also exhibit vowel lengthening; however the phenomena show little to no mutual interaction. In a rule-based theory permitting simultaneous application (**anderson1974**) lengthening and deletion would apply simultaneously; neither rule feeds or bleeds the other.[3] See **round** for an analysis of Yidiny lengthening.

Word-final deletion is sensitive to foot placement, and foot placement is sensitive to phonological word boundaries. In Yidiny, phonological words commence at the left edge of each root and each disyllabic suffix (**dixon1977a**).[4] Phonological words therefore begin with either a polysyllabic root or a disyllabic suffix and are followed by zero or more monosyllabic or entirely consonantal suffixes. Word-final deletion targets unfooted syllables and therefore only affects prosodic words which, *modulo* deletion, would be at least trisyllabic. As a consequence, we are interested here in three kinds of phonological word: those comprised of bare roots of three or more syllables; those comprised of roots plus one or more monosyllabic suffixes; and those comprised of a disyllabic suffix plus one or more additional, monosyllabic suffixes. The third kind is rare,[5] and so discussion will focus on the first two.

Word-final deletion applies only if the word thereby avoids surfacing with an unfooted syllable. For example, the roots *gindanu-* 'moon' and *gubuma-* 'black pine' both contain three vowels, each of which is a potential syllabic nucleus at the surface. In (3a, 4a) they have undergone deletion of their final vowel to prevent it from surfacing in an unfooted syllable; compare (3b, 4b) where the roots are non-final in the word, and the final vowels surface.

(3)	a.	'moon[ABS]'	/gindanu/	→	(gin da:n)
					*(gin da:) nu
	b.	'moon-ERG'	/gindanu-ŋgu/	→	(gin da) (nuŋ gu)

[2] Available from 10.6084/m9.figshare.4579696

[3] Deletion counter-bleeds lengthening, thus in a strictly serial analysis lengthening would precede word-final deletion (**dixon1977a**; **dixon1977b**; **hayes1985**; **crowhurst1995**).

[4] Yidiny's only prefix, [ɟa:-] 'in a direction' occupies its own phonological word (**dixon1977a**).

[5] For an illustration, see example (25).

(4) a. 'black pine[ABS]' /gubuma/ → (gu bu:m)
*(gu bu:) ma
b. 'black pine-PURP' /gubuma-gu/ → (gu bu) (ma gu)

Final vowel deletion may also affect suffixes. In (5a,c, 6a), the vowels of the nominal comitative suffix *-yi* and verbal comitative suffix *-ŋa* have undergone deletion, thereby preventing the surfacing of an unfooted syllable. In (5b, 6b) the suffixes are non-final in the word, and the vowel surfaces.

(5) a. 'woman-COM' /buɲa-yi/ → (bu ɲa:y)
*(bu ɲa:) yi
b. 'woman-COM-ERG' /buɲa-yi-ŋgu/ → (bu ɲa) (yiŋ gu)
c. 'black bream-COM' /gulugulu-yi/ → (gu lu) (gu lu:y)
*(gu lu) (gu lu:) yi

(6) a. 'come-COM[IMP]' /gada-ŋa/ → (ga da:ŋ)
*(ga da:) ŋa
b. 'come-COM-PST' /gada-ŋa-lɲu/ → (ga da:) (ŋal ɲu)

Word-final deletion interacts with restrictions on word-final consonants, and the interaction plays out differently in roots versus suffixes. In roots, deletion will fail to apply if the result would be an illicit word-final coda, containing either a stop or /w/ (7) or a cluster (8). One conceivable alternative, to also delete the consonant, is not attested in roots (7, 8).[6]

(7) a. 'man[ABS]' /waguɟa/ → (wa gu:) ɟa
*(wa guɟ)
*(wa gu:)
b. 'dog[ABS]' /gudaga/ → (gu da:) ga
*(gu da:g)
*(gu da:)
c. 'sugar ant[ABS]' /balawa/ → (ba la:) wa
*(ba la:w)
*(ba la:)
d. 'place name[ABS]' /ŋalumba/ → (ŋa lu:m) ba
*(ŋa lu:mb)
*(ŋa lu:m)

(8) 'warn[IMP]' /binarŋa/ → (bi na:r) ŋa
*(bi na:rŋ)
*(bi na:r)

[6] Neither Dixon's grammar (**dixon1977a**) nor dictionary (**dixon1991** which cites underlying forms) records a surface form for the roots in 7c and 7d, or for roots illustrating the same pre-final consonant or comparable consonant clusters. However, **dixon1977a** specifically reports that the roots *balawa-* and *gindalba-* do not undergo deletion; the surface forms provided here are what we would expect if this is so.

In contrast, deletion in suffixes applies not only to the final vowel, but also to a single consonant that precedes it, if that consonant would be illicit word-finally, as in (9). This form of CV deletion respects phonotactic constraints while also avoiding unfooted syllables.[7]

(9)	a.	'grey possum-ERG'	/margu-ŋgu/	→	(mar gu:ŋ)
	b.	'see-PST'	/wawa-lɲu/	→	(wa wa:l)
	c.	'warn-DAT.SUB'	/binarŋa-lɲu-nda/	→	(bi nar) (ŋal ɲu:n)

However, word-final deletion never deletes the initial segment of a suffix (and consequently, it will never delete an entire suffix), as illustrated in (10).

(10)	a.	'woman-SET'	/buɲa-ba/	→	(bu ɲa:) ba
					*(buɲ ba)
	b.	'bandicoot-GEN'	/guygal-ni/	→	(guy ga:l) ni
					*(guy ga:ln)
					*(guy ga:l)

Deletions do not occur word internally (11a,b), nor do word-final, licit codas delete (11b). All Yidiny roots and suffixes that are consonant-final end underlyingly with licit coda consonants, so no morph undergoes spontaneous deletion of an underlyingly-final consonant (11c).

(11)	a.	'woman-SET'	/buɲa-ba/	→	(bu ɲa:) ba
					*(buɲ ba)
	b.†	'name[ABS]'	/bagiram/	→	(ba gi:) ram
					*(ba gi:rm)
					*(ba gi:r)
	c.		*/bagirag/	→	*(ba gi:r)

To summarize, word-final deletion applies only so as to avoid the surfacing of unfooted syllables. It may delete the final vowel from a root and the final (C)V sequence from a suffix, but will not delete a suffix-initial segment. Deletion is blocked (in roots) or expanded (in suffixes, from V deletion to CV deletion) in order to obey phonotactic restrictions on word-final codas. These are the regular conditions under which word-final deletion occurs.

In addition to its regular application, Yidiny contains roots and suffixes which are exceptional non-undergoers of word-final deletion. In (13), the non-undergoer roots *mulari-*, *guɟara-*, *ɟudulu-*, *baŋgamu-* all resist word-final deletion despite their pre-final consonant being permissible as a coda, and despite the fact that the consequence is an unfooted, word-final syllable.

[7] The 'dative subordinate' is marked by what **round2013** has called 'compound suffixation', comprising two monosyllabic suffixal morphs, /-lɲu; -nda/. That these are not a single, disyllabic suffix is evident in the fact that they fail to be parsed into a their own phonological word, separate from the root.

(12) a. 'initiated man[ABS]' /mulari/ → (mu la:) ri
*(mu la:r)
b. 'broom[ABS]' /guɟara/ → (gu ɟa:) ra
*(gu ɟa:r)
c. 'brown pigeon[ABS]' /ɟudulu/ → (ɟu du:) lu
*(ɟu du:l)
d. 'potato[ABS]' /baŋgamu/ → (baŋ ga:) mu
*(baŋ ga:m)

dixon1977a reports 115 trisyllabic roots whose phonotactic shape would, under regular conditions, expose them to word-final deletion. Of these, 34, or around 30%, are exceptional non-undergoers. The distinction is idiosyncratic; neither **dixon1977a** nor subsequent researchers have found any phonological, semantic or grammatical factor that categorically determines whether a root will be a non-undergoer.[8]

Suffixes also may be exceptional non-undergoers. In (17) the non-undergoer suffixes *-nda*, *-lɟi* and *-na* resist word-final deletion and allow an unfooted syllable to surface. Avoidance of regular, word-final CV deletion is seen in (13a,b) and V deletion in (13c).

(13) a. 'grey possum-DAT' margu-nda → (mar gu:n) da
*(mar gu:n)
b. 'see-LEST[ABS]' wawa-lɟi → (wa wa:l) ɟi
*(wa wa:l)
c. 'go-PURP' gali-na → (ga li:) na
*(ga li:n)

Tables 2 and 3 list all suffixal allomorphs in Yidiny which, on phonotactic grounds, could plausibly delete.[9] Regular undergoers are in Table 2 and non-undergoers in Table 3.

[8] Historically speaking, borrowed forms may account for many of these items (Barry Alpher p.c.); synchronically, however, their motivation is opaque.

[9] Such suffixes must be vowel-final and monosyllabic. If just the final vowel is to delete, then it must leave behind a single, licit-coda consonant in word final position. This will require the suffix to be -CV, and be preceded by a vowel, not a consonant. Alternatively, if the final CV is to delete, then the suffix must be -CCV, since suffix-initial segments do not delete, and it too must attach to a vowel-final stem. Data here is from a comprehensive search of **dixon1977a** in which relevant information can be found on pp.50–54, 151. 'Emphatic' *-ɲa* (**dixon1977a**) is excluded. It behaves as a phonological clitic that occupies a distinct phonological word, and does not undergo final deletion.

Table 2: Monosyllabic suffixes which undergo word-final deletion.

Function		**-CV**	**-CCV**
Case	ERGATIVE		-ŋgu
	LOCATIVE	-la	
	ACCUSATIVE	-ɲa	
	COMITATIVE	-yi	
	GENITIVE	-ni, -nu	
Verbal	PAST tense inflection	-ɲu	-lɲu, -ɻɲu
	COMITATIVE derivation	-ŋa	
	DATIVE SUBORDINATE inflection		-nda[10]

Table 3: Monosyllabic suffixes which escape word-final deletion.

Function		**-CV**	**-CCV**
Case	DATIVE		-nda
Verbal	PURPOSIVE inflection	-na	-lna, -ɻna
	LEST nominalizing derivation		-nɟi, -lɟi, -ɻɟi

Exceptional non-undergoers, both roots and suffixes, only block the deletion of their own segments; the exceptionality does not spread to neighboring morphs. Accordingly in (14), the exceptional non-undergoer LEST does not block deletion in the following, regular undergoer, ERGATIVE suffix.

(14) wiwi-ːɟi-nɟi-ŋgu ‘give-ANTIP-LEST-ERG’ → (wi wiː) (ɟin ɟiːŋ)
*(wi wiː) (ɟin ɟiːŋ) gu

Likewise, the presence of a regular undergoer will not undo the blocking effect of an exceptional non-undergoer. In (15) the regular undergoer COMITATIVE does not undermine the blocking of deletion in the exceptional non-undergoer PURPOSIVE, which follows it.

(15) maɟinda-ŋa-lna ‘walk up-COM-PURP’ → (ma ɟin) (da ŋaːl) na
*(ma ɟin) (da ŋaːl)

It will be recalled that roots in Yidiny can undergo word-final deletion of vowels, but not of the consonants that precede them. More specifically, roots that end in CCV do not delete final CV, whereas some suffixes do, and nor does final C′V delete from roots that end in VC′V, where C′ is an impermissible coda. Two conceivable accounts for this

[10] The dative subordinate is marked by a string of two monosyllabic suffixes *-lɲu-nda*, cf. fn.7.

may be distinguished. On one account, the grammar of Yidiny expressly prohibits root-final CV deletion. On the other, it happens just by chance that all CCV-final and VĆV-final roots are exceptional non-undergoers. On the latter account, the grammar WOULD enforce CV deletion from roots, if only the lexicon provided the right inputs; on the former account it would not. The level of empirical support for these hypotheses can be assessed statistically. Table 4 compares counts of CCV- and VĆV-final roots and CCV-final suffixes which either do or do not delete. The distribution is strongly unbalanced, and we can reject with confidence the null hypothesis that it is due to chance ($\chi^2_{df=1}$ = 47.9 $p < 10^{-10}$). Table 5 compares counts of roots that are CCV- and VĆV-final with those that are VCV-final, i.e., where C is a permissible coda. Again, the counts are highly unbalanced and we reject the hypothesis that the absence of deletion in CCV- and VĆV-finals is by chance ($\chi^2_{df=1}$ = 125.8. $p < 10^{-10}$). The only empirically-supported conclusion is that the lack of consonant deletion in Yidiny roots is systematic, not due to chance. A satisfactory formal analysis should reflect this.[11]

Table 4: Deletion of coda-ilicit pre-final C in roots versus suffixes.

	CCV- and VĆV-final roots	**CCV-final suffixes**
No deletion	116	6
Deletion	0	4

Table 5: Deletion in roots with pre-final coda-illicit C versus prefixal coda-licit C.

	CCV- and VĆV-final roots	**VCV-final suffixes**
No deletion	116	34
Deletion	0	81

2.2 Constraint rankings

A briefly sketch now follows of how the facts above would be analysed in OT. Foot placement in Yidiny is due to FOOTBINARITY ≫ PARSESYLLABLE ≫ ALIGN(FT,L,PRWD,L) (**princesmolensky2004**[1993], **mccarthy1993 mccarthy1995r**). Of these, only PARSESYLLABLE (PRS) will be of interest for our purposes; I assume that other prosodic constraints are satisfied optimally. Absolute restrictions against obstruents and

[11] As a reviewer observes, there is an interesting historical background to be clarified here, and an account of it is planned. Naturally, the object of a synchronic analysis differs ontologically from that of a historical one. The two are complementary, but neither account would substitute for or serve as a counter-analysis to the other.

/w/ in codas are due to SONORANT/CODA (e.g. **lombardi2002**) and *w/CODA; I assume these are unviolated.

Regular word-final deletion in Yidiny can be analysed straightforwardly by ranking PRS ≫ MAXIMALITY (MAX, **mccarthy1995r**). This causes deletion of final vowels in preference to the surfacing of unfooted syllables, but not if an illicit coda results.

Segments may delete from the right edge of the word only, not the left or word-internally. High-ranking ANCHOR-LEFT(morph) penalizes deletion from the left edge of any morph and CONTIG-IO(PRWD) penalizes deletion internally (**mccarthy1995r**).

Yidiny permits complex codas word-internally, but not word-finally. Ranking CNTG ≫ *COMPLEXCODA (**bernhardt1998**) accounts for this; ranking both above PRS accounts for the absence of deletion after pre-final clusters in roots and the defeat of candidates which delete only a final vowel from word-final CCV suffixes.

Word-final deletion applies differently to roots and suffixes. Roots will not undergo consonant deletion, even if the consequence is an unfooted syllable. The ranking of undominated MAX-C/ROOT (**mccarthy1995r**) above PRS accounts for this. Suffixes do not violate MAX-C/RT and consequently are free to undergo consonant deletion, however highly-ranked ANC penalizes the deletion of morph-initial segments. This accounts for the fact that a consonant may delete from a -CCV suffix but not from -CV.

At this point, regular word-final deletion occurs whenever satisfaction of the markedness constraint PRS requires the violation of the lower-ranked faithfulness constraint MAX. Deletion is blocked unexceptionally whenever PRS itself is violated in order to satisfy higher-ranking constraints, which are of two kinds: those which penalize marked codas, SON/CODA, *w/CODA, *CPLX; and those which penalize deletion in specific morphological contexts, namely at left edges of morphs, ANC, and consonants in roots, MAX-C/RT. We see that the driver of word-final deletion in Yidiny is the ranking of PRS ≫ MAX. Deletion occurs when PRS is satisfied but MAX is not. Regular blocking results when PRS must be violated, in which case MAX can be satisfied.

Exceptional non-undergoers avoid deletion. For them, MAX is always satisfied, even at the expense of PRS. Consequently, while regular undergoers are subject to a ranking of PRS ≫ MAX, exceptional non-undergoers must be subject to MAX ≫ PRS. In §4 I consider two approaches that will ensure this is the case, one morphological and one phonological. First though, a remark about constraint violations.

3 Relativized constraint violation

I introduce here a simple expression for relating the violations of certain pairs of constraints, which will aid discussion in later sections.

For any constraint C and candidate CAND, there will be zero or more violations of C. Given the definition of C, those violations will be due to certain parts, or loci, in CAND, either in the output of CAND or in the correspondences between input and output elements (**mccarthy1995r**). We can define the set of LOCI OF VIOLATION, V(C, CAND), as the loci in CAND which cause violations of C (**mccarthy2003 lubowicz2005**). Now, some pairs of constraints C_1, C_2 are related such that for any CAND, the loci of violation of C_2

are a subset of the loci of violation of C_1. In many cases, the latter are precisely those members of the former which also contain some particular kind of phonological element. For example V(MAX-C, CAND) are those members of V(MAX, CAND) which also contain input consonants. In that case, we can express V(C_2, CAND) terms of the INTERSECTION of the set V(C_1, CAND) and some appropriately defined second set, that picks out loci containing the criterial elements. Let us define the set of 'φ-loci', L_φ(D(φ), CAND), as the set of loci in CAND that contain a phonological element φ of the kind denoted by predicate D(φ). For example, V(MAX-C, CAND) can be defined in relative terms, as in (16), where the predicate INPUT_CONSONANT(φ) denotes input consonants. (For brevity I omit the 'CAND' from the expression for each set.)

(16) V(MAX-C) $=_{def}$ V(MAX) $\cap$ L_φ(INPUT_CONSONANT(φ))

This relativized method will be used below to define new constraints C_N in terms of a reference constraint, C_R, and a set of phonological elements which restrict the violations of C_N relative to those of C_R.

4 Preliminary analysis of word-final deletion

4.1 A morphological approach

We now consider an OT implementation of the morphological approach to Yidiny exceptionality, using lexically indexed constraints (**pater2000r**; **pater2006**; **pater2009r**). A lexically indexed constraint C_M behaves precisely like its unindexed counterpart, C, except that it can be violated only by structures which contain exponents of a specific set *M* of morphs, each of which has been assigned a diacritic mark which I will term a LEXICAL M-INDEX, that co-indexes it to C_M. The definition can be expressed relatively as in (17), following a similar formulation by **finley2010**

(17) V(C_M) $=_{def}$ V(C) $\cap$ L_φ($m \in M$ & EXP(φ, m)), *where*:
M is the set of morphs co-indexed to C_M.
EXP(φ, m) states that element φ is an exponent of morph m

If we now define two sets of Yidiny morphs, *U* the set of regular undergoers of word-final deletion, and *N* the set of exceptional non-undergoers, then either of the rankings in (18) will ensure that the correct sets of morphs is subject to the desired partial ranking of PRS and MAX.

(18) a. $PRS_U \gg MAX \gg PRS$
b. $MAX_N \gg PRS \gg MAX$

In (18a), all phonological exponents of undergoer morphs will be subject to $PRS_U \gg$ MAX, and non-undergoers to MAX $\gg$ PRS. In (18b), all phonological exponents of non-undergoer morphs will be subject to $MAX_N \gg$ PRS, and undergoers to PRS $\gg$ MAX. For

now I will use ranking (18a); the reason for this will become clear in §5.[12,13] Examples in (19a–b) illustrate word-final deletion of regular undergoers which are indexed *U*, the root *malanu-*$_U$ and suffix ERGATIVE *-ŋgu*$_U$, while (19c–d) show the absence of deletion for exceptional non-undergoers *mulari-* 'initiated man' and DATIVE *nda*.

(19)

		PRS-$_U$	MAX	PRS
a.	/malanu$_U$/ 'right hand[ABS]' (ma la:n) ≻ (ma la:) nu	W	L	W
b.	/margu-ŋgu$_U$/ 'grey possum-ERG' (mar gu:ŋ) ≻ (mar gu:ŋ) gu	W	L	W
c.	/mulari/ 'initiated man[ABS]' (mu la:) ri ≻ (mu la:r)		W	L
d.	/margu-nda/ 'grey possum-DAT' (mar gu:n) da ≻ (mar gu:n)		W	L
e.	/maɟinda-ŋa$_U$-lna/ 'walk up-COM-PURP' (ma ɟin) (da ŋa:l) na ≻ (ma ɟin) (da ŋa:l)		W	L

Example (19e) illustrates the fact that violations of PRS$_U$ require not merely the presence of a *U*-indexed morph in the word, but a locus of violation which contains a phonological exponent of a *U*-indexed morph (17). Namely, the final syllable of (19e), *na*, is unfooted. However since that syllable contains no phonological exponent of a *U*-indexed morph, no violation of PRS$_U$ results. This is true despite the presence of a *U*-indexed morph elsewhere in the word.

4.2 A phonological approach

The phonological approach correlates the (un)exceptionality of a segment with representational properties of the segment itself. Implementations differ as to which property is used. **zoll1996** analyses segments which resist deletion as having root nodes in their input, whereas segments that delete more readily lack root nodes, and are termed SUBSEGMENTS. Under these assumptions, a ranking MAX(SEG) ≫ PRS ≫ MAX(SUBSEG) ensures that segments with input root nodes are subjected to MAX(SEG) ≫ PRS, while those without are subjected to PRS ≫ MAX(SUBSEG).[14] Examples are in (20), where segments without root nodes are underlined.

[12] Briefly, procedures for learning OT grammars improve in performance if they opt to rank markedness higher than faithfulness when given a choice. Consequently the ranking in (18a) will be learned in preference to (18b); see §5.

[13] An early proposal that only faithfulness constraints be indexable (**benua1997**; **ito1999**; **fukazawa1999**) has proven untenable (**pater2000r**; **pater2006**; **flack2007r**; **flack2007li**; **inkelas2007**; **gouskova2007**; **mahanta2008**; **jurgec2010**).

[14] Assuming undominated *FLOAT (**myers1997**), which prohibits surface subsegments, and low-ranked DEP(ROOT) (**zoll2001**).

(20)

		Max(Seg)	Prs	Max(Subseg)
a.	/malanu/ (ma la:n) ≻ (mu la:) nu		W	L
b.	/marguŋgu/ (mar gu:ŋ) ≻ (mar guŋ) gu		W	L
c.	/mulari/ (mu la:) ri ≻ (mu la:r)	W	L	W
d.	/margu-nda/ (mar gu:n) da ≻ (mar gu:n)	W	L	W

I wish to draw a distinction now between two conceivable kinds of phonological analysis. A concrete phonological analysis represents exceptionality using regular phonological material, such a features, root nodes and prosodic units, or perhaps their absence. An abstract phonological analysis uses diacritic lexical indices, which I will term lexical ☒-indices, on segments, much like the morphological analysis uses lexical M-indices on morphs. Some objections which have been raised to phonological analyses are specific to the concrete approach. These include doubts over whether sufficiently many concrete phonological contrasts would be available in languages with very many exceptional patterns (**gouskova2012**), and concerns over whether learners can choose between multiple, alternative concrete representations (**kiparsky1973r pater2009r**). I will set these concrete-specific concerns aside for now, and instead assume an abstract phonological approach. I return to the concrete approach in §9, where I argue on independent grounds that it is poorly adapted to efficient learning.

Accordingly, I will use lexical ☒-indices *u* and *n* to index undergoer and non-undergoer segments respectively, and define ☒-indexed constraints, $C_{☒}$, in relative terms as in (21).

(21) $V(C_{☒}) =_{def} V(C) \cap L_{\varphi}(\varphi \in ☒)$, *where:*
☒ is the set of phonological elements co-indexed to $C_{☒}$.

Returning to the phonological account of Yidiny exceptionality, a constraint ranking Max-*n* ≫ Prs ≫ Max, or Prs-*u* ≫ Max ≫ Prs, will be sufficient for our purposes. Tableau (22) shows examples using the latter ranking; *u*-indexed segments are underlined.

(22)

		Prs-*u*	Max	Prs
a.	(ma la:n) ≻ (mu la:) nu	W	L	W
b.	(mar gu:ŋ) ≻ (mar guŋ) gu	W	L	
c.	(mu la:) ri ≻ (mu la:r)		W	L
d.	(mar gu:n) da ≻ (mar gu:n)		W	L
e.	(ma ɟin) (da ŋa:l) na ≻ (ma ɟin) (da ŋa:l)		W	L

A recent criticism of the phonological approach to exceptionality in OT is that it overgenerates (**gouskova2012**). Adapting Gouskova's arguments to the facts of Yidiny: if we adopt the ranking Prs-*u* ≫ Max ≫ Prs, then it is no longer necessary to assign a high ranking to the morphologically-sensitive constraints Anc and Max-C/Rt, which penalize the deletion of morph-initial segments and root consonants. Rather, so long as all morph-initial segments and all root consonants lack a lexical *u*-index, then by virtue of the partial ranking Max ≫ Prs, they will resist deletion irrespective of the ranking of Anc and Max-C/Rt. By the same token however, if Anc and Max-C/Rt do receive a low ranking, then the analysis will fare poorly in the context of Richness of the Base (**princesmolensky2004**-1993]), since without high-ranked Anc and Max-C/Rt ensuring that morph-initial and root-consonant deletion is impossible, there is nothing to prevent segments from deleting in those positions if they are *u*-indexed in the lexicon. For example, a root such as **binarŋa* could undergo CV deletion; a suffix **-ni* could delete entirely; and **mulari* could delete from the left, thereby failing to capture the generalization that the absence of such forms is not an accident of the lexicon, but a systematic property of the grammar. This is perhaps the most significant apparent flaw of the phonological approach: it fails to rule out unattested patterns. This is in contrast to the morphological approach, which does rule them out. Or at least, so it would seem. In §5 I show that the true situation can be otherwise, once learning is taken into account.

4.3 Alternatives

Before proceeding to learning, I mention two OT alternatives to the analysis of exceptionality in Yidiny word-final deletion.

Co-phonological approaches handle exceptionality as a type of cyclicity effect (**orgun1996 kiparsky2000r inkelas2007 bermudez-Otero2016**). On each morphological cycle the result of a morphological operation is submitted to an appropriate phonological subgrammar, of which the language may possess many. Problematic for any cyclicity-based approach to exceptionality in Yidiny word-final deletion is that the Yidiny case is non-cyclic. Instead, undergoers are subject to deletion only if word-final. For example, in building both words in (23a,b) the first step would be to introduce the undergoer root *bigunu-* 'shield'. However at that point, the 'deleting' subgrammar should only be applied if the root will end up word final, as in (23a) but not in (23b).

(23)	a.	'shield[Abs]'	/bigunu/	→	(bi guːŋ)
	b.	'shield-comit-erg'	/bigunu-yi-ŋgu/	→	(bi gu) (nu yiːŋ)
					*(bi gun) (yiŋ gu)

Selecting the correct subgrammar in (23) thus requires information about the next step in the derivation. Crucially though, it requires forewarning, not only of whether or not there is more morphology to come, but also of what the phonological ramifications will be. This is because the relevant domain for word final-deletion in Yidiny is not the morphological word but the prosodic word. For example, in (24) the roots *gaɟula-* 'dirty' and *gumaɻi-* 'red' are followed by suffixes. Since the suffixes are monosyllabic, just one

prosodic word results and the roots are non-final in their prosodic word. In (25) however, the roots are followed by the disyllabic INCHOATIVE suffix *daga*, which commences a second prosodic word. As a consequence, the roots are final in their prosodic word and deletion is possible: the undergoer *gaɟula-* deletes while the non-undergoer *gumaɻi-* does not.

(24) a. 'dirty-CAUS-PST'
/gaɟula-ŋa-lɲu/ → [(ga ɟu) (la ŋa:l)$_{PWd}$]
b. 'red-CAUS-PST'
/gumaɻi-ŋa-lɲu/ → [(gu ma) (ɻi ŋa:l)$_{PWd}$]

(25) a. 'dirty-INCHO-PST'
/gaɟula-daga-ɲu/ → [(ga ɟu:l)$_{PWd}$] [(da ga:ɲ)$_{PWd}$]
b. 'red-INCHO-PST'
/gumaɻi-daga-ɲu/ → [(gu ma:) ɻi $_{PWd}$] [(da ga:ɲ)$_{PWd}$]

Any cyclic, look-ahead mechanism in Yidiny would therefore need to know how the word would be prosodically parsed on the NEXT cycle, before it can decide whether or not to apply the 'deleting' subgrammar on the current cycle. The look-ahead mechanism would therefore require the power of a subgrammar itself, yet if the theory were augmented in this manner, then other core mechanisms such as scope, or 'bracket erasure', effects (**inkelas2007**) would be undermined. I conclude that co-phonology theory as it stands cannot analyse exceptionality in Yidiny word-final deletion.

Another approach would be to lexically list two allomorphs for all undergoer morphs in the language, and have the grammar select them either optimally (**mester1994 kager1996 mascaro1996** and **tranel1996a**; **tranel1996b**) or with some degree of stipulation (**bonet2007**; **round2013**; **wolf2015**). On this approach, 'deletion' is apparent only, due in reality to the selection between two input allomorphs, one of which contains only a subset of the segments in the other (for a proposal not unlike this for Yidiny, see **hayes1997**). An example is shown in (26), where the grammar optimally selects between two input allomorphs of the undergoer root *bigunu-* 'shield'.

(26)

{/bigunu/, /bigun/} 'shield[ABS]'		ANC	MAX-C/RT	PRS	MAX
a. ☞	/bigun/ :: (bi gu:n)				
b.	/bigun/ :: (bi gu:) nu			∗W	
c.	/bigunu/ :: (bi gu:n)				∗W
d.	/bigunu/ :: (bi gu:) nu			∗W	

Two objections can be raised. First, because the approach simply lists alternant pairs, it misrepresents their resemblances as accidents, rather than relating them systematically. Relatedly, in the context of Richness of the Base, the analysis would allow the apparent deletion of morph-initial and -medial segments as well as root consonants, by leaving them out of an underlying allomorph, in a pair such as {/bigunu/, /gunu/}. Ranking ANC and MAX-C/ROOT highly would not ameliorate the problem, as shown in (27).

(27)

	{/bigunu/, /gunu/}	Anc	Max-C/root	Prs	Max
a. ☞	/gunu/ :: (gu nu)				
b.	/bigunu/ :: (bi gu:) nu			∗W	

Second, it is unclear how the analysis would prevent apparent deletion in word medial positions in the event that it is optimsing, as in (28), where the true output *buɟala-ŋa:-lna* violates Prs while the more optimal false winner **buɟal-ŋa-lna* does not. The constraint Cntg will not prevent this occurring.

(28)

	†{/buɟala, buɟal/}-ŋa-lna/ 'finely ground-cause-purp'	Cntg	*Cplx	Prs	Max
a. ☞	/buɟala-ŋa-lna/ :: (bu ɟa) (la ŋa:l) na			∗L	
b. ∗	/buɟal-ŋa-lna/ :: (bu ɟal) (ŋal na)				

I conclude that neither the co-phonological approach nor the allomorph-selection approach offers a viable alternative for Yidiny word-final deletion.

5 Learning exceptionality

5.1 Biased Constraint Demotion

I turn now to consider how exceptionality is, or isn't, learned. After introducing Prince and Tesar's (**princetesar2004**) Biased Constraint Demotion (BCD) algorithm and adaptations of it for the learning of indexed constraints, I show that the learning of Yidiny word-final deletion does not proceed as one might expect from the discussion in §4. A solution is then offered in §6.

Prince and Tesar's BCD is a computationally efficient algorithm for the learning of OT grammars. It builds upon Tesar's earlier Recursive Constraint Demotion (RCD) algorithm (**tesar1995 tesar2000**), deterministically learning a grammar, conditional on the data, by ranking constraints in a series of steps, or recursions. At the first step, one or more constraints is assigned to the highest-ranked constraint stratum in the grammar. A stratum is a set of constraints whose relative ranking against one another is indeterminate given the data, but whose ranking relative to constraints in other strata is significant. The act of assigning constraints to a stratum is termed installation. At each subsequent step, one or more additional constraints are installed in the next-highest stratum, and so on, until all constraints are ranked. The determination of which constraint(s) are installed next is based on evidence from winner–loser pairs (WLPs). For each WLP, any constraint yet to be installed will favor the winner in the pair, the loser, or neither. The full table of WLPs and constraints yet to be installed is termed the support. A fragment of a support is shown in (29). The relative order of constraints and WLPs in a support is inconsequential, though for ease of inspection I set out markedness constraints to the left of a vertical double line, and faithfulness to the right.

(29)

	FtBin	Prs-*u*	Prs	*Cplx	Max	Cntg	Anc
a. /margu-ni/ (mar gu:n) ≻ (mar gu:) ni		W	W		L		
b. /guygal-ni/ (guy ga:l) ni ≻ (guy ga:l)		L	L		W		W
c. /guygal-ni/ (guy ga:l) ni ≻ (guy ga:ln)		L	L	W	W		
d. /guygal-ni/ (guy ga:l) ni ≻ (guy ga:l ni)	W	L	L				
e. /bulmba/ (bulm ba) ≻ (bul ba)			L	L	W	W	

In the original RCD algorithm, the sole criterion for installing a constraint was that it favor no losers. This is true of the constraints FtBin, Cntg and Anc in (29). When a constraint, C, is installed, all of the WLPs for which C favors the winner are removed from the support, since the constraint ranking has now accounted for them. In the RCD, all constraints meeting this criterion at any recursion are installed, and the result at the end of all recursions is a correct grammar for the data. Nevertheless, the grammars inferred by the RCD are not optimal (**princetesar2004**). The suboptimality relates to the subset problem (**baker1979r**; **angluin1980**), a general problem in algorithmic learning from positive evidence, namely that the system which results from learning will correctly assess as grammatical all attested items, but will fail to rule out certain systematically unattested items. This in turn relates to the notion of restrictiveness: a learning algorithm ought ideally to learn the most restrictive grammar consistent with the data. The RCD does not do this. In practice, meeting this desideratum is challenging for an efficient algorithm. However **princetesar2004** demonstrate that good headway can be made by enhancing the RCD with a small set of biases, hence the name Biased Constraint Demotion, or BCD. The BCD differs from the RCD in two main respects. The first is the principle of faithfulness delay. According to this, at every recursion faithfulness constraints are not installed, even when they favor no losers, unless there are no other installable constraints. In (29) for example, the BCD would install the markedness constraint FtBin but not the faithfulness constraints Cntg and Anc. If we do this, and remove from (29) all the WLPs for which FtBin favors the winner, namely (29d), and remove FtBin, we have (30), in which only faithfulness constraints, Cntg and Anc favor no losers; under these conditions, faithfulness delay would permit the installation of Cntg and Anc.

(30)

	Prs-*u*	Prs	*Cplx	Max	Cntg	Anc
a. /margu-ni̲/ (mar guːn) ≻ (mar guː) ni	W	W		L		
b. /guygal-ni̲/ (guy gaːl) ni ≻ (guy gaːl)	L	L		W		W
c. /guygal-ni̲/ (guy gaːl) ni ≻ (guy gaːln)	L	L	W	W		
e. /bulmba/ (bulm ba) ≻ (bul ba)		L	L	W	W	

However, there is a second principle to consider also. A principle of 'freeing up markedness' states that when there is a choice between installing several faithfulness constraints, the algorithm should install the smallest subset possible, whose installment would cause a markedness constraint to become installable in the next recursion. For example, in (30), installing Cntg would remove WLP (30e), thereby freeing up the markedness constraint *Cplx at the next recursion; no comparable gain would flow from installing Anc. On those grounds, from (30) the BCD would install Cntg.

5.2 A support for learning Yidiny exceptionality

I now consider several learning scenarios for Yidiny exceptionality. Each begins directly after the installation of undominated constraints. **??** contains a set of WLPs that is representive of all combinations of roots and suffixes which are relevant to the grammar of word-final deletion: it is not the complete support, but it represents the complete support well. Segments which can delete are underlined. To economize on space below, WLPs will be referred to by the letters in the first column of **??**.

5.3 Learning the phonological account (preliminary version)

We begin with the learning of the phonological account of Yidiny exceptionality described previously in §4.2. For the moment, I assume that input segments are already lexically ☒-indexed as *u* or *n*. We begin after undominated constraints have been installed, with a support as in (31).

Table 6: Support for learning Yidiny exceptionality.

a.	/margu-ni̱/	(mar guːn)	≻(mar guː) ni
b.	/guygal-ni̱/	(guy gaːl) ni	≻(guy gaːl)
c.	/guygal-ni̱/	(guy gaːl) ni	≻(guy gaːln)
d.	/margu-ŋgu̱/	(mar guːŋ)	≻(mar guːŋ) gu
e.	/margu-ŋgu̱/	(mar guːŋ)	≻(mar guːŋg)
f.	/bigunu̱-yi̱-ŋgu̱/	(bi gu) (nu yiːŋ)	≻(bi gun) (yiŋ gu)
g.	/wawa-lɲu̱/	(wa waːl)	≻(wa waːl) ɲu
h.	/gali-ŋa̱/	(ga liːŋ)	≻(ga liː) ŋa
i.	/gaɟara̱/	(ga ɟaːr)	≻(ga ɟaː) ra
k.	/margu-nda/	(mar guːn) da	≻(mar guːn)
l.	/wawa-lna/	(wa waːl) na	≻(wa waːl)
m.	/gali-na/	(ga liː) na	≻(ga liːn)
n.	/guɟara/	(gu ɟaː) ra	≻(gu ɟaːr)
o.	/maɟinda-ŋa̱-lna/	(ma ɟin) (da ŋaːl) na	≻(ma ɟin) (da ŋaːl)
p.	/bulmba/	(bulm ba)	≻(bul ba)

(31)

	Prs-*u*	Prs	*Cplx	Max	Max-C	Cntg	Anc
a, h, i.	W	W		L			
b.	L	L		W	W		W
c.	L	L	W	W			
d, g.	W			L	L		
e.			W	L	L		
f.				L	L	W	
j, k, l, o.		L		W	W		
m, n.		L		W			
p.			L	W		W	

Support (31) does not contain any markedness constraints that favor no losers. Two faithfulness constraints favor no losers: Cntg, which would free up *Cplx if installed, and Anc, which would not free up any markedness constraints. Consequently, Cntg is installed next, removing WLPs (f) and (p) from the support. After that, the newly freed-up *Cplx is installed, removing WLPs (c) and (e), and leaving (32).

(32)

	Prs-*u*	Prs	Max	Max-C	Anc
a, h, i.	W	W	L		
b.	L	L	W	W	W
d, g.	W	W	L	L	
j, k, l, o.		L	W	W	
m, n.		L	W		

In 32 only Anc favors no losers, and so is installed. This removes (b), freeing up Prs-*u*, which is installed next, removing (a,h,i) and (d.,g), leaving (33). From (33), Max will be installed since it frees up Prs. This leaves Prs and Max-C, which according to faithfulness delay, will be ranked last as Prs ≫ Max-C, as in (34).

(33)

	Prs	Max	Max-C
j, k, l, o.	L	W	W
m, n.	L	W	

(34) Cntg ≫ *Cplx ≫ Anc ≫ Prs-*u* ≫ Max ≫ Prs ≫ Max-C

Some comments are in order. First, the BCD algorithm has learned the key constraint ranking Prs-*u* ≫ Max ≫ Prs responsible for the core of Yidiny exceptionality. Secondly however, it has also ranked Anc ≫ Prs-*u*, in which case the learned grammar expressly prohibits morph-initial deletion. Indeed, had Max-C/rt been included in (31), it would also have been ranked highly since it only ever favors winners, meaning the grammar would also expressly prohibit CV deletion in roots (the reasons for my excluding Max-C/rt are clarified in §6). This means that the algorithm is learning precisely the rankings required to prevent the phonological solution from overgenerating, thereby voiding the major criticism of the phonological approach which was introduced in §4.2. This is perhaps surprising, so why is the ranking learned? It is learned because the BCD algorithm attempts to construct a restrictive grammar. The typical assumption, that grammars implementing a phonological approach would not assign redundant, high rankings to constraints like Anc, is predicated on an implicit assumption that the learner would be seeking a permissive grammar; doing so leads to overgeneration. However no successful learner would adopt that assumption, because successful learning in general requires a restrictive approach. For the theory of exceptionality, this is significant. It means the result obtained here, in which a phonological approach to exceptionality has been learned without overgeneration, is not dependent on some minor detail of the BCD, or the constraints used, or even OT. Rather, it follows from a general principle of learning. Consequently, the adoption of realistic assumptions about learning narrows the performance gap between the phonological and morphological approaches. I will examine the phonological approach further in §7.3.

5.4 Learning indexed constraints and the morphological analysis

We consider next the learning of the morphological approach. The support begins, after installation of undominated constraints, as (35). These are the same constraints and WLPs as in the previous section, but without Prs-*u*. The support begins with no lexically indexed constraints; how they are learned is considered shortly. I also do not include Max-C/rt in the support. Max-C/rt is essentially a variant of Max-C, indexed to all root morphs. This is the kind of constraint we might reasonably expect the morphological approach to learn.

(35)

	Prs	*Cplx	Max	Max-C	Cntg	Anc
a, h, i.	W		L			
b.	L		W	W		W
c.	L	W	W			
d, g.	W		L	L		
e.		W	L	L		
f.			L	L	W	
j, k, l, o.	L		W	W		
m, n.	L		W			
p.		L	W	W	W	

Turning now to the BCD algorithm, neither of the markedness constraints in support (35) favors no losers. Cntg does, and would free up *Cplx. Anc also does, but would not free up any markedness constraints. Accordingly, Cntg is installed next, removing WLPs (f) and (p) are from the support, and *Cplx after that, removing (c) and (e), leaving (36).

(36)

	Prs	Max	Max-C	Ans
a, h, i.	W	L		
b.	L	W	W	W
d, g.	W	L	L	
j, k, l, o.	L	W	W	
m, n.	L	W		

Anc is installed next, removing WLP (b), which leaves (37), a support in which there is no constraint which favors no losers.

(37)

	Prs	Max	Max-C
a, h, i.	W	L	
d, g.	W	L	L
j, k, l, o.	L	W	W
m, n.	L	W	

Supports in this state are said to have reached INCONSISTENCY. An inconsistency, however, is not a failure.

Inconsistencies indicate that the combination of data and assumptions currently under consideration have not led to a working grammar. Accordingly (assuming the data is correct), a revision of the assumptions is warranted. Suppose, in this case, that a revision could be made which leaves intact all previously installed constraints and their rankings, and the validity of all previously accounted-for WLPs, that is, a revision that would change only what is in the support. Suppose also that as a result of this revision the support came to contain a constraint that favors no losers. Such a revision would resolve the inconsistency. The BCD could restart and, one hopes, lead to a working grammar. Revisions that meet these criteria can be considered a type of learning. One such revision is to add a new, lexically M-indexed constraint to Con.

pater2009r describes a method for learning M-indexed constraints and assigning coindices to morphs, which takes a BCD inconsistency as its starting point. **coetzee2009** extends this to Output-Output constraints, which I will not consider here. Becker's modifications (**becker2009**; **becker2011r**) are addressed in §8.

Central to Pater's method is the operation of CONSTRAINT CLONING, a process I describe informally here and return to in detail in §8. Within the stalled support, a constraint C is sought which, if it were indexed to some set *M* of morphs, would (i) favor at least one winner[15] and (ii) favor no losers. Assuming such a constraint C can be identified, it is then cloned, which is to say, a lexically M-indexed version of it, C_M, is added to the support. Because C_M favors no losers, it is installed next. For example, support (38) is the same as (37) but now displays information about which morphs are involved. I have annotated relevant undergoers as *U* and non-undergoers as *N*.

[15] The new constraint needs to favor at least one winner to have any chance of freeing up another constraint once it is installed.

(38)

		Prs	Max	Max-C
a.	/margu-ni$_U$/	W	L	
d.	/margu-ni$_U$/	W	L	L
g.	/margu-ŋgu$_U$/	W	L	L
h.	/gali-ŋa$_U$/	W	L	
i.	/gaɟara$_U$/	W	L	
j.	/binarŋa/	L	W	W
k.	/margu-nda$_N$/	L	W	W
l.	/wawa-lna$_N$/	L	W	W
m.	/gali-na/	L	W	
n.	/guɟara$_N$/	L	W	
o.	/maɟinda-ŋa-lna$_N$/	L	W	W

According to the criteria for cloning, all three of Prs, Max and MaxC are candidates for cloning (indexed to sets *U*, *N* and *N* respectively). I assume that owing to faithfulness delay, markedness constraints are cloned in preference to faithfulness when both are available, in which case Prs will be cloned. In (39) the cloned, lexically M-indexed constraint Prs$_U$ is added to the support. Installing it removes WLPs (a,d,g,h,i) which frees up Max, whose installation is followed by Prs and Max-C. The resulting ranking is (40), which requires comment.

(39)

	Prs-$_U$	Prs	Max	Max-C
a, h, i.	W	W	L	
d, g.	W	W	L	L
j, k, l, o.		L	W	W
m, n.		L	W	

(40) Cntg ≫ *Cplx ≫ Anc ≫ Prs$_U$ ≫ Max ≫ Prs ≫ Max-C

The algorithm has successfully learned the key constraint ranking Prs$_U$ ≫ Max ≫ Prs. However, it did not create an indexed version of Max-C for roots, and thus has not learned to expressly prohibit CV deletion in roots. To be sure, no individual roots ending in CCV or VC′V (where C′ would be an illicit coda) will have been co-indexed to Prs$_U$ during the cloning operation (see §8 for details) and so none of those roots will be subject to CV deletion, however the ranking in (40) predicts that if the lexicon did contain a root such as **binarŋa$_U$*, then that root and any like it would undergo CV deletion. This is overgeneration of the same kind which was believed to beset phonological accounts. Thus, while §5.3 showed that grammars learned for the phonological account may suffer less than expected from overgeneration once learning is taken into consideration, §5.4 shows that grammars for the morphological account may suffer from overgeneration more than expected. In the next section, I propose a solution.

6 Morphological analytic bias: the Morphological Coherence Principle

In §5.4 the grammar which was learned for a morphological analysis of Yidiny exceptionality suffers from a manifestation of the subset problem. Although the algorithm correctly handled all attested data, it did not learn the more restrictive generalization which applies also to unattested data, that roots in Yidiny do not undergo consonant deletion. The problem arises because the cloning procedure assesses morphology on a morph-by-morph basis only, whereas the true generalization in Yidiny applies to a class of morphs, in this instance, to roots. The remedy to be pursued here has two parts. It adds a new kind of constraint cloning, which indexes a constraint not to an idiosyncratic lexical list of morphs, but to a general class. It then biases constraint cloning so that class-indexed (or K-INDEXED) cloning is preferred over lexically indexed cloning. Effectively, this introduces an analytic bias (**moreton2008r**) from morphology to phonological learning at BCD inconsistencies.

Now, supposing that the algorithm is seeking a constraint that it will clone and K-index to some non-idiosyncratic class of morphs, which classes should be available for the learner to consider? Important here is the fact that human phonological learning will need to proceed in parallel with, and interleaved with, morphological learning (**tesar2007 merchant2008**). Accordingly, I assume the learner has access both to universally-defined classes such as 'root', and those classes which have been morphologically learned, such as ERGATIVE CASE. The biasing principle, which I term the Morphological Coherence Principle is stated in (41), where criterion 2 provides an additional bias towards maximal restrictiveness.

(41) The Morphological Coherence Principle:
   1. At a BCD inconsistency, attempt to create a K-indexed constraint, co-indexed to some universal or learned morphological class *K*, before attempting to create a lexically-indexed constraint.
   2. If multiple constraints are eligible for K-indexation, select the one whose co-indexed class is most general.

The MCP has some desirable theoretical properties. If the universal state of CON at the commencement of learning is CON_{init}, then the MCP obviates the need for CON_{init} to contain any constraints that are relativized to universal or learned morphological classes, since such constraints will be learned on demand, if and only if needed. In effect, this reduces the size of CON_{init} without any change in the explanatory capacity of the theory. And, since it allows the grammar to BUILD constraints for language-specific morphological classes it makes those constraints available to the learner without problematically assuming them universal (**russell1995 hammond2000** see also **smith2004 flack2007r**). The MCP operationalizes, in a specific manner, the kind of insight into linguistic theory that **anderson2008** argues ought to follow from an improved understanding of the learning device.

Let us now return to Yidiny exceptionality, equipped with the MCP. Learning begins and proceeds as in §5.4 until the inconsistency in (38), at which point a constraint is sought for cloning. The MCP states that if possible, a constraint should be cloned and K-indexed. In (38) MAX-C would favor no losers if it were K-indexed to the entire class of roots, so it is cloned and accordingly K-indexed. This is the functional equivalent of adding MAX-C/RT to CON, and the reason why in §5 I did not include MAX-C/RT in the support at the outset. Adding MAX-C/RT to the support results in (42). From (42), MAX-C/RT is installed and WLP (j) is removed, whereupon we return to inconsistency, in (51). As in §5.4, the process from that point results in the cloning of PRS and the installation of PRS_{U}, then MAX, PRS and MAX-C, yielding the desired constraint ranking (43).

(42)

	PRS	MAX	MAX-C/RT	MAX-C
a, h, i.	W	L		
d, g.	W	L		L
j.	L	W	W	W
k, l, o.	L	W		W
m, n.	L	W		

(43)

	PRS	MAX	MAX-C
a, h, i.	W	L	
d, g.	W	L	L
k, l, o.	L	W	W
m, n.	L	W	

(44) CNTG ≫ *CPLX ≫ ANC ≫ MAX-C/RT ≫ PRS_{U} ≫ MAX ≫ PRS ≫ MAX-C

To summarize, results from §5.3 suggested that, provided a learner is seeking a restrictive grammar, the phonological approach to exceptionality may not suffer from overgeneration. This contradicts recent arguments, which on examination appear to adopt the implausible assumption that a learner would be seeking a permissive grammar. That being said, I have not yet clarified how the learner would arrive at the requisite ☒-indices required by the phonological approach. That will be discussed in §7.3. Meanwhile, §5.4 revealed that without further refinement, the BCD is prone to learning grammars that overgenerate even in a morphological approach to exceptionality, due to an overly atomistic method of morphological generalization. This was remedied in §6 by the Morphological Coherence Principle (41), which solves the learning problem and simplifies CON_{init}.

7 The theoretical status of lexical indices

7.1 Lexical M-indices

In §7 I set Yidiny to one side and consider some matters of theory.

Lexical M-indices are representations which are visible to the phonology, but they are not phonological elements *per se*. In OT, GEN cannot alter M-indices. It cannot add or remove them, or displace them from one morph to another. There is therefore no need for mechanisms such as M-index 'faithfulness', rather it is simply assumed that the lexical affiliation of a morph *m* with an M-index *M* is identical in the input and output. This set of properties is shared with other kinds of lexical affiliation, such as the affiliation of a phonological element with its morph, and is termed Consistency of Exponence (**mccarthy1993a vanderived2007**).

Taking a historical view, M-indices closely resemble the RULE FEATURES and ALPHABET FEATURES of early generative phonology (GP) (**chomskyhalle1968 lakoff1970 coats1970 zonneveld1978** *inter alia*). Both sets of formalisms fulfill the function of determining for cases of exceptionality whether a morph *m* participates in certain phonological patterns or not, by ensuring that *m* is visible or not visible as required, to OT's constraints or GP's phonological rules. Diacritic features were investigated extensively in GP. It was argued that the theory should not allow the phonology to manipulate diacritic features (**kiparsky1973r zonneveld1978**). The same applies to M-indices in OT. It was argued that not all idiosyncrasies in the phonology can be analysed satisfactorily in terms of rule exception features, and that there is an additional role for cyclicity (**chomskyhalle1968 kiparsky1982r**) and the same has been recognized for M-indices (**pater2009r**). In GP, it was also assumed that the diacritic features of morph *m* were distributed across, and directly characterized, each of the phonological elements (namely, segments) in *m*. We might ask whether this is also true of M-indices in OT. Suppose that it is, so that the *M*-indices of a morph *m* directly characterize each phonological element φ that is lexically affiliated with *m* (that is all φ which are EXPONENTS of *m*). In that case, the relative definition of an M-indexed constraint (25), repeated here as (45), can be revised and simplified as (46).

(45) $V(C_M) =_{def} V(C) \cap L_{\varphi}(m \in M \,\&\, Exp(\varphi, m))$, *where*:
M is the set of morphs co-indexed to C_M.
$Exp(\varphi, m)$ states that element φ is an exponent of morph *m*

(46) $V(C_M) =_{def} V(C) \cap L_{\varphi}(\varphi \in \boxtimes_M)$, *where*:
$\boxtimes_M$ is the set of phonological elements co-indexed to C_M.

It will be recalled that the relative definition of a constraint C_M is expressed as the set intersection between the loci of variation of the unindexed constraint C, written V(C), and the set of loci, $L_{\varphi}(D(\varphi))$ which contain some criterial type of phonological element φ, described by predicate $D(\varphi)$. Importantly, this means that M-indexed constraints are defined DIRECTLY in terms of phonological elements, φ, and only indirectly in terms of morphs *m*. The indirectness shows up in the complexity of $D(\varphi)$ in (45), which links morphs to their exponent φ elements via the function $Exp(\varphi, m)$. This is in contrast with (46), where the assumption is that all φ elements are directly characterized by the M-index borne by their affiliated morph. The constraint definition no longer refers to the morph itself, and so the predicate $D(\varphi)$ is simpler.

At risk of laboring the point, the phonology itself assesses violations of M-indexed constraints directly in terms of φ elements, not morphs. While it is possible to refer to the morphs in the definitions of M-indexed constraints as in (17)/(45), it is not necessary. Nor is it possible to refer only to the morphs and not to the φ elements, since the loci of violation of these constraints are defined inherently at a sub-morphological, phonological level.

7.2 Lexical ☒-indices

Let us now consider the nature of lexical ☒-indices of the type I invoked in §4.2 and §5.3. My proposal is that these are exactly like M-indices: they are non-phonological indices of lexical affiliation, visible to, but not manipulable by, the phonology and used for making particular phonological elements visible or not, as required, to OT's constraints in order to provide a coherent account of exceptionality. The only distinction between ☒-indices and M-indices lies in the supplementary assumption attached to M-indices, in (47).

(47) The M-index assumption:
A lexical index which characterizes phonological element φ_i will also characterize all other phonological elements φ_j affiliated with the same morph *m*.

☒-indices are not subject to this redundancy; they are affiliated with only those φ elements for which the affiliation makes any difference to the analysis of language. As I will show in §8, that makes ☒-indices somewhat simpler to learn, since they correspond more directly to the evidence in the data.

The reader may also have noticed that the definition of a ☒-indexed constraint in (21) is almost exactly like the simplified definition of an M-indexed constraint in (46). This reflects the fact that for the operation of the phonology, it is φ elements, and the indexation of specific φ elements, that matter. Whether or not one chooses to adopt supplementary assumption (47) in fact has no material consequence for the evaluation of an individual indexed constraint. The question of whether there are other consequences, and whether they are desirable, is taken up in §9.

7.3 Learning lexical ☒-indices

Given the proposal above, the learning of ☒-indices is quite parallel to the learning of M-indices. I assume that the MCP still applies, so that class-based exceptionality and K-indexed constraints continue to be learned with priority over idiosyncratic exceptionality, even though the latter will now be accounted for by ☒-indexed constraints, not M-indexed. This is a coherent assumption to make. The MCP is concerned with the learning of class-based generalizations, whereas ☒- and M-indexed constraints are alternative devices for learning idiosyncrasies. Accordingly, in a stalled support once there are no K-indexed constraints available for cloning, the algorithm seeks a constraint C which, were it indexed to some set ☒ of phonological elements, would (i) favor at least one winner and (ii) favor no losers. All else proceeds as for M-indexed constraints. In

the learning of Yidiny word-final deletion, the process begins as in §6, leading to a first inconsistency resolved by the addition of MAX-C/RT to CON, and proceeding from there to the second inconsistency (43), repeated here in part and in more detail as (48).

(48)

		PRS	MAX	MAX-C
a.	/margu-ni̲/ (mar guːn) ≻ (mar guː) ni	W	L	
d.	/margu-ŋgu̲/ (mar guːŋ) ≻ (mar guːŋ) gu	W	L	L
h.	/gali-ŋa̲/ (ga liːŋ) ≻ (ga liː) ŋa	W	L	
i.	/gaɟara̲/ (ga ɟaːr) ≻ (ga ɟaː) ra	W	L	
k.	/margu-nda/ (mar guːn) da ≻ (mar guːn)	L	W	W
m.	/gali-na/ (ga liː) na ≻ (ga liːn)	L	W	
n.	/guɟara/ (gu ɟaː) ra ≻ (gu ɟaːr)	L	W	
o.	/maɟinda-ŋa̲-lna/ (ma ɟin) (da ŋaːl) na ≻ (ma ɟin) (da ŋaːl)	L	W	W

In (48), no K-indexed constraint is available for cloning.[16] Turning to potential ☒-indexed constraints, we see that the constraint PRS would, if it were co-indexed to all underlined phonological elements, favor at least one winner and favor no losers, and so it is cloned and co-indexed resulting in (49). From there the algorithm proceeds in the now-familiar fashion, resulting in grammar (50). With its high-ranking MAX-C/RT and ANC, (50) does not overgenerate. Moreover, given the argument in §7.2, that for EVAL there is no detectable difference between M-indexed and ☒-indexed constraints, we can see that grammar (50) is in all material aspects identical to grammar (44) learned in §6.

[16] Actually this is not strictly true. All PAST suffixes for example are undergoers, in which case the MCP would generate and rank PRS$_{PST}$. Notwithstanding this, the essential argument remains, since other morphological classes exist, such as ERGATIVE and 'root', that are not uniformly (non)undergoers, and still need to be handled by lexically-indexed, not K-indexed, constraints. This minor correction applies equally to the learning process in §6.

(49)

		Prs	Prs_U	Max	Max-C
a.	/margu-ni̲/	W	W	L	
d.	/margu-ŋgu/	W	W	L	L
h.	/gali-ŋa̲/	W	W	L	
i.	/gaɟara̲/	W	W	L	
k.	/margu-nda/	L		W	W
m.	/gali-na/	L		W	
n.	/guɟara/	L		W	
o.	/maɟinda-ŋa̲-lna/	L		W	W

(50) Cntg ≫ *Cplx ≫ Anc ≫ Max-C/rt ≫ Prs-*u* ≫ Max ≫ Prs

8 Constraint cloning

8.1 Assessing eligibility for cloning

It is necessary now to examine more precisely the processes by which constraints are deemed eligible for cloning (§8.1), by which a viable set of co-indexed elements is identified (§8.2), and by which a selection is made between multiple eligible constraints (§8.3).

Earlier, I introduced criteria by virtue of which a constraint becomes eligible for cloning. These are restated in (51) in a generalized from, so that the set *S* is: a coherent class of morphs for K-indexing; an idiosyncratic set of morphs for M-indexing; or an idiosyncratic set of lexical phonological elements for ☒-indexing.

(51) A constraint should be sought for cloning which, if it were indexed to set *S*, would (i) favor at least one winner, and (ii) favor no losers.

Criterion (51ii) ensures that once the cloned constraint is added to the support, it can be installed; (51i) ensures that its installation will remove at least one WLP from the support, and thereby have some hope of freeing up other constraints. The formulation in (51) improves upon Pater's (**pater2009r**) criterion, which is to seek a constraint that favors no losers 'for all instances' of some morph.[17] To see why Pater's criterion fails, consider WLPs (h,l,o) from the stalled support (38), reproduced in part and in detail in (52). For the purposes of discussion, I assume we are attempting to learn an M-indexed constraint, though the argument generalizes to other kinds.

[17] Pater's phrase 'favors only winners' is equivalent to my 'favors no losers'.

(52)

		Prs	Max	Max-C
h.	/gali-ŋa/ 'go-com[imp]' (ga li:ŋ) ≻ (ga li:) ŋa	W		
l.	/wawa-lna/ 'see-pure' (wa wa:l) na ≻ (wa wa:l)	L	W	W
o.	/maɟinda-ŋa-lna/ 'walk up-com-purp' (ma ɟin) (da ŋa:l) na ≻ (ma ɟin) (da ŋa:l)	L	W	W

In (52), WLPs (h) and (o) both contain the suffix *-ŋa*, a regular undergoer which our procedure ought to co-index to the M-indexed constraint Prs_{U}. In WLP (h) word-final *ŋa* is subject to deletion, and Prs favors the winner. In WLP (o) non-final *ŋa* is parsed into a foot and escapes deletion. Nevertheless, for WLP (o) Prs favors the loser. This has nothing to do with *ŋa*, but is due to the non-deletion of the unparsed, word-final non-undergoer *-lna*. Pater's co-indexing criterion asks whether Prs favors no losers 'for all instances' of *-ŋa* in the support. The answer is 'no', because (o) contains an instance of *-ŋa* and Prs favors the loser for (o). This is the wrong result; the suffix *-ŋa* ought to get co-indexed to Prs_{U}. It comes about because Pater's criterion does not discriminate between morphs that contribute to violations and those which are present in the word, but do not contribute. The criteria in (51) avoid this problem because they refer directly to how the co-indexed constraint would perform, were it created. The next two sections detail how to operationalize them.

8.2 Specifying co-indexed sets

The question considered here is which set *S* ought to be co-indexed to a given constraint C if we wish to clone C? The answer varies depending on which kind of indexed constraint we are constructing. One possible answer is that no such set exists, and C cannot be cloned. Seen from that angle, the question here is also: is C eligible for cloning?

K-indexed constraints can be co-indexed only to the morphological classes $K_1, K_2 \ldots K_n$ in the language (§6). In (41) I suggested that the preferred class for co-indexation is the most general one. Thus, to efficiently assess if constraint C is eligible for cloning and K-indexing, the learner should proceed stepwise through the available classes, ordered by decreasing generality. The process is one of trial and error. At each step, the constraint C_K is built and applied to all WLPs in the support. If C_K meets criteria (51) then it is successful; the process halts and C_K is used, otherwise the trial and error continues. If by the end, no successful constraint $C_{K1} \ldots C_{Kn}$ is found, then C is ineligible for cloning.

For M-indexed and ☒-indexed constraints, the desired set *S* can be identified by focusing attention on loci of violation. Suppose we are considering constraint C for cloning. For any WLP, *p*, its loci of violation of constraint C fall into three classes: the class w(*p*), responsible for violations of C that favor the winner (i.e., the locus occurs in the loser only), class l(*p*) which favor the loser (locus occurs in the winner only) and class n(*p*) which favor neither (occurs in both). Next define $☒_{\text{w}(p)}$ as the set of phonological ele-

ments φ contained in any of the loci in $\textsc{w}(p)$, and $\boxtimes_{\textsc{l}(p)}$ as the set of φ elements contained in any of the loci in $\textsc{l}(p)$. Finally, define $\boxtimes_{\textsc{w}}$ as the union of $\boxtimes_{\textsc{w}(p)}$ for all WLPs, p_1, p_2 ... p_n, in the support, and $\boxtimes_{\textsc{l}}$ as the union of all $\boxtimes_{\textsc{l}(p)}$ in the support. Now, consider the set $(\boxtimes_{\textsc{w}} - \boxtimes_{\textsc{l}})$, the set difference between $\boxtimes_{\textsc{w}}$ and $\boxtimes_{\textsc{l}}$. This is the set of all φ elements which both (i) appear in at least one locus that in at least one WLP causes C to favor a winner, and (ii) never appear in a locus that causes C to favor a loser. For a $\boxtimes$-indexed constraint this is an optimal set S. If for a given constraint C, $(\boxtimes_{\textsc{w}} - \boxtimes_{\textsc{l}})$ is the null set, then we may conclude that C is ineligible for cloning.[18]

To find the equivalent for an M-indexed constraint, it is necessary to extrapolate from $\boxtimes_{\textsc{w}}$ and $\boxtimes_{\textsc{l}}$ to morphs: set S will be the set $(M_{\textsc{w}} - M_{\textsc{l}})$ where $M_{\textsc{w}}$ is the set of all morphs m_w, such that any of m_w's phonological exponents is an element of $\boxtimes_{\textsc{w}}$; and $M_{\textsc{l}}$ is the set of all morphs m_l, such that any of m_l's phonological exponents is an element of $\boxtimes_{\textsc{l}}$. Note that $M_{\textsc{w}}$ and $M_{\textsc{l}}$ can be calculated only after the calculation of $\boxtimes_{\textsc{w}}$ and $\boxtimes_{\textsc{l}}$ is performed.

In §7.1 I considered what is involved computationally in assessing violations of $\boxtimes$- and M-indexed constraints, and argued that the calculations for both are essentially concerned with φ elements, not morphs. Here we see that the same is true when learning the co-indexed set. As in §7.1, one can bring morphs into the picture, to be sure, but in both cases doing so requires additional computational effort, for no effective difference in how the grammar will work. In §9 I will argue the theory to be preferred is one which admits lexically $\boxtimes$-indexed constraints, but not M-indexed.

8.3 Selecting among eligible constraints

Suppose there are multiple lexically-indexed constraints which are eligible for cloning; which do we choose? The principles of faithfulness delay and freeing-up of markedness constraints will eliminate some options (§5.1). Beyond that, I suggest the learner chooses the constraint which favors the most winners, and whose installment would therefore remove the greatest number of WLPs. A desirable consequence will be a bias toward restrictiveness. For example, suppose Max is eligible. If so, then so too is Max-C, Max-V, Max-p, etc. This 'maximize-winners' criterion would select Max, and increase the restrictiveness of the grammar, relative to the other options.

Interestingly, **becker2009** proposes a MINIMIZE-winners criterion, whose effect is to generate many, very specific cloned constraints, each indexed to highly specific subclasses in the lexicon. The aim is to account for a particular phenomenon, which I describe here. I argue that other accounts are possible, and that Becker's solution has undesirable consequences.

When language learners assign novel words to existing grammatical categories, they do so on the basis of statistical correlations that exist in the lexicon, for example between category membership and aspects of the members' phonological forms (**poplack1982**;

[18] To be precise, if $(\boxtimes_{\textsc{w}} - \boxtimes_{\textsc{l}})$ is the null set then it is possible that there still exists some additional, viable set S which contains fortuitous elements φ_i which are elements of both $\boxtimes_{\textsc{w}}$ and $\boxtimes_{\textsc{l}}$ such that in EVERY WLP p in which φ_i is contained in some number n of the loci $\textsc{w}(p)$ there are at least n offsetting loci in $\textsc{l}(p)$ which contain other elements φ_j which are also in S. Identifying these fortuitous elements φ_i, or even determining if any exist, would very likely be prohibitively expensive computationally.

albright2002b). One such task is to assign a word as exceptional or non-exceptional, given evidence which underdetermines that choice. The key question here is, what existing statistical knowledge do speakers use, and what do they ignore? In Turkish, speakers appear to ignore correlations between the (non)alternation of a stop's laryngeal features and the quality of its neighboring vowel. It is proposed (**becker2009**; **becker2011r**) that this is because speakers do not access lexical statistics *per se*, rather they attend to the statistics of constraint indexation. Importantly, Con lacks constraints such as *[+high]tV which refer to a stop and the quality of its vocalic neighbor. Consequently, no such constraint can be indexed, making such correlations invisible and hence irrelevant to a speaker when she assigns a novel word to a (non)exceptional lexical category. Assuming this is the case, then in order for fine-grained knowledge to be available to speakers, an atomizing, 'minimize-winners' criterion for cloning is needed. However, this solution would seem neither necessary nor warranted.

Notwithstanding the facts of Turkish, speakers in other languages and performing other novel-word tasks do use lexical correlations which lack a corresponding constraint in Con (**moreton1999 albright2002b albrighthayes2002 ernestus2003**), indicating that speakers are capable of such computation. In that case, atomized indexed constraints alone are not enough to produce the Turkish results. An additional stipulation is required, that this ability is suppressed when assigning novel words to exceptionality classes; yet this leads to a curious view of phonology. Whereas the grammar is usually the store of generalizations, just in the case of exceptionality, it is a store of highly detailed idiosyncrasy, and just in that case speakers ignore their usual, lexical store of idiosyncrasy and turn to the grammar. More satisfying would be to find some other explanation of the Turkish data. While that would take us well beyond this scope of this paper, it can be noted that what is required is a mechanism that can filter the lexical information in some way. That mechanism needn't be part of the OT grammar. Indeed, if it is true that learners build certain constraints during learning (**flack2007r hayes2008 hayes2014**), then there must exist extra-grammatical generalization devices, which may provide the lexicon-filtering power needed. For now I conclude that that Becker's proposal follows from just one possible solution to an interesting puzzle, however both the puzzle and solution are outliers relative to what else we know. In contrast, a 'maximize-winners' criterion leads to the learning of restrictive grammars, and on those general grounds would appear correct.

9 Discussion

9.1 The case against concrete accounts

Throughout this paper, I have considered only the abstract phonological approach to analyzing exceptionality, gradually building the argument that its superiority to the morphological approach lies in the fact that it localizes exceptionality to specific φ elements, which are the elements in terms of which the relevant computation must be carried out. Concrete phonological approaches also localize exceptionality at a sub-morphological

level, but compared to the abstract approach they are ill-suited to learning, and to seriality, as follows.

Lexical indexation is an ideal response to BCD inconsistency, because it annotates the lexicon with indices which are invisible to all previously installed constraints. This guarantees, without needing to check, that all previously accounted-for WLPs remain accounted for. Even if some of them contain lexical φ elements which acquire a new index, their violations of all previously ranked constraints remain unchanged, since no previously-ranked constraint is sensitive to the new index. In contrast, the alteration of phonological form — for example, removing a root node from certain segments — may very well alter the evaluation of WLPs by already-ranked constraints, thus it requires a re-evaluation of the entire ranking. It is not possible to simply repair an inconsistency and resume the BCD process. An ABSTRACT phonological account is therefore easier to learn.

In serial theories, concrete phonological approaches face the problem that in non-initial strata, it is possible that a preceding stratum will have removed, altered, moved or introduced, those aspects of phonological form which should function as pseudo-indices, which lack Consistency of Exponence. This opens up the possibility of all manner of phonological manipulations of exceptionality, for which I am unaware of any evidence.

Taking a more historical view, **chomsky1964** criticized concrete phonological accounts espoused by structuralists (e.g. **bloomfield1939**) for the proliferation of underlying segments that they entailed. To the extent that such concerns matter to modern phonological theories, ⊠-indexation avoids such proliferation by augmenting representations with non-phonological indices (cf §7), rather than additional underlying phonological distinctions.

9.2 The case against M-indexing

In §7 and §8 I showed that for both constraint evaluation and constraint learning, exceptionality is calculated in terms of phonological elements, not morphs. Morphs can be brought into the picture, but at additional computational cost and to no effect. Perhaps, however, it is nevertheless empirically true that exceptionality is inherently morph-bound. If that were so, then phonological exceptionality in any morph *m* would always be either (i) uniform throughout all phonological exponents of *m* or (ii) entirely predictably located within *m*. Yet this is not the case. If we accept something along the lines of Anderson's (**anderson1982**) analysis of French schwa as an exceptionally-deleting /ø/ vowel, then that exceptional property is neither uniform throughout morphs nor does it have a predictable location. Similarly, in Turkish, non-high round vowels are phonotactically exceptional outside the first syllable (**clements1982r**; **Van1991**), yet the location of the exception is not predictable, as seen in a comparison of *otoban* 'highway', *monoton* 'monotone', *fenomen* 'phenomenon' and *paradoks* 'paradox'. There is no doubt that in most known cases, exceptionality does happen to be either uniform or predictable within a morph, but this follows uninterestingly from the fact that most exceptional morphs are short, or that most phonological alternations are either local, in which case their location

inside a morph is predictably restricted to an edge, or domain-spanning, in which case the morph acts uniformly. However, when such uninformative cases are set aside, the small, informative residue of evidence does not support the morph-based view.

A second argument in defense of M-indices might be that morphs, and not φ elements, belong to lexical strata, and that a single morphological diacritic can therefore coherently index a whole set of phonological exceptionality patterns, patterns which impact different parts of the morph and which therefore would be only incoherently represented by individual diacritics on φ elements. Yet the empirical falsity of this claim has long been recognized. SPE (**chomskyhalle1968**) permitted both stratal diacritics, later labeled MORPHOLOGICAL FEATURES (**postal1968**) and more specific RULE FEATURES (**lakoff1970**), in view of the fact that distinct phonological patterns associated with strata are not uniformly attested in all morphs. For more recent work, see for example **labrune2012** on Japanese.

A third argument in defense of M-indices might be that since some kinds of phonological exceptionality are cyclic (§7.1), and since cycles are inherently tied to morphology, not φ elements, then something like M-indices are required anyhow, in which case ⊠-indices are redundant. I would suggest that this is a category mistake. While it is true that cycles are inherently tied to morphology, they are tied not to morphs, but to morphological operations. Some operations are non-concatenative and hence morph-free (**anderson1992**). Cyclicity effects, therefore, are about how phonological subgrammars correlate with OPERATIONS; in contrast, ⊠-indices are about correlation with FORMS. M-indices fall uncomfortably in between. Since they are inherently attached to morphs, they will be unavailable for the triggering of cyclicity effects associated with non-concatenative operations. And, as we have seen above, they are inefficient, and in all likelihood insufficient, devices for exceptionality of forms.

10 Conclusion

For most of the generative period, an implicit assumption has been that we must choose between a concrete phonological and a diacritic morphological approach to phonological exceptionality.[19] But the argument from learning is that the correct theory is phonological and diacritic, based on lexical phonological indices which are visible to the phonology but not manipulable by it. The concrete phonological approach, whose pseudo-indices are manipulable by the phonology, is ill-suited to efficient learning (§9.1). Diacritic approaches are well suited to learning; however the computation of exceptionality is simply not carried out in terms of morphs, rather its currency is lexical phonological elements. This is true for both constraint evaluation (§7.1) and the learning of co-indexation (§8.2). Concurrently, plausible assumptions about learning ensure that a diacritic phonological account does not suffer from overgeneration (§5.3), and reveal the need for a morphological analytic bias, operationalized here as the Morphological Coherence Principle (§6). Finally, a morph-based diacritic theory appears empirically insufficient in the in-

[19] Except, trivially, in purely abstract theories (e.g. **lamb1966 fudge1967**).

evitably small number of cases that are informative (§9.2). No doubt there is much more to be said on the topic of exceptionality, but I hope to have established that the nature of exceptionality is, in essence, phonological and diacritic.

Abbreviations

Abbreviations conform with the Leipzig glossing rules; in addition are: LEST 'lest' and SET 'inclusion/one of a group' (**dixon1977a**).

Chapter 3

Split-morphology and lexicalist morphosyntax: The case of transpositions

Andrew Spencer
University of Essex

One of Anderson's many contributions to morphological theory is the claim that morphology is split between syntactically mediated inflection and lexically mediated derivation. In Minimalist morphosyntax all morphology is syntax. This means that the split morphology proposal is not meaningful for that model. In lexicalist models, however, the split morphology hypothesis manifests itself as a distinction between direct accessibility to syntactic representations (inflection proper) and lack of accessibility. However, there are construction types which bring the inflection-derivation distinction into question. One of these is the transposition, as illustrated by the ubiquitous deverbal participle. This is a mixed category, being at once a form of a verb yet having the external syntax of an adjective. It is thus unclear which side of the split participles fall. Similarly, participles seem to be an embarrassment for the Word-and-Paradigm models of inflection which have become dominant since Anderson first introduced them to contemporary theorizing. This is because they seem to require us to define a 'paradigm-within-a-paradigm' (or 'quasi-lexeme-within-a-lexeme').

I provide an analysis of Russian participles within Stump's PFM2 model, deploying the model of lexical representation developed in **Spencer13:book** which fractionates representations into more finely grained subcategories than is usual. I take a participle to be the adjectival representation of a verb, coded directly by means of a set-valued feature, REPR. I show how a set of rules can be written which will define the adjectival paradigm as a set of forms belonging to the overall paradigm of the original verb lexeme. The rules define a partially underspecified lexical entry for the participle ('quasi-lexeme'), which has essentially the same shape as the lexical entry for an (uninflected) simplex adjective. Thus, the participle's lexical entry is that of an adjective, just as though we were dealing with derivation, but it realizes the verbal properties of voice/aspect and it shares its semantics and lexemic index with its base verb, just as in the case of verb inflection. The participle thus straddles the split, but in a principled fashion.

1 Introduction: Morphological architecture

Since **AndersonSR77:inflection** (re-)introduced to generative grammar the traditional notion of 'word-and-paradigm', and particularly the ground-breaking work of **Matthews72** on Latin inflection, morphologists have been grappling with the challenge of providing an adequate characterization of the key notions 'word' and 'paradigm'.

Central to this debate has been the fate of the Bloomfield/Harris interpretation of the morpheme concept (**AndersonSR15:morpheme**) and the notion of 'Separationism' (**Beard95:book**). The 'word' notion presupposes (at least) a word/phrase distinction. (In morpheme-based approaches no such distinction is necessary and all morphology and syntax can be subsumed under a model of morphotactics.) There are very well known problems with any attempt to find necessary and sufficient conditions for the 'concrete' instantiations of word — the phonological word, (inflected) word form, syntactic word, even. This is generally on account of incomplete grammaticalization, which strands constructions and formatives in a limbo between the status of function word– clitic– affix, compound element– affix, analytical syntactic construction – periphrasis and so on. However, the 'abstract' notions of word are no less problematic, specifically the lexeme and the morphosyntactic word (i.e. an inflected word form together with the morphosyntactic property array that it realizes). Defining the set of morphosyntactic words often requires us to make decisions about what constitutes a word form, which brings us back to the issue of clitics, periphrasis and so on. It also requires us to make sometimes arbitrary decisions about morphosyntactic property sets (MPSs) in the light of form-content mismatches such as (some types of) syncretism (**Baerman:etal05:book**), overabundance (**Thornton12:overabundance**), defectiveness (**Sims15:book**), and deponency (**Baerman:etal07**), as summarized in **Stump16:book**

Inflectional properties, and word-oriented functional categories generally, such as definiteness (for nouns) or modality (for verbs) in English, seem to presuppose an inflection∼derivation dichotomy that is notoriously hard to pin down. Broadly speaking it distinguishes the creation of new lexical items/units (generally, Saussurean signs pairing a cognitive meaning with a set of forms) from forms of a lexical item/unit. The component of grammar that defines new lexical units or lexemes is derivational morphology. However, as **Spencer13:book** itemizes in some detail, such a (canonical) inflection/derivation dichotomy represents just two poles of a scale of types of lexical relatedness. Some of the intermediate types of relatedness pose problems for any clean characterization of the lexeme concept (**Spencer16:individuating**).

One way of characterizing the core of the inflection/derivation distinction is the notion of split morphology (**AndersonSR82**). The essence of the distinction can be thought of as an interface claim: inflection interfaces directly with syntax ('inflection is what is relevant to syntax'). The obverse to this claim is that derivation interfaces with the lexicon, in the sense that derivation is what gives rise to expansion of the lexical stock (as well as defining relatedness between already fixed lexical entries), in other words derivation is 'what is relevant for the lexicon'. Anderson implements this architectural claim by saying that it is the rules of syntax themselves which define inflectional morphology.

This move raises the important question of what model of syntax we are presupposing. Most versions of the Minimalist Program presuppose something very close to the model of morphotactics proposed by the American Structuralists: the atoms of representation are morphemes, morphology and syntax are identical (it is therefore a terminological choice whether we think of sentence construction as morphotactics or syntax), and the notions lexeme, word form, inflection, derivation, inflectional paradigm are at best heuristic descriptive terms which cannot be given a coherent definition within the model. The natural syntactic framework for investigating a word-and-paradigm, or rather, lexeme-and-paradigm approach to inflection is, perhaps, a lexicalist, or constraints-based, model (**Miller:Sag97**; **Sadler:Spencer01**; **Sadler:Nordlinger06:stacking**). In that case the question of split morphology takes on a somewhat different aspect. Rather than claiming that syntactic rules *construct* inflected forms as such, we must say that inflected forms, compared with derived lexemes, are permitted to interact in specific ways with syntactic representations, or equally that inflected forms bear properties which are visible to syntactic representations and principles.

The obvious way to implement this idea is to say that the abstract characterization of an inflected word includes a morphosyntactic description which overlaps with that of a corresponding syntactic representation. A concrete version of this type of overlap is seen in the form-content mapping, as defined in Stump's notion of paradigm linkage (**Spencer:Stump13:Hungproncase**; **Stewart:Stump07**; **Stump02:paradigmlinkage**; **Stump06:heteroclisis**; **Stump16:book**; **Stump16:MorphMetatheory**). Stump distinguishes morphological properties, the FORM paradigm, from syntactic properties, the CONTENT paradigm. By default these are homologous, but there are many instances of mismatch. For example, Latin syntax distinguishes singular and plural number and a variety of cases, including dative and ablative, but those two cases are never distinguished morphologically for any lexeme in the plural. On the other hand, Spanish verbs have two distinct subparadigms for the imperfect subjunctive but that distinction is nowhere reflected in the syntax. Other mismatches include deponency and periphrasis. To a limited extent we can say that the form-content paradigm distinction is a reflex of covert split morphology: such a distinction is not definable for derivational morphology.

In lexicalist models, the derivational morphology $\sim$ lexicon interface operates over property sets which don't play a direct role in syntax. The hedge 'direct' is important: typically, derivation *is* relevant to syntax, in the sense that it changes a lexeme's morphosyntactic class. More subtly, derivation may make appeal to argument structure realization (witness English Subject Nominalizations, *able*-Adjective formation and so on). But if it is assumed that lexemes have a representation of their word class argument structure and other relevant properties, then such syntactically expressed relations can be defined over lexical representations, as extensively argued in the constraints-based literature (**Wechsler14:book**). This is effectively a statement of the doctrine of lexical integrity, at the abstract level of representation as defined by **Ackerman:LeSourd97**

The conclusion to be drawn is that morphology interfaced with a constraints-based syntax needs to be split in essentially the way proposed by Anderson, but as an abstract architectural property, which sometimes bears a rather complex relation to concrete mor-

phophonological expression. Inflectional morphology maps to syntactic representations in a way in which derivational morphology is unable to, while derivational morphology serves to define (specific kinds of) lexical relatedness. However, there remain interesting cases of violations of lexical integrity with derivational relations, in which syntax appears to have access to the internal structure of the derived word. This paper will argue that such phenomena require us to extend the scope of the split in morphology in a way that takes the notion of 'lexeme' as syntactic atom seriously, and which ultimately provides conceptual motivation for a lexicalist interpretation of split morphology.

A case in point is the class of denominal (relational) adjectives in many languages, which allow one noun to modify another by taking on the morphosyntax of an adjective. In European languages, including English and Russian, such adjectives respect lexical integrity, in the sense that the base noun is opaque to syntax. For example, the base noun *kniga* 'book' in the Russian relational adjective *knižnyj* 'pertaining to a book/books' is opaque to agreement, government or any other syntactic process, just as in English the noun *tide* in *tidal* is opaque. For instance, the phrases *poderžannaja kniga* 'second-hand book' and *knižnyj magazin* 'book shop' do not gives us **poderžannaja/poderžannyj knižnyj magazin*, and although we can say *high tide* and *tidal barrier* we can't say **high tidal barrier*. The importance of these observations is that there are languages in which just such attributive modification into a derived adjective is possible (see the discussion of Tungusic and Samoyedic examples in **Nikolaeva08** and also the detailed discussion of the Samoyedic language Selkup in Spencer 2013, chapter 10).

The relational adjectives of Russian and English, however, share one important property with the Tungusic and Samoyedic pure relational adjectives, namely, they have precisely the same lexical semantics (cognitive content) as the base noun. This leaves us with the question of how to explain why in some languages relational adjectives are opaque and in others they are transparent to attributive modification.

Spencer (2013) argues that the crucial difference between true transpositional relational adjectives of, say, Selkup, and the 'fake' transpositions of English/Russian is that true transpositions are effectively forms of the base noun lexeme, while the English/Russian relational adjectives are distinct lexemes, though ones which have a semantic representation identical to that of their base, what Spencer (2013: 275) calls a transpositional lexeme. Other types of transpositional lexeme include English property nominalizations (*kindness, sincerity, …*), deverbal nominalizations such as *destruction*, and participial forms which have been converted in qualitative adjectives such as *(very) boring/bored, charming, excited, ….* A relational adjective which is a true transposition permits inbound attributive modification because it is, in an important sense, still a noun, just as a noun stem marked for number, case, possession or definiteness is still a noun.

One consequence of this reappraisal of the morphology∼syntax interface is that the crucial divide can no longer be straightforwardly equated with a traditional inflection/derivation distinction. It is not appropriate to think of a deverbal participle or a relational adjective as merely an inflected form of a verb or noun, because that participle or adjective will in general inflect like an adjective, not like a verb/noun. However, following **Haspelmath96** Spencer (2013, chapter 10) argues for an enrichment of the traditional notion of the inflec-

tion paradigm to include an attribute REPRESENTATION (taken from Russian descriptive practice). Thus, a participle is the adjectival representation of the verb, and as such it can have its own adjectival inflectional paradigm. It thus has the outward appearance of an autonomous lexeme, but appearances are deceptive. Rather, the transposition is a 'quasi-lexeme', and the transpositional relationship therefore represents a particularly striking instance of a deviation from inflectional canonicity.[1] Like transpositions, these are not usually described under the heading of morphology-syntax mismatches, and like transpositions they are not discussed in Stump's (**Stump16:book**) otherwise very detailed survey.

In the model of lexical representation argued for in Spencer (2013) the notion 'form of a lexeme' in this somewhat extended sense is reflected very simply: all forms of a lexeme share their Lexemic Index. This leads us to propose a (no doubt too strong) hypothesis about the nature of the split in morphology:

Principle of Lexemic Transparency
Let D be a word derived by some regular morphological process from a word B, possibly of different morphosyntactic category. If morphosyntactic processes treat D in the same manner as they would treat the base word, B, even where the category of D is such that we would not otherwise expect it to be subject to such processes, then D is a form of the lexeme B (shares B's Lexemic Index).

I shall argue that the architectural equivalent of splitting inflectional from derivational morphology is this modification of the notion of lexical integrity: if morphology defines a set of forms of a lexeme, rather than defining a new, autonomous lexeme, then those forms will show lexical transparency. Derived lexemes, however, show lexical opacity (one reflex of which is the more familiar property of lexical integrity). This paper will illustrate that proposal on the basis of the behaviour of Russian deverbal participles. These are particularly useful. First, Russian adjectival morphosyntax is very clearly distinguished from noun or verb morphosyntax, so it is easy to show that the participles behave like adjectives. Second, the Russian past tense and conditional mood are expressed by a form which is historically a participle and which show participle-like agreement, but which has been reanalysed as a verb form (the l-participle). This contrasts in important ways with the true participles. Third, like many languages, Russian often converts its participles into true qualitative adjectives. These have (almost) exactly the same set of forms as the participles but their syntax is no different from that of a simplex adjective. The true participles are like the l-participle in showing lexical transparency with respect to the base verb. They therefore both appear on the inflectional side of the split morphology. This is because they are forms of the verb's paradigm, and do not constitute independent lexemes in their own right. They contrast with the converted participles,

[1] Other such deviations are certain forms of evaluative morphology (cf 'the diminutive *form of* a noun') and grammaticalized argument structure alternations (cf 'the passive/anti-passive/applied/causative *form of* a verb').

which are autonomous lexemes and hence opaque with respect to the verb properties of their (etymological) base lexeme.

2 Lexical representations and lexical relatedness

2.1 Introduction

Our discussion will require us to be explicit about a number of aspects of lexical representation and the way that words, in the broadest senses of the term, are related to each other. I shall adopt a generalized form of Stump's (**Stump01:book**) Paradigm Function Morphology, which I have called *Generalized Paradigm Function Morphology*, GPFM (Spencer 2013). The GPFM model is designed to permit us to use the machinery of PFM to describe types of lexical relatedness which go beyond normal inflectional morphology. It thus extends the lexical representations that morphology has access to by incorporating representations of syntactic properties and the lexical semantic representation of words. In GPFM lexemes have to be individuated by means of an arbitrary index, the Lexemic Index (defining something like the key field in a database). One of the reasons for this is because it is arguably not possible in the general case to individuate lexical representations of lexemes in terms of any of the linguistically relevant properties that can be ascribed to a lexical representation. In addition, however, the index serves an important role in distinguishing certain types of morphological relatedness.

2.2 Lexical representations

I begin with the descriptive representational apparatus required to characterize an inflected word form, taking inflection to be an uncontroversial category for the sake of exposition. I then generalize the representational format to provide a characterization of the lexemic entry.

A word has a minimum of three contentive attributes (together with a fourth, its Lexemic Index, LI): FORM, SYN(TAX), SEM(ANTICS). The SYN attribute records idiosyncratic selectional or collocation properties, but its main component is the argument structure attribute, ARG-STR. This records thematic argument arrays in the standard fashion (notated here as x, y, … variables). However, it also includes a semantic function (sf) role.

For nouns and verbs the sf roles are the 'R' and 'E' roles respectively, familiar from the literature. The 'R' (for 'referential') argument is canonically associated with lexical entries whose SEM value belongs to the ontological class of *Thing*. It thus identifies those predicates that typically denote (concrete or abstract) objects and that can serve as the lexical head of a referring expression, i.e. a canonical noun. Thus, the 'R' argument of TREE[2] corresponds to the 'x' variable in the semantic representation λx.TREE(x). It is the argument that is the target of attributive modifiers: *(tall) tree* (**Spencer99:transpositions**). See **Lieber04:book** for concrete examples of the R role being deployed in morphology.

[2] Where relevant, I adopt the standard convention of putting the name of a lexeme in SMALL CAPS.

The 'E' (for 'event(uality)') argument (sometimes written as 'e' or 's' (for 'situation')) is canonically associated with lexical entries whose SEM value belongs to the ontological class of *Event*. It thus identifies those predicates that typically denote states or events (eventualities) and that can serve as the lexical head of a clause, i.e. a canonical verb. Thus, the 'E' argument of FALL corresponds to the 'e' variable in the (neo-Davidsonian) semantic representation λe,x.FALL(e,x). For attributive modification (principal role of the traditional adjective class) I assume a semantic function role labelled 'A'. This is coindexed to a noun's R sf role to represent attributive modification. All adjectives which function as attributive modifiers, including relational adjectives and participles, have an 'A' semantic function role.

I assume the SEM attribute is essentially a formula in predicate calculus defined over the ontological types *Thing, Event, Property* (**Jackendoff90**), and perhaps others, corresponding loosely to the morphosyntactic categories of Noun, Verb, Adjective. I remain here agnostic as to whether the categories N, V, A are universal and if so, in what sense. I assume that some languages also have a category of Adverb, and also transpositional morphosyntax, adjective-to-adverb (as in English *ly*-suffixation), verb-to-adverb (gerund) and noun-to-adverb (found in Selkup, for example), but I will not have much to say about that category here.

Adpositions may mandate a further ontological category of, say, *Relation*, but I ignore that too. The SEM attribute can be thought of as a label for an encyclopaedic representation, such as λx.CAT(x) or λx,y.WRITE(x,y), but sometimes including linguistically encoded information relevant to semantic interpretation that cannot simply be consigned to an undifferentiated encyclopedia, for instance, λx,y,z.SIMILAR_TO(x,y,δ) ∧ CAT(y) ∧ DIMENSION(δ), 'similar to the property of 'cat' in some dimension, δ', or λx,y.AGAIN(WRITE(x,y)) 'to re-write something'.

The FORM attribute is essentially a record of the word's morphology. Assuming an articulated inflectional system, complete with arbitrary inflectional classes and possibly other purely morphological, paradigm-based properties, the FORM attribute needs to specify all the information needed to locate the word form in the appropriate inflectional paradigm. The first property is the morpholexical category, MORCAT. This will typically be derived by default from the syntactic category of the representation (the SYN|CAT attribute),[3] but that default mapping is not infrequently overridden, sometimes in rather complex ways.

The next property is largely defined by reference to the syntactic category of the word form/lexeme, namely, the morpholexical signature, MORSIG. This specifies all those morphosyntactic properties for which an element of that MORCAT is obligatorily inflected. An inflected word has to have a specification of the morphosyntactic property set (or sets), MPSs, that it realizes. In the case of syncretism this may be a (natural or unnatural) class of MPSs. For example, a Russian adjective is obligatorily inflected for at least

[3] Given the complexities of category mixing it is better to dispense entirely with morphological or syntactic category labels such as 'verb', 'adjective'. The required lexical classes can be defined over other aspects of representation much more efficiently and it is not difficult in constraints-based models to ensure that those aspects of representation are accessible to rules of morphosyntax. However, for convenience of exposition I will continue to talk of (morphological or syntactic) verbs, adjectives and so on.

the properties of number, gender and case, and these features are therefore listed in the MORSIG (a gradable adjective is also inflected form comparative and superlative forms). We will see in §?? that the conception of MORSIG assumed in Spencer (2013) can be enriched and extended to include the FORM-CONTENT paradigm distinction introduced in Stump (2002) and subsequent work.

Finally, the representation has to specify a phonological form for the word, through an attribute FORM|PHON. The precise characterization of the PHON entry, in general, will be given by the rules of inflectional morphology. I will assume that one aspect of the PHON representation will be a specification of the STEM on which the inflected form is based, but I ignore this refinement because it will not be relevant to the question of split morphology.

The actual inflected word forms of the lexeme are defined by inflectional rules, which apply to the pairing $\langle £, \sigma \rangle$, where σ is a complete, permissible feature set for the lexeme with Lexemic Index £. The lexemic representation has to include all the idiosyncratic morphological information relevant to a lexeme's realized paradigm. In the next subsection I summarize the way that the Paradigm Function can be generalized to define not only inflection but all the systematic forms of relatedness.

One aspect of these representations is worth noting. In keeping with the inferential-realizational assumptions underlying our model of inflection the SEM attribute remains constant for all inflected forms. In particular, there is no characterization at the level of the lexical representation of a word form (much less the level of lexemic representation) of the semantics of, say, tense or number. What this means is that in the syntax the VP which is headed by a past tense form verb may, ceteris paribus, be interpreted as referring to an event situated prior to speech time. However, since 'past tense' forms are also used in sequence of tense constructions, irrealis conditionals and so on, 'past time' is only the default interpretation.

2.3 Lexical relatedness

We can now ask what types of systematic relatedness lexemic entries (i.e. lexemes) can exhibit. Spencer (2013) argues extensively that we can find pretty well all the logically possible types as defined by the very crude but simple artifice of defining non-trivial differences in the four principal attributes, FORM, SYN, SEM, LI. For instance, if we consider pairs of representations of distinct lexemic entries, $£_1$, $£_2$, i.e. those with distinct LIs, then we can identify several different types of relatedness (usually all treated as derivation).

Suppose that the lexemes $£_1$, $£_2$ are distinct in FORM, SYN, SEM representations. Suppose also that the FORM/SEM representations of $£_2$ subsume or properly include (in some sense) those of $£_1$. Then we have standard (canonical) derivational morphology, DRIVE ⇒ DRIVER. Languages sometimes define derived lexemes without changing the FORM attribute at all, however. A case in point is that of adjectives which are converted wholesale to nouns without any change in morphology (Spencer 2002; see also the discussion of *Angestellte(r)* nouns in Spencer 2013). We will later see examples of deriva-

tion in which FORM/SYN/LI attributes are changed but without changing the meaning (transpositional lexeme).

Now let's consider what happens if we keep the LI constant, that is, we consider intra-lexemic relatedness. To begin with, let us assume that only the FORM attribute can change. In the canonical case this is the same as inflectional morphology. Ignoring for the present the transpositional interpretation of the participles as a verbal noun or as verbal adjectives, we can say that all inflected forms of SING are forms of a verb. We have to be a little cautious when referring to syntactic properties: the syntactic distribution of any given inflected form is, in general, distinct from that of other forms. The 3sg subject agreement form of a verb does not occur in the same syntactic positions as the 3pl form. The properties that give rise to these differences, however, are precisely the MPSs which bifurcate into FORM/CONTENT paradigms. This means that we must enrich lexical representations in the obvious way: FORM paradigms are defined over features typed as FORM features, and CONTENT paradigms are defined over features typed as SYN features, with the proviso that the two sets of features are identical by default. Modulo the CONTENT paradigm, then, in canonical inflection the SYN value of a given word form is identical to that of the other forms of that lexeme. This means that most inflection is what Booij (**Booij94**) would call contextual inflection.[4]

Likewise, the lexeme SING denotes the same event type in all of its inflected forms, and in that sense all word forms share the same SEM representation. In Spencer (2013) I argue that there are certain types of inflection that enrich the semantic representation of the base lexeme, whilst remaining inflectional. Certain kinds of Aktionsart marking, as well as semantic case marking often have this characteristic, as do causative argument structure alternations. In traditional descriptions of languages with such inflection we often find terminological vacillation, as linguists are unsure whether to label, say, the iterative form or the causative of a verb inflectional or derivational (and similar problems afflict evaluative morphology).

The GPFM descriptive framework proposed in Spencer (2013) generalizes the PFM model so that all forms of lexical relatedness, from contextual inflection to derivation, are defined over four principal attributes of a lexical representation. This requires us to generalize the notion of the Paradigm Function to that of a Generalized Paradigm Function (GPF), which is like the Paradigm Function except that it consists of four component functions, f_{form}, f_{syn}, f_{sem}, f_{li}. For canonical derivation the GPF introduces non-trivial changes to all four components (including the LI). For the converted adjectives and *Angestellte(r)* nouns the f_{form} function has no effect (it can be thought of as a kind of identity function). For most inflection, the $f_{syn,sem,li}$ functions are the identity function, because the GPF simply realizes inflectional properties of the lexeme at the FORM level.

In the GPFM model the lexemic representation is defined in terms of the Lexemic Index and a completely underspecified (empty) feature set, *u*, for example, ⟨PUT, *u*⟩, a special case of the GPF. This maximally underspecified GPF defines just those properties of a

[4] This includes Booij's parade examples of inherent inflection, past tense and plural number. See Spencer (2013: 77–82) for critical discussion of Booij's distinction.

lexeme (identified by its Lexemic Index) that are completely idiosyncratic and completely unpredictable. However, although such a representation reflects the traditional notion of a maximally compact, non-redundant dictionary entry, it is not a representation that can serve as the direct input to rules of inflection. This is because the lexemic entry has to be specified for those inflectional properties that it can and must inflect for. This set is defined by the morphological signature, MORSIG. In Spencer (2013: 199) I make this rather obvious point explicit in the *Inflectional Specifiability Principle*. In the current context this can be stated as follows: a lexeme is inflected for a given MPS, iff that MPS is defined in its MORSIG.

In Spencer (2013) I treat the MORSIG attribute as part of the FORM paradigm of a lexeme. However, we know that the FORM and CONTENT paradigms of a lexeme can differ substantially. For this reason, it is necessary to enrich the SYN attribute of a lexemic entry with a (possibly distinct) MORSIG attribute. The values of the (FORM and CONTENT) MORSIG attribute are for the most part predictable from other aspects of the lexical representation. First, the FORM MORSIG attribute is generally projected from the CONTENT MORSIG and by default the two attributes are identical. Second, to some extent the meaning of the lexeme can determine the content of the MORSIG attribute. Most importantly, however, the MORSIG (which, recall, is essentially a record of the properties for which a lexeme inflects) is largely projectable from various aspects of the SYN attribute, notably the ARG-STR attribute. Thus, a lexeme with the SYN|ARG-STR value ⟨E⟨ x, …⟩⟩ (i.e. a verb) will by default have the syntax and morphology of a verb. The lexemic representation needs to be enriched to include purely idiosyncratic information, such as irregular stem forms, irregular inflections, defective cells or subparadigms, and so on. Technically, this can easily be achieved in GPFM by defining very specific functions for particular properties over the LIs of the lexemes concerned. For instance, the irregular past tense of PUT can be defined by a function defining the STEM_{pst} form for the pairing ⟨PUT, *u*⟩: GPF(⟨PUT, {STEM_{pst}|PHON}⟩) = /pʊt/ or similar. This will override any less specific (in practice, any other) statement of past tense morphology. Similarly, a defective lexeme such as FORGO (lacking a past tense form) will have a tense-specific GPF(⟨FORGO, {tense:pst }⟩) = undefined. This again will override any other statement, including the GPF(⟨GO, {tense:pst }⟩) = /wɛnt/, which applies to one other verb based on GO (cf *underwent*) and hence is less specific.

The role of the MORSIG attribute can be simply illustrated by the English plural. Any lexeme with the SYN|ARG-STR|⟨R ⟩ value licenses MORSIG|num:{sg, pl}, provided that its SEM attribute specifies it as a count noun. For a mass noun the MORSIG value will be just num:{sg}, while for a plurale tantum noun it will be num:{pl}. The {pl}, resp. {sg} values for such nouns are therefore undefined, so that a GPF(⟨SINCERITY, {num:pl}⟩) or GPF(⟨SCISSORS, {num:sg}⟩) will not correspond to any legal output.

In summary, I assume a representation for a dictionary entry of a very traditional kind: it is minimally redundant, specifying just the unpredictable phonological and semantic information, together with any morphosyntactic information that cannot be projected from the phonological and semantic specifications. For an entirely well-behaved lexeme belonging to the default inflection class for that word category, this is all the information

that is required, but that is only because the morphological signature can be projected from that entry too.

To all intents and purposes GPF collapses with PFM for most cases of inflectional morphology. However, the model is designed to cover all types of lexical relatedness, up to regular derivation, but using essentially the same machinery as PFM. For instance, for the derivation of a Subject Nominal such as DRIVER from DRIVE we would have the partial GPF shown in (??), where $\mathcal{V}$ is the Lexemic Index of a verb lexeme and δ is the derivational feature which defines the Subject Nominal formation process.

(1) a. $f_{li}(\langle \mathcal{V}, \delta \rangle) = \boxtimes(\mathcal{V})$, where $\boxtimes$ is a function over LIs corresponding to the derivational feature δ

b. $f_{sem}(\langle \mathcal{V}, \delta \rangle) = [_{\textit{Thing}}\ \lambda x, \text{PERSON}(x) \wedge \mathcal{P}'(x)]$, where $\mathcal{P}'$ is a suitable form of the semantic representation of the lexeme $\mathcal{V}$

c. $f_{syn}(\langle \mathcal{V}, \delta \rangle) = u$

d. $f_{form}(\langle \mathcal{V}, \delta \rangle) = Z \oplus er$, where Z is the δ-selected stem form of $\mathcal{V}$

This corresponds to a novel lexemic entry which is exactly like that in (??) except that it is defined over the pairing $\langle \boxtimes(\mathcal{V}), u \rangle$. If DRIVE is the LI of the verb DRIVE, then $\boxtimes(\mathcal{V})$ is the LI of the derived lexeme DRIVER and $\langle \boxtimes(\mathcal{V}), u \rangle$ defines its lexemic entry.

The representation in (??) lacks any syntactic specification. This is because that specification can be given by a default mapping from the SEM attribute, by virtue of the Default Cascade (Spencer 2013: 191–194). Under that principle, regularly derived lexemes have their principal morphosyntactic properties projected from their semantic representations, in accordance with the notional model of parts of speech. Thus, DRIVER is of ontological category *Thing* and therefore by default has the semantic function role R. I assume that the SEM attribute includes an indication that the lexeme denotes a countable *Thing* so that the lexeme licenses the full MORSIG|NUM:{sg, pl}. However, the lexeme in (??) is derived, not simplex. In Spencer (2013) I propose a Category Erasure Principle, under which the morphosyntactic properties of a base lexeme in derivation are deleted so that they can be overwritten by the Default Cascade. However, given maximally underspecified lexemic entries to start with this is probably not necessary: the word formation rule interpretation of the GPF takes the only information it has available in a lexical entry, that is the phonology of the root and the meaning, and modifies it, say, by adding an affix, and by enriching the SEM representation systematically, for instance, by adding a predicate. The Default Cascade then specifies the underspecified properties, including the two MORSIG representations. If the base lexeme's entry includes non-default specifications such as irregular inflected forms or non-standard argument realization, these will override the Default Cascade. Illustrative examples of simple inflection and derivation are provided in the Appendix.

I have sketched the GPFM approach to inflection on the one hand, and standard derivation on the other hand. Non-standard types of derivation are handled by overriding some of the defaults in the GPF. However, this still leaves us with the somewhat intriguing, but very widespread set of relatedness types in which FORM/SYN attributes are altered, much as in derivation, but without changing either the SEM representation or, crucially,

the LI attribute, one of the class of relatedness types often called a 'transposition'. The parade example of this type is, perhaps, the deverbal participle, the 'adjectival representation' of a verb.

Transpositions pose particular difficulties for a simple interpretation of the inflection-derivation divide (and, indeed, for any architecture with the equivalent of split morphology): on the one hand, a participle is generally taken to be 'a form of' the base verb (regular participles are never given their own entry in a traditional dictionary, for instance). On the other hand, a participle is morphologically and syntactically a kind of adjective. If morphology is split, on which side do transpositions fall?

3 Paradigm Linkage: the problem of transpositions

Spencer:Stump13:Hungproncase (**Bonami:Stump16:PFM**; **Stump16:book**) describe a variant of the Paradigm Function Morphology model, PFM2, which explicitly distinguishes two types of paradigm. FORM paradigms determine the mapping between MPS's and the word forms realizing them; CONTENT paradigms determine the set of grammatical distinctions a lexeme needs to make in a syntactic representation. The two paradigms are related by rules of paradigm linkage. In the default case the two paradigms align perfectly: a morphological plural noun is syntactically plural and vice versa. However, there are numerous mismatches. A simple case in point is provided by English perfect and passive participles. These have distinct syntax (they collocate with different auxiliaries, and only the passive participle can be used attributively), but they are never distinguished morphologically. Thus, the (single) FORM property (**Aronoff94:book**) has to map to two CONTENT properties. **Stump16:book** provides an extensive survey of the principal types of mismatch encountered in the world's inflectional systems.

I will argue that we can regard the FORM/CONTENT paradigm distinction as a reflex of split morphology for constraints-based lexicalist models of syntax. The reasoning is very simple — no corresponding FORM/CONTENT distinction can be mandated for derivation. Indeed, given standard PFM2 assumptions it is difficult to imagine what such a distinction could mean.

Incorporating the FORM-CONTENT paradigm distinction into GPFM has consequences for the way in which lexical representations are organized. We have seen that the FORM attribute of a lexical representation is defined in part in terms of its MORSIG, which determines precisely those properties a lexeme must inflect for. In the original GPFM model the MORSIG attribute is only defined at the level of FORM paradigms, therefore. Since CONTENT paradigms are not always congruent to FORM paradigms we will need to draw the appropriate distinction in lexical representations. The obvious way to do this is to assume that each lexeme is associated with *two* (at least) MORSIG attributes, one of which is a value of the FORM attribute and the other of which is a value of the SYN attribute. We can call these f-MORSIG, c-MORSIG respectively.[5]

[5] A number of authors have proposed that cells in a lexeme's paradigm can be realized by multi-word, periphrastic, constructions (**Sadler:Spencer01**; **Brown:etal12:periphrasis**; **Bonami15:collocation**). **Popova:Spencer17:FDSL11** following suggestions by **Bonami15:collocation** propose that such construc-

Stump16:book defines paradigm linkage in terms of the Paradigm Function and a ***Corr*** function. For lexemic index £, a set of MPSs σ, τ, a stem form, Z, the function ***Corr***(⟨£,σ⟩) delivers a form correspondent ⟨Z,τ⟩, such that if PF(⟨Z,τ⟩) = ⟨w,τ⟩, then PF(⟨£,σ⟩) = ⟨w,τ⟩. By default the FORM and CONTENT features are identical, that is the set $\sigma = \tau$, as defined by a set of property mappings (***pm***), that is, ***pm***(σ) = σ. However, the function ***pm*** is called upon whenever there is a mismatch between FORM and CONTENT properties. This is the case, for instance, with deviations such as syncretism, deponency and so on. In the case of active∼passive deponency the property mapping maps the morphological passive voice forms to the syntactic active paradigm and leaves the morphological active voice forms undefined.

We can ask how paradigm linkage would work for derivation, by taking the example of a derivational feature, δ, such as the privative denominal adjective feature, *privadj*, which derives *friendless* from *friend*, as described in Stump (2001:252–60).[6] Here, we are dealing with trivial (two-celled) paradigms, so that the featural mismatches of inflectional paradigms will not be found, and we can work with just a single derivational feature, δ.

(2) Let ***Corr***(⟨FRIEND, *privadj*⟩) be the correspondence function for *privadj* applied to the lexeme FRIEND.
Corr(⟨FRIEND, *privadj*⟩) = ⟨Z,δ⟩, where Z = |friend|, the stem of FRIEND, and δ = *privadj*. If PF(⟨Z,δ⟩) = ⟨w,δ⟩, then PF(⟨£,δ⟩) = ⟨w,δ⟩, hence, PF(⟨FRIEND, *privadj*⟩) = ⟨friendless, *privadj*⟩.

This seems very straightforward, but there is a hidden difficulty, not immediately apparent from a language like English, with limited inflection. Consider a hypothetical language just like English but in which nouns and adjectives have entirely different inflectional paradigms, or, indeed, consider a derivational process such as that of Subject Nominalization which derives DRIVER from DRIVE. What we have to ensure is that the output lexeme is (automatically, by default) inflected as a noun, as opposed to the base lexeme which is inflected as a verb. As it stands, the Paradigm Function applied to a pairing of Lexemic Index and derivational feature will not deliver what Stump calls the realized paradigm of the output lexeme. At best, it might define the (or an) uninflected stem form of the derived word. Moreover, without significantly altering the nature of the ***Corr*** function it will be impossible to define additional lexical information such as inflectional class membership, idiosyncratic stem forms or other deviations from default inflection, or indeed any non-default or purely morphological property of the output.[7]

A direct solution to this problem might be to enrich the content of the derivational feature, δ, so that it incorporates the full set of paradigmatic oppositions realizable by the derived lexeme. Thus, the Subject Nominal feature (say *subjnom*) could be defined

tions demand additional CONTENT feature sets specifically to define the content of the periphrastic expression, which is often at odds with the default feature content of the words which make up that expression. Periphrasis therefore provides substantial motivation for a FORM/CONTENT or m-/s-feature (**Sadler:Spencer01**) distinction.

[6] Stump does not discuss derivation, or, indeed, transpositions in his works on paradigm linkage.

[7] It is also not clear how the additional semantic predicate of the output would be defined.

as the complex [*subjnom*, NUMBER:α]. This would mean that we would have to define the Paradigm Function so that it defines each inflected form of the output lexeme, rather than defining an underspecified lexemic entry, which then gets inflected by the standard inflectional rules applying to words of that category. For instance, the Paradigm Function could take the form PF(⟨DRIVE,{*subjnom*, NUMBER:sg}⟩) = ⟨driver, {*subjnom*, NUMBER:sg}⟩, PF(⟨DRIVE,{*subjnom*, NUMBER:pl}⟩) = ⟨drivers, {*subjnom*, NUMBER:pl}⟩. We could call this approach the 'full-listing approach'.

Now, the full-listing approach would have the rather peculiar consequence that the word forms {*driver, drivers*} would be inflected forms of the verb DRIVE, and there would be no such thing as a DRIVER lexeme. Not only is this entirely counter-intuitive, and at variance with any sensible distinction between inflection and derivation, it becomes even more counter-intuitive when we see recursive derivation. It would entail that the noun *reprivatizability* was an inflected form of the adjective *private*. The problem is that the ***Corr*** function needs to be able to define not a cell in a realized paradigm (what syntacticians refer to, misleadingly, as the lexical entry of a word(form)), but it has to define a featurally underspecified lexemic entry (an autonomous dictionary entry in traditional terms). I therefore reject the full-listing approach in favour of that adopted in the GPFM model.

We are now in a position to examine the case of transpositions. The problem we must address is how to ensure that the transposition is assigned to its own inflectional paradigm, proper to its new morphosyntactic category, whilst still in some sense remaining part of the inflectional paradigm of the base lexeme. Recall that I have proposed adopting a class of features, [REPRESENTATION:{...}] to define the paradigm space of a transposition. I now consider the way this feature can be deployed to define the inflectional paradigm-within-a-paradigm of a transposition. Let ρ = [repr:κ] for some transpositional relation κ, e.g. a verb-to-adjective (participle) transposition. I assume that the GPF applied to the pairing $\langle £,\rho \rangle$ defines a partially specified representation for the transposition. Normally, when the Paradigm Function applies to a pairing of LI and MPS the MPS has to represent a complete and coherent set of features sufficient to define the inflected form. However, this is not strictly speaking a property of the PFM system itself. In principle, we could define a partial paradigm for a lexeme by reference to just a proper subset of the features required to define any fully inflected word form. For instance, suppose that verbs in a language inflect for a variety of tense-aspect-mood-voice (TAMV) series, and that in addition they show subject agreement. We could, in principle, define the stem sets for the TAMV categories independently of the agreement morphology, by simply not specifying the subject agreement properties, effectively defining a set of 'screeves' for the language (**AndersonSR92:book**). This makes use of the same notion of feature underspecification used in GPFM to define derivation, but relativized to a specific feature set.

The GPF for a transposition has to specify the derived morphosyntactic category, κ, and, by default, the f-/c-MORSIG attribute, associated with that category, together with additional morphological properties inherited from the base, such as TAMV properties. If the Paradigm Function were to apply in the same way for transpositions as for other

inflected forms then the ***Corr*** function would have to provide the form correspondents for each inflected form, but it would not be able to provide a lexemic entry for the uninflected transposition. In other words, standard application of PFM2 principles would give rise to a full listing interpretation of the paradigm. It would not be possible to provide any characterization of the transposition as a quasi-lexeme (see p. **??**). This generates many of the same conceptual and technical problems as the application of ***Corr*** to derivation. One additional consequence is that it would be difficult to describe the very common situation in diachronic change in which a participle is reanalysed as a qualitative adjective, often without significantly changing the lexical semantics of the original verb lexeme (transpositional lexeme). If we treat the participle as akin to a lexeme, then we can easily define the diachronic reanalysis over that representation, just as we would for ordinary derivational conversion. One of the problems with the full-listing approach is that it is not clear how we could permit the transposition to inherit MORSIG properties of the derived category (Adjective in the case of participles). We therefore need to define the GPF in such a way that it allows us to define a quasi-lexeme as part of the paradigm of the base.

First note that application of the Default Cascade to define derived categories is entirely excluded in the case of (pure) transpositions, since these preserve the meaning, and hence the ontological category, of their base; indeed, this is the whole point of a transposition. The most natural assumption is that the transpositional morphosyntax effects a shift in the syntactic categorization and that morphological recategorization falls out as a consequence. (In fact, the extent to which the transposition acquires derived categorial properties and loses those of its base is subject to a good deal of cross-linguistic variation, tempered by poorly understood typological tendencies. See **Malchukov04**:**book** for discussion in the context of action nominal transpositions.) This is effectively a weaker instantiation of the Default Cascade. Following Spencer (**Spencer99**:**transpositions** 2013), I will assume that the shift in syntactic representation is actually a modification of the argument structure representation; specifically, the definition of a complex sf role (see below). The crucial point is that the transpositional GPF changes the SYNCAT value of the base lexeme to that of a mixed category and this, ceteris paribus, will automatically entrain a shift in the morphological category and hence the MORSIG attributes.

The basic machinery is only hinted at in Spencer (2013, chapter ten). In order to develop an explicit account we first need to refine the definition of the REPR(ESENTATION) attribute. I will therefore assume that the REPR feature takes an ordered pair as value, not a singleton element, [REPR:⟨κ,λ⟩], where κ, λ range over lexical categories. Thus, a participle represents a verb as an adjective and hence will bear the specification [REPR:⟨V,A⟩], while a predicatively used noun will have the specification [REPR:⟨N,V⟩]. We then couple these feature values to syntactic category specifications in the obvious way, stated informally in (**??**).

(3) Given [REPR:⟨κ,λ⟩], for λ = N, V, A.
Then SYNCAT = λ,
and by default, MORCAT = λ

In the GPFM framework the [REPR] attribute will be associated with an appropriate enrichment of the semantic function role of the ARG-STR attribute. A verb base ARG-STR includes E, the event semantic function role, ⟨E, ⟨x,...⟩⟩. That of the participle is enriched by addition of the A semantic function role, to become (simplifying somewhat) ⟨A_i⟨E, ⟨x_i,...⟩⟩⟩, where the co-indexing indicates that the element of which the participle is predicated is an element of the argument array of the base verb lexeme (the highest such argument for Indo-European type participles, though not necessarily so in other languages). This representation reflects the mixed categorial nature of the participle. While its main external syntax (**Haspelmath96**) is now that of an adjective, it retains the eventive ARG-STR of the base verb, which permits it, on a language specific basis, to realize a number of verb properties. These typically include the realization of internal arguments as verb dependents (complete with quirky case marking). The E semantic function role also permits modification by adverbials targeting event semantics.

Exactly which adjectival properties are acquired and, more crucially for participles, which verb properties are lost or retained, differs cross-linguistically. The Indo-European participle, for instance, typically retains the active∼passive alternation, but may also retain aspect (Slavic) or tense (**Lowe15:book**). As attributive modifiers, the Indo-European participles can only modify a noun which expresses the participle's highest (subject) argument, effectively making them into heads of subject-oriented relative clauses, but in other languages there is much greater freedom in the choice of argument that can be relativized on by the participle (**Spencer16:ptcprels**).

In §**??** 'Paradigm linkage rules for Russian' I present a more detailed analysis of Russian participles in which this basic schema for transpositions is expanded upon.

4 Russian participles

In this section we look at the set of four participles that are regularly associated with Russian verbs. Before we can consider these, however, we need to understand Russian verb inflection and the place of the participles in that system. I begin with an overview of the grammatical distinctions as a whole made by verbs, in other words, the CONTENT paradigm, before considering the actual morphological forms themselves. We encounter the familiar problem that there is no consensus on just what the oppositions are and how they relate to each other, and so some of my descriptive decisions will be motivated in part by expositional convenience. I illustrate with the second conjugation transitive verb UDAR′IT′ 'hit', whose imperfective aspect series is formed by shifting to the first conjugation, UDAR′AT′.

The simplest categories are the infinitive and imperfective ('present') and perfective ('past') gerunds. These are indeclinable. Telic verbs such as UDAR′IT′/UDAR′AT′, 'hit', require imperfective and perfective aspect forms for most of their paradigm. We can distinguish three moods, indicative, imperative, conditional. The imperative is straightforward and I will ignore it here. The indicative distinguishes three tenses, present (for imperfectives only), past, future. I return to the conditional below.

Transitive verbs show an active∼passive voice opposition. However, the voice system

is complicated by the fact that imperfective verbs are able to take the reflexive suffix *-s´a/s´*. This has the basic function of giving a reflexive/reciprocal meaning (though similar meanings can also be expressed with fully-fledged reflexive/reciprocal pronouns). The reflexive form also has a wide variety of other uses, including the passive (**Gerritsen90:book**). This means that we should set up a reflexive voice category and define the imperfective passive in terms of this, but I set this task aside since it is not directly relevant. The perfective verbs express the passive alternation periphrastically, with BE + perfective passive participle. At the level of the CONTENT paradigm we need to be able to define [VOICE:{act,pass}] for both perfective and imperfective verbs, therefore.

Present tense is expressed morphologically, but only for imperfective verbs. There is no dedicated tense marker, and in effect, present tense is realized by the person/number subject agreement morphology. The future tense is expressed periphrastically by imperfective verbs, by means of BE + the infinitive form. For perfective verbs the future is expressed by a paradigm which is essentially the same as the present tense paradigm for imperfective verbs. Thus, the present tense of the imperfective verb P´ISAT´ 'write' is *p´išu, p´išeš,...* and the future tense of the prefixed perfective verb form NAP´ISAT´ is *nap´išu, nap´išeš,....*

One of the main challenges of the Russian verb system is the representation of the past tense. This is derived historically from a periphrastic perfect tense series formed by BE and a resultative participle expressed by a suffix *-l*, the l-participle. Syntactically, the l-participle behaved like a predicative adjective, agreeing with the subject in number and gender, but not in person. The auxiliary BE was lost, leaving the l-participle and its adjective-like agreement as the sole exponent of past tense. The agreement inflections on an l-participle such as *(na)p´isal* 'wrote' are almost identical to those on a predicative (short-form) adjective such as MAL 'short': *(na)p´isal, (na)p´isala, (na)p´isalo, (na)pisal´i* vs *mal, mala, malo, maly*. The only real difference is that in the plural the l-participle stem is palatalized but not the predicative adjective stem: *(na)p´isal´i* vs *maly*.

The simple way of analysing the past tense construction would be to take the *-l* formative to be an exponent of past tense and define two distinct sets of subject agreement rules for present/future and past tenses. However, the l-participle has one other significant usage which precludes this direct analysis. The conditional mood is expressed by means of the invariable particle *by*, a kind of freely distributed enclitic (it can occur anywhere in the clause except absolute initial position). This can co-occur with the infinitive, (**??**).

(4) Esl´i by skazat´ pravdu, ... 'To be honest, ...'
if BY say.INF truth

However, it is much more often found with the l-participle, (**??**).

(5) Esl´i by ty skaza-l pravdu, ... 'If you told/were to tell/had told the truth, ...'
if BY you say-L-PTCP truth

As is indicated in the gloss, the conditional is tenseless, and serves as the translation equivalent of past or non-past conditional in English. The existence of the conditional

construction means that we cannot regard the l-participle simply as the exponent of past tense. In fact, it is an instance of what, since **Aronoff94:book** morphologists have called a 'morphome', that is, a pure morphological form, which serves as a stem for building up inflected word forms, but which realizes no MPSs on its own and whose distribution is not mandated by any semantic, phonological or other non-morphological properties. The CONTENT paradigm for Russian is summarized in Table **??**, but ignoring the true participles.

Table 1: Russian verb CONTENT paradigm for UDAR′IT′/UDAR′AT′ 'hit'

ASPECT	imperfective	perfective
INFINITIVE	udar′a-t′	udar′i-t′
GERUND	udar′a-ja	udar′i-v(ši)
IMPERATIVE	udar′a-j(te)!	udar′(te)!
TENSE		
present	udar′a-ju, -eš, …	<none>
future	bud-u, -eš, …udar′at′	udar′-u, -iš
past	udar′a-l, -a, -o, -′i	udar′i-l, -a, -o, -′i
CONDITIONAL	udar′a-l, -a, -o, -′i + by	udar′i-l, -a, -o, -′i + by
PASSIVE	udar′at′-s′a, etc	(byl) udaren, -a, -o, -y

I turn now to the morphological or FORM paradigm. We can divide Russian verb forms into five groups. The first is the infinitive and the second the set of two indeclinable gerund forms. The third is the set of finite forms, including the imperative mood. In practice, these are limited to the present (non-past) forms showing subject agreement in person/number. The fourth is the l-participle. Finally, we have the set of (declinable) active and passive participles discussed earlier. These are tabulated in Table **??**.

Table 2: Russian verb FORM paradigm for UDAR′IT′/UDAR′AT′ 'hit'

ASPECT	imperfective	perfective
INFINITIVE	udar′a-t′	udar′i-t′
GERUND	udar′a-ja	udar′i-v
IMPERATIVE	udar′a-j(te)!	udar′(te)!
TENSE		
present/future	udar′a-ju, -eš, …	udar′-u, -iš
L-PTCP	udar′a-l, -a, -o, -′i	udar′i-l, -a, -o, -′i
REFLEXIVE	udar′at′-s′a, etc	<none>

Summarizing Table **??**, the verb can/must inflect for aspect. The verb system also shows voice alternations. However, these are effected either through reflexive morphol-

ogy (imperfective verbs) or periphrastically (perfective verbs) so voice proper lacks a dedicated, purely morphological exponent, except for the true participles. The infinitive, gerunds and the l-participle do not show any tense oppositions. The subject agreement shown by l-participles differs from that of finite verb forms in that it is defined in terms of MPSs proper to predicative adjectives, not those of finite verbs.

Table 3: CONTENT feature array

ASPECT	{ipfv,pfv}
VFORM	INFINITIVE
	TENSE:{prs, fut, pst}
	IMPERATIVE:{sg, pl}
	CONDITIONAL:{yes, no}
REFLEXIVE	{yes, no}
AGRSUBJ	PERSON:{1, 2, 3}
	NUMBER:{sg, pl}
	GENDER:{m, f, n}
VOICE	{ACTIVE, PASSIVE}
REPR	{⟨V,A⟩, ⟨V,Adv⟩}

In Tables **??**, **??** I provide a summary list of the features which populate the CONTENT and FORM paradigms. I have provided only basic labels for the various MPSs. Ideally, we would want to know how they are grouped together, if at all, in the two paradigms. This is a difficult question, and I finesse it by just assuming what is effectively a list structure for the MPSs, with a number of dependency statements between them. Thus, I have not distinguished, say, a finite from a non-finite set of forms or constructions. However, for the purposes of giving a broad-brush characterization of the morphology∼syntax mapping this is probably not a problem.

Table 4: FORM feature array

ASPECT	{ipfv,pfv}
VFORM	INFINITIVE
	TENSE:{prs-fut}
	IMPERATIVE:{sg, pl}
	L-PTCP
REFLEXIVE	{yes, no}
AGRSUBJ	PERSON:{1, 2, 3}
	NUMBER:{sg, pl}
	GENDER:{m, f, n}
REPR	{⟨V,A⟩, ⟨V,Adv⟩}

From this overview we can see that there is a very clear divide between the CONTENT paradigm MPSs required to describe the system as a whole and the FORM paradigm MPSs required to describe the individual word forms. These tables ignore the inflectional paradigms of the participles, of course, but adding them will just serve to emphasize the CONTENT∼FORM disparity.

Russian has four participles, active∼passive and perfective∼imperfective, which are typical attributive modifiers with the agreement morphosyntax of standard adjectives. However, they retain a variety of verb properties, making them into typical examples of mixed categories. Thus, in (**??**), the imperfective active participle *upravl′ajuščij* takes a temporal PP adjunct and assigns instrumental case to its complement, just like the finite verb (**??**).

(6) a. (čelovek-a), upravl′a-jušč-ego v tečen′ie mnogo let
(the.person[M]-GEN.SG) run-PRSPTCP-M.GEN.SG in course many of.years
mestnoj školoj
local.INSTR school.INSTR
'of (the person) (who has been) running the local school for many years'

b. Ivanov upravl′aet v tečen′ie mnogo let mestnoj školoj
Ivanov runs in course many of.years local.INSTR school.INSTR
'Ivanov has been running the local primary school for many years'

The participles are often described as present/past tense forms, but their semantics is essentially aspectual and they fit somewhat better into the overall verb scheme if they are treated as perfective∼imperfective pairs. They are summarized in Table **??** (**Wade92:book**).

Table 5: Russian participles of the verb UPRAV′IT′/UPRAVL′AT′ 'control'

Aspect	imperfective	perfective
Active	*upravl′aju-šč-*	*uprav′i-vš-*
Passive	*upravl′a-em-*	*upravl′-on(n)-*

Perfective aspect participles, for semantic reasons, usually only have past time reference.[8] The imperfective participles realize a time relative to the main verb of the clause, so that 'present tense' is a particularly misleading label for these forms (Wade 1992: 375).

(7) Ja v′idel/v′ižu sobak-u, bega-jušč-uju po beregu
I saw/see dog[F]-ACC.SG run-ACTPRSPTCP-F.ACC.SG along the shore
'I saw/see the dog running along the shore'

[8] But see the counter-examples in the Academy of Sciences grammar *Russkaja Grammatika, I*: 667, *pred'javl′ajuščij* 'presenting', *vzvolnujuščij* 'exciting', *sdelajuščij* 'doing', *smoguščij* 'being able'.

The participle differs from the finite form in this respect. Example (**??**) would only be possible with either the meaning '(dog) which usually runs along the shore' or as a somewhat marked form of the historic present (cf *Russkaja Grammatika I*: 665).

(8) Ja v′idel sobak-u, kotor-aja begaet po beregu
I saw dog[F]-ACC.SG which-F.NOM.SG runs along the shore
'I saw the dog which is running along the shore'

In this respect, Russian participles are just like their English counterparts, of course.

The participles have a number of properties aligning them with verbs. In addition to realizing the purely verbal (eventive) categories of tense-aspect-voice, the active participles can take reflexive forms, either as reflexive variants of non-reflexives, inheriting all the semantics of the reflexive forms, *upravl′at′(s′a)* 'manage, control' ∼ *upravl′ajuščijs′a*, or as inherent reflexives (with no non-reflexive counterpart), *bojat′s′a 'fear'* ∼ *bojaščijs′a*. As we have seen, syntactically, the participles retain the verb's argument structure, including quirky case assignment to complements, such as instrumental in the case of *upravl′ajuščijs′a* (see examples (**??**)) and genitive in the case of *bojaščijs′a*.

On the other hand, the participles have a number of adjectival properties. The most salient morphosyntactic property is that of attributive adjective agreement together with the morphological property of belonging to a well-defined adjectival inflectional class. (As attributive modifiers the participles can be restrictive or non-restrictive, like other attributive modifiers, including relative clauses.) However, they can also be used as predicates with the copula BYT′ 'be' or with 'semi-copulas' such as STAT′ 'become', KAZAT′S′A 'seem', OSTAT′S′A 'remain', and others. Most commonly it is the perfective passive participles that can be used as predicates but active participles can also be found in this role (**RussGramm80-II**). The passive (though not the active) participles can also appear as predicates, in the so-called short form, a typically adjectival property.[9]

(9) užin uže poda-n
supper.M.SG already serve-PASSPTCP.M.SG
'Supper has already been served'

Given this brief descriptive summary of the basic facts of Russian participles we can turn to the central questions: how do we represent participles in a formal, constraints-based grammar with an inferential-realizational morphology? How do these representations relate to the inflection∼derivation divide and the issue of split morphology?

5 Paradigm linkage rules for Russian

In the extension to GPFM presented here, the FORM paradigm and the CONTENT paradigm are modelled by the attributes f-MORSIG, c-MORSIG. However, those types of

[9] Present active participles can be used in the short form, however, when they are converted into true, qualitative, adjectives (**RussGramm80-I**).

lexical relatedness which modify the MPSs of a representation, such as transpositions, will ipso facto modify the content of Stump's FORM/ CONTENT paradigms and the f-/c-MORSIG attributes. The GPF for such types of relatedness therefore has to specify that novel content, by stipulation, if necessary. In the rules I propose below I show how this can be achieved for Russian conjugation. The reference to MORSIG is taken to mean c-MORSIG by default and by default, this is identical to f-MORSIG.

I shall begin by specifying the CONTENT and FORM MORSIG attribute for non-participial verb categories. We need to define the f-MORSIG in part in terms of the c-MORSIG and in part independently. The default is the identity mapping from c-MORSIG to f-MORSIG. The c-MORSIG list shown earlier in Table **??** is defined for any lexical representation whose ARG-STR includes the E semantic function role. The MPSs that are shared across the CONTENT and FORM paradigms of Russian verbs are fairly limited (cf Tables **??**, **??**). They are (ignoring for the present the participles and gerunds): ASPECT:{ipfv, pfv}, VFORM:{INFINITIVE, IMPERATIVE:{sg, pl}}, VOICE:{act, pass}, and AGRSBJ:{PERSON/NUMBER/GENDER}.

The c-features ASPECT, VFORM:{INFINITIVE, IMPERATIVE:{sg, pl}} have relatively straightforward f-feature correspondents. I shall ignore INFINITIVE and IMPERATIVE for present purposes. AGRSUBJ is also a FORM property but with some complications and I return to it when I discuss the l-participle.

The status of the TENSE feature is a little unclear. At the CONTENT level there are clearly three values, *prs, fut, pst*, but only the *prs* and *fut* values have a FORM correspondent, and even then the value of the FORM correspondent is a composite *prs/fut*, and therefore not a direct correspondent of either c-[TENSE:prs] or c-[TENSE:fut]. The c-TENSE:{pst} property is realized by the morphomic l-participle, and not by any dedicated f-TENSE:{pst} property. I shall therefore assume a univalent FORM property, TNS, itself a value of VFORM, realizing c-TENSE:{prs, fut} depending on the value of ASPECT. This replaces the atom-valued TENSE:{present-future} shown in Table **??** above. The property VOICE:{act, pass} is intriguing. It is a CONTENT property of the verb system but it is not a FORM property of any part of the verb system proper, outside of the participle subsystem. However, the participles distinguish active and passive sets, and the perfective passive participle is actually part of the periphrastic exponence of the syntactic (CONTENT) VOICE property. Moreover, the two passive participles have the passive SYN|ARG-STR representation, namely, …⟨E⟨(x), y, …⟩, where (x) denotes the demoted active subject argument role. Therefore, VOICE is both part of the CONTENT and the FORM paradigm, albeit with a somewhat complex realization, which will require the f-MORSIG attribute to be modified by a feature co-occurrence statement restricting f-VOICE to the participles. The other CONTENT paradigm features are represented by forms which are effectively periphrastic. It is for these reasons that the FORM MPSs have to be defined independently, as shown below.

One attribute that is shared across FORM/CONTENT paradigms is REPRESENTATION. The only reason for a language to define a true transpositional category is to allow a lexeme to assume the syntactic distribution of a word of a different class, so REPR clearly has to be a CONTENT feature. In languages such as Russian, in which

participles are marked morphologically, this also means that REPR is a FORM feature. I shall assume that the REPR attribute as applied to verbs has three values. The first is *plain* (or ⟨V,V⟩), the 'identity representation', and the default value. Where no indication of REPR is given the default is to be understood. The second value of REPR is ⟨V, A⟩, defining the four participles. The third value is ⟨V,Adv⟩,[10] defining the imperfective and perfective gerunds. I shall ignore the gerunds for simplicity of exposition. This means that I shall only mark the participles explicitly.

We now define the content of the c-MORSIG attribute by reference to the verb's SYN value. The aspects, and voice are defined with the mapping shown informally in (??). This will apply to any representation whose ARG-STR contains the E sf role. In practice, this means verbs, participles, and gerunds.

(10) ... E ... ⇒ ASPECT, VOICE ⊂ MORSIG

I return to the c-MORSIG|CONDITIONAL, TENSE:{pst} mappings below when I discuss the status of the l-participle.

The feature array defining participles is given in Fig. ??. This is the 'derived' MORSIG attribute for any lexical representation defined by the feature REPR:⟨V,A⟩. The property sets labelled [1] in Fig. ?? come from the MORSIG attribute of the base verb by virtue of (??). The property sets [3], [4] are those which ensure that the participles are inflected like adjectives. I return to these when I have introduced adjective inflection.

$$\left[\begin{array}{ll} \text{MORSIG} & \left[\begin{array}{ll} \text{ASPECT} & \{\text{ipfv, pfv}\}\,\boxed{1} \\ \text{VOICE} & \{\text{act, pass}\}\,\boxed{2} \\ \text{CONCORD} & \left[\begin{array}{ll} \text{NUMBER} & \{\text{sg, pl}\} \\ \text{GENDER} & \{\text{m, f, n}\} \\ \text{CASE} & \{\text{nom, ...}\} \end{array}\right]\boxed{3} \\ \text{MORCLASS} & \text{adj}\,\boxed{4} \end{array}\right] \end{array}\right]$$

Figure 1: Feature structure for Russian participles

The key to my analysis of transpositions is to incorporate part of the analysis of derivational morphology into the definition of the transposition's entry. Given a feature pairing ⟨£,ρ⟩, where £ is a lexemic index and ρ contains a value of REPR (for instance, [REPR:⟨V,A⟩] for participles), the GPF applied to this pairing leaves the LI and the SEM representations of the base unchanged, but enriches the SYN|ARG-STR attribute by creating a complex semantic function role ⟨A_i⟨E⟨x_i, ...⟩⟩⟩. The coindexation guarantees

[10] The semantic function role label Adv stands proxy for whatever the appropriate way is of defining adverbs as distinct from adjectives.

that the noun head modified by the participle is identified with the highest thematic argument of the base verb's argument array, i.e. its SUBJECT. The FORM function component of the GPF defines the stem set for the participle, but underspecifies all other FORM information, including the MORSIG (and the CONTENT paradigm MORSIG also underspecified by the SYN function).

The lexemic entry for a typical transitive Russian verb such as UDAR´IT´/UDAR´AT´ is that shown in Fig. **??**, where $\mathcal{V}$ stands for the verb's lexemic index. The lexemic entry's value for LI is just the LI of the lexeme, of course (the GPF here does not describe a process of derivational morphology). Note that for this lexeme the SYN attribute, too, is completely underspecified, lacking even the ARG-STR attribute. The value of that attribute is determined by default from the *Event* ontological type of the SEM representation. However, the REPR feature which defines transpositions introduces a realization rule which has to be defined over a specified ARG-STR representation. Therefore, the ARG-STR attribute has to be part of the lexeme's (SYN attribute's) morpholexical signature (inflection-like argument structure alternations such as passive or antipassive impose a similar requirement). In this respect, the transpositions are like inflection and not like derivation.

We can informally state the realization rule which defines adjectival representations of lexemes (transpositions-to-adjective) as a function α from ARG-STR representations to ARG-STR representations, as shown in (**??**), where (**??**) defines a participle's ARG-STR and (**??**) defines that of a relational adjective.

(11) a. $\alpha(\langle \mathrm{E}\langle \mathrm{x},\ldots\rangle\rangle) = \langle \mathrm{A_i}\langle \mathrm{E}\langle \mathrm{x_i},\ldots\rangle\rangle\rangle$
b. $\alpha(\langle \mathrm{R}\rangle) = \langle \mathrm{A}\ \langle \mathrm{R}\rangle\rangle$

Given the lexemic entry in Fig. **??** and the realization rules for Russian morphology, the GPF for the imperfective active participle, *udar´ajušč-* will map to a partially underspecified lexical representation, as shown in Fig. **??**. A representation such as this is the 'quasi-lexeme' discussed earlier. Like a simple lexemic entry, or an entry defined by a derivational GPF, it needs to have its MORSIG attributes specified in order to be inflectable. I turn now to how those MORSIG entries are defined.

The partially specified MORSIG we need to be able to define for participles is that shown in Fig. **??**. The ASPECT/VOICE properties are shared with verbs. This can be achieved by writing the rules defining the MORSIG of verbs and participles in such as way as to refer either to the 'outermost' E semantic function role of the ARG-STR attribute, or the 'embedded' E role found with participles.[11] The CONCORD attribute, $\boxed{3}$, comes from the generic c-MORSIG of an adjective, shown in (**??**).[12]

(12) ... A ... $\Rightarrow$ [CONCORD:{NUMBER, GENDER, CASE}] $\subset$ MORSIG

The [MORCLASS adj] specification, $\boxed{4}$, is strictly speaking a stipulation, except that

[11] Russian action nominals are transpositional lexemes and not true transpositions.

[12] Members of the semantically defined class of quality or scalar adjectives will also have the feature COMPARISON added to their MORSIG to define comparative/superlative morphology.

$$
\begin{aligned}
f_{form}\,(\langle \mathcal{V},\, u\,\rangle) \quad &= \quad \left[\begin{array}{ll} \text{STEM0} & \left[\begin{array}{ll} \text{PHON} & \text{udar}' \\ \text{MORCLASS} & \text{V}\,|\,\text{CONJ2} \end{array}\right] \\ \text{STEM1} & \left[\begin{array}{ll} \text{PHON} & \text{udar}'\text{aj} \\ \text{MORCLASS} & \text{V}\,|\,\text{CONJ1} \end{array}\right] \end{array}\right] \\
f_{syn}(\langle \mathcal{V},\, u\,\rangle) \quad &= \quad u \\
f_{sem}(\langle \mathcal{V},\, u\,\rangle) \quad &= \quad \lambda\text{x,y[HIT(x,y)]} \\
f_{li}(\langle \mathcal{V},\, u\,\rangle) \quad &= \quad u
\end{aligned}
$$

Figure 2: Lexemic entry for UDAR′IT′/UDAR′AT′

$$
\begin{aligned}
\text{FORM}(\langle \mathcal{V}, \{\text{REPR:}\langle \text{V,A}\rangle\}\rangle) \quad &= \quad \left[\begin{array}{ll} \text{STEM0}\,|\,\text{PHON} & \text{udar}'\text{ajušč} \\ \text{MORSIG} & u \end{array}\right] \\
\text{SYN}(\langle \mathcal{V}, \{\text{REPR:}\langle \text{V,A}\rangle\}\rangle) \quad &= \quad \left[\begin{array}{ll} \text{ARG-STR} & \langle \text{A}_i \langle \text{E}\langle \text{x}_i\text{,y}\rangle\rangle\rangle \\ \text{MORSIG} & u \end{array}\right] \\
\text{SEM}(\langle \mathcal{V}, \{\text{REPR:}\langle \text{V,A}\rangle\}\rangle) \quad &= \quad \text{identity function} \\
\text{LI}(\langle \mathcal{V}, \{\text{REPR:}\langle \text{V,A}\rangle\}\rangle) \quad &= \quad \text{identity function}
\end{aligned}
$$

Figure 3: Underspecified entry for *udar′ajušč(ij)*

$$
\left[\begin{array}{ll} \text{MORSIG} & \left[\begin{array}{l} \text{ASPECT:}\{\text{u}\} \\ \text{VOICE:}\{\text{u}\} \\ \text{CONCORD:}\{\text{u}\} \\ \text{MORCLASS: ADJ}\,|\,\text{DECL1/2} \end{array}\right] \end{array}\right]
$$

Figure 4: MORSIG for Russian participles

in Russian (in contrast to, say, Latin) all participles belong to the default adjectival class, DECL1/2.

Given this machinery we can now account for the transpositional mixed category of participle by application of the (quasi-inflectional) generalized paradigm function applying to a verb and delivering its participial forms, triggered by the [REPR] feature. These representations will be underspecified for (adjectival) CONCORD features. For concreteness, consider the imperfective active participle, *udarʹajušč-*. Given ρ = {REPR:⟨V,A⟩, ASPECT:ipfv[1], VOICE:act[2]}, then for $\mathcal{V}$ = UDARʹITʹ/UDARʹATʹ, the GPF(⟨$\mathcal{V}$,ρ⟩) will apply to a lexical representation which is derived from the lexemic entry for UDARʹITʹ/UDARʹATʹ whose MORSIG attributes have been fully specified, allowing the lexeme to be inflected (in the broad sense of this term, including 'inflection' for participle formation). This GPF will deliver a partially underspecified lexical representation for the participle. The GPF will specify the participle's stem form(s), the ASPECT, VOICE features which define that particular participle, and the enriched ARG-STR attribute with complex semantic function role.

The lexical representation which is input to the GPF is shown in Fig. **??** and the lexical representation of the participle is shown in (**??**). In Fig. **??**, STEM0 denotes the perfective stem, which is effectively the lexeme's root. CONJ2 is second conjugation, and this means that most inflectional rules will be defined over another stem, *udarʹi-*, derived by regular rules of stem formation. STEM1 denotes the imperfective stem, a member of CONJ1, the first conjugation, whose inflectional stem is therefore *udarʹaj-*. Attributes which belong to both FORM and SYN (CONTENT) paradigms are tagged to make them more easily identifiable.

(13) f_{form} = STEM-iap ⊕ šč = udarʹaju-šč-
f_{syn} = ARG-STR:⟨A_i⟨E⟨x_i,y⟩⟩⟩

where STEM-iap denotes the imperfective active participle stem, derived from STEM1.

We have now achieved our goal of defining participles. In effect, we have defined the lexemic entry for a class of adjectives, whose peculiarity is that they are marked for verbal voice and aspect and they share their semantics and lexemic index with a parent verb.

It is instructive to compare the behaviour of the true participles with that of the l-participle. While the true participles are mixed categories, the l-participle is essentially a verb form with unusual agreement morphology. It is not entirely clear how best to account for the peculiar agreement properties of the l-participle in the past tense and conditional constructions, but the simplest (if somewhat crude) way to do this is to assume that verbs in general agree with their subjects in person, number, gender features, but that in the past/conditional forms agreement is restricted to number, gender, while in the present tense forms it is restricted to person, number.[13]

[13] A more sophisticated approach might define agreement in a morphology-driven fashion by stating that the agreement features that the syntax can manipulate are restricted to those that can be expressed by a particular morphological form, so that it is the morphological MORSIG that determines which features trigger agreement, even in identical syntactic positions.

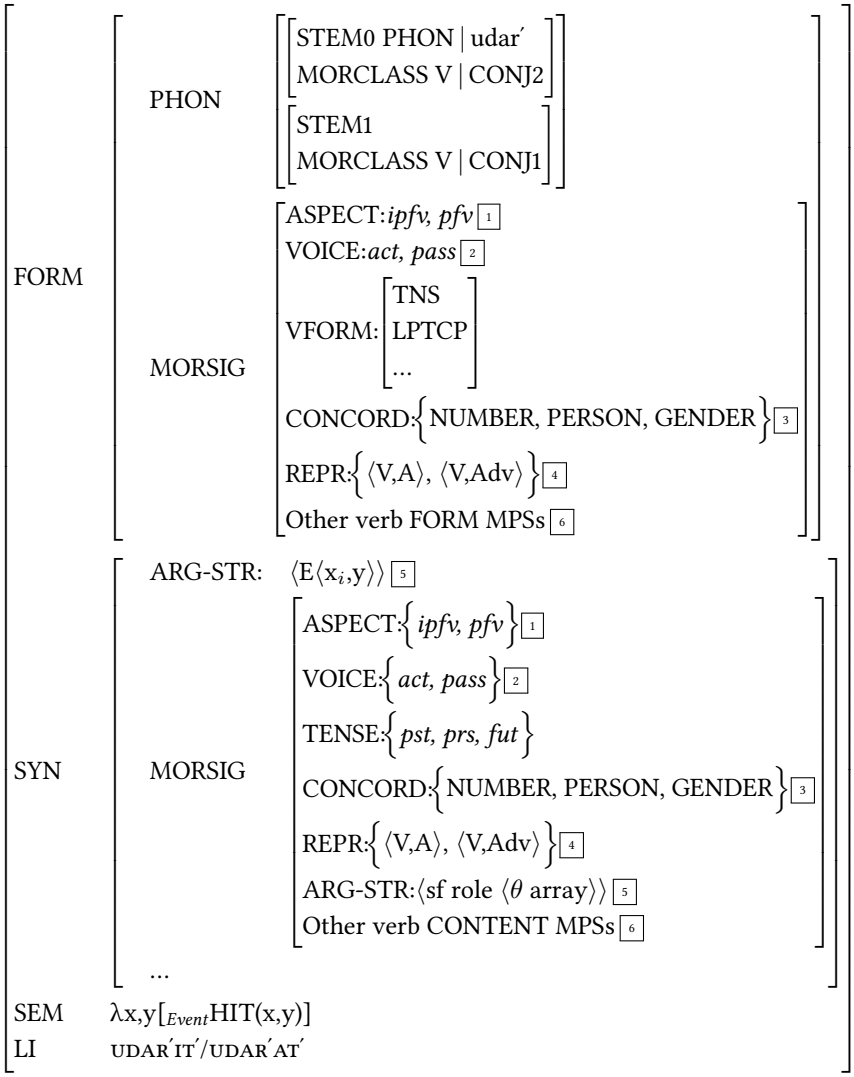
FORM
PHON
STEM0 PHON | udar´
MORCLASS V | CONJ2
STEM1
MORCLASS V | CONJ1
MORSIG
ASPECT: ipfv, pfv 1
VOICE: act, pass 2
VFORM: TNS LPTCP ...
CONCORD: NUMBER, PERSON, GENDER 3
REPR: ⟨V,A⟩, ⟨V,Adv⟩ 4
Other verb FORM MPSs 6
SYN
ARG-STR: ⟨E⟨x_i,y⟩⟩ 5
MORSIG
ASPECT: ipfv, pfv 1
VOICE: act, pass 2
TENSE: pst, prs, fut
CONCORD: NUMBER, PERSON, GENDER 3
REPR: ⟨V,A⟩, ⟨V,Adv⟩ 4
ARG-STR: ⟨sf role ⟨θ array⟩⟩ 5
Other verb CONTENT MPSs 6
...
SEM λx,y[Event HIT(x,y)]
LI UDAR´IT´/UDAR´AT´

Figure 5

I conclude, then, that the l-participle is a verb form that inflects just like a (subtype of) predicative (short-form) adjective. Assuming a feature adjdecl covering adjectival morphology generally, we can distinguish several sub-types of declension, including that for the predicative adjectives, [adjdecl:predadj]. The l-participles will belong to a subtype of this class, [adjdecl:predadj:lptcpdecl]. The adjdecl feature is part of the MORSIG attribute of ordinary adjectives, as defined by the paradigm linkage rule in (**??**).[14]

(14) ... A ... ⇒ ADJDECL ⊂ MORSIG

We also compare true participles with qualitative adjectives derived by conversion from participles. These are a type of transpositional lexeme. The theoretical significance of this type of lexical relatedness for current morphological models was first identified, as far as I am aware in Spencer (2013: 275), where I discuss English words such as PREPOSITIONAL, from PREPOSITION. These look like relational adjectives (noun-to-adjective transpositions), because their lexical semantic content seems to be identical to that of their base noun (*prepositional phrase* means the same as the compound *preposition phrase*, for instance). However, the English adjectives are syntactically opaque: *monosyllabic prepositional phrase* does not mean 'phrase headed by a monosyllabic preposition' but only 'monosyllabic phrase headed by a preposition' (though that interpretation is possible for the compound, *monosyllabic preposition phrase*). In other words, the adjectival expression only has the structure *monosyllabic [prepositional phrase]*, not *[monosyllabic preposition]al phrase*.

English has a large number of qualitative adjectives derived by conversion from participles, which are also instances of transpositional lexemes (**Spencer16**:**MorphMetatheory**): *amazed/amazing, bored/boring, challenged/challenging, interested/interest-ing, ...* In very many cases it is not possible to identify a meaning difference between the adjective and the etymological verb base: *This book bores me/This book is boring*. Russian participles likewise are often converted into qualitative adjectives: *potr´asajuščij* 'amazing', *vyzyvajuščij* 'provocative, defiant, challenging'. In Russian it is often possible to determine that a word with the shape of a participle is actually an independent adjective, since only true adjectives have the short predicative form: *uspexi potr´asajušči* (plural, from *potr´asajušč*) 'the-progress is-amazing' (lit. 'the-successes are-amazing'), *ego poveden´ie vyzyvajušč* 'his behaviour is-defiant'. True participles do not have the predicative form (Russkaja Grammatika II, p.666). English converted participles fail to inherit the complementation properties of the base verb: *The obstacle course challenged the stamina of the athletes*∼*The obstacle course was very challenging (*the stamina of the athletes)*. Russian transpositional adjective lexemes behave likewise. The transitive base verb POTR´ASAT´ 'to amaze' gives us *uspexi potr´asal´i nas* 'the-successes amazed us', but the (active) transpositional lexeme cannot take a direct object: **uspexi potr´asajušči nas.* The true participle remains transitive: *potr´asajuščie nas uspexi* 'progress (successes) which amaze us'.

[14] As adjectives, we might expect the participles to have predicative forms, too. This is true, however, only for the passive participles, especially the perfective passive, which has a special stem form ending in a singleton /n/, *nap´isan* 'written', in contrast to the attributive form with geminated /nn/: *(v speške) nap´isannaja (zap´iska)* '(a hastily) written (note)'.

We thus have a double dissociation of properties: on the one hand, we have the l-participle which has the form of a predicative adjective but which realizes finite (tense/mood) verb properties and retains the full complementation properties of the verb, and on the other hand we have participle-like adjectives which, while allowing predicative adjective forms, lack all the crucial morphosyntax of verbs. In between we have the true participles, with the external morphosyntax of an adjective but the complementation properties of the base verb.

6 Conclusions: Transpositions and split morphology

I have argued in this paper for a view of morphosyntax which recognizes a word/phrase distinction and, given that, a distinction between (abstract) lexemes and (concrete) inflected word forms of those lexemes, valid for very nearly all known languages. I have also assumed that languages can increase their stock of lexemes by means of derivational morphology, and that in some cases this is sufficiently regular to be regarded as paradigmatic, hence, part of the grammatical system proper. The inflection/derivation distinction is controversial, however, because it is often difficult to know where the boundary actually lies and where to place intermediate types of lexical relatedness. The transpositions, as exemplified by the Russian participial system discussed here, represent a particularly troublesome case-in-point.

Participles and other transpositions are often treated as a type of derivational morphology, because they involve a shift in word class, but this is a wrong characterization. Participles are part of a verb's paradigm, they are not a type of lexical stock expansion (**Beard95:book**). Nonetheless, participles inflect like adjectives, not like verbs and thus seem to straddle the inflection/derivation divide in a way that calls that very distinction into question, and, on the face of it, even provides support for models in which all morphology is just syntax by other means (Minimalism/Distributed Morphology) or in which all syntax is just morphology by other means (American morphemics). Here I have claimed that, on the contrary, we can only make sense of participles against a set of background assumptions that contradict monolithic models in which there is no autonomous morphology module (Aronoff's 'morphology-by-itself'), so that morphology is no more than syntax by other means. The crucial observation is that derivational morphology induces a type of lexical opacity which is lacking in transpositions, which, by contrast, show the kind of lexical transparency associated with inflected forms.

In the GPFM model, canonical derivation is a relation between maximally underspecified (minimally redundant) lexical representations (lexemic entries), consisting of a specification of the basic root form (FORM|PHON) and the ontological/semantic representation (SEM). The morphosyntactic properties are then projected from these by default mappings. Canonically, derivation enriches the PHON and the SEM representations, and the Default Cascade then specifies the morphosyntactic properties of the derived lexeme. One consequence is that there will then be no 'trace' left of the morphosyntactic properties of the base lexeme, such as its word class or its argument structure. This automatically guarantees most of the lexical opacity/lexical integrity effects familiar from

the literature. In some (noncanonical) cases it may be necessary to stipluate overrides of lexical information as part of the derivational Generalized Paradigm Function. This might be true if, for instance, a base lexeme belongs to a lexical category which is not the default for its ontological class. For instance, a language with a distinct lexical category of (qualitative) adjective may also have non-derived stative verbs whose denotations are of the ontological type Property, and which by the Default Cascade should be adjectives, not verbs. Such a verb would have to have its ARG-STR|⟨E …⟩ value prespecified. If such a lexeme were the input to a verb-to-noun derivational function (nominalization), then that ARG-STR would have to be overridden by the nominalization function, replacing it wholesale with the ARG-STR|⟨R⟩ value. Exactly how such cases are to be handled has to remain a matter for future research.

No such opacity is found with canonical inflection: in general, the syntax treats a noun as a noun no matter what its number, case, definiteness, … value. Thus, although a locative case marked noun would normally be restricted to functioning as an adjunct or the complement of a class of adpositions, it can still be modified by an adjective, just like any other noun form, so that for a noun 'house', *new house-*LOC means 'at a new house'. In this respect, a locative case marked noun typically differs from a derived denominal lexeme denoting a location. Many languages have a denominal derivational marker meaning 'place where there is/are NOUN, place associated with NOUN': N-PLACE. Typically, the base noun, N, is not accessible to morphosyntax, so that an expression such as *new house-*PLACE could only mean 'new place where there is a house/are houses' and not 'place where there is/are a new house(s)'. Thus, inflection differs from derivation in being lexically transparent.

The significance of these rather obvious points about inflected forms is that we are far less able to make similarly categorical claims where transpositions are concerned. A participle behaves in the syntax *to some extent* as though it were an inflected verb form, but not entirely. In GPFM the lexical representation of a transposition exhibits transparency by virtue of being a member of the base lexeme's (extended) paradigm, that is, by bearing the same LI as the base. In this respect it is not an autonomous lexeme. On the other hand, the transposition exhibits the external syntax of a distinct (derived) lexical category. The extension to the GPF proposed here permits us to model this 'inflectional-paradigm-within-a-paradigm' effect in a way that reflects the lexical transparency of the transposition while also allowing us to state restrictions on full transparency as a restriction on the MORSIG of the transposition.

A crucial aspect of the analysis is the distinction between FORM/CONTENT properties (m-/s-features). Without at least this level of differentiation we cannot make sense of Russian participles and we cannot distinguish them from the verbal l-participle form or the transpositional lexemes. A second crucial aspect of the analysis of any type of transposition is the Lexemic Index (LI). In the general case, there is no combination of lexical properties that will uniquely serve to individuate lexemes, but it is nonetheless essential to impose such an individuation to account for the patterns of lexical relatedness that are found across languages, specifically, to distinguish true participles from departicipial converted adjectives (transpositional lexemes).

The combination of the FORM/CONTENT paradigm distinction (or its equivalent) and the Lexemic Index allow us to define not just canonical inflection but also non-canonical intra-lexemic types of relatedness such as that shown by transpositions. That combination also serves to reconstruct the split in the morphology argued for originally by Anderson. Indeed, split morphology is entailed by this set of assumptions, except that in a constraints-based model the split is defined in terms of access: inflectional morphology is that which permits syntax to retain access to the properties of the base lexeme (lexical transparency), while derivational morphology permits no such access (lexical opacity/integrity).

The more articulated view of morphosyntax proposed here allows us to pose a question which was not at the forefront of debate over the question of split morphology, as far as I am aware: which side of the split do transpositions fall on? The answer, given the foregoing, is 'both'. The transposition is inflectional by virtue of preserving the base's LI and by virtue of the, at least partial, transparency of the base's properties. It is derivational by virtue of the fact that it defines the paradigm of a quasi-lexeme, within the paradigm of the base. But it would be difficult to make sense of this conclusion without assuming the basic split in the first place.

1 APPENDIX: Illustration of the lexical representations assumed

Figure **??**: Illustration of the application of the GPF for standard inflection.

DRIVE = GPF(⟨DRIVE,u⟩) =

To specify, say, the 3sg form *drives*, the GPF(⟨DRIVE,{3sg}⟩) applies to the output of Figure **??** to specify the FORM value |STEM0⊕z|, leaving other aspects of the representation unchanged. This is equivalent to the operation of the paradigm function in PFM1 and the output of the ***Corr*** function in PFM2.

$$\begin{bmatrix} \text{FORM} & \text{STEM0} \ |\, \text{draɪv} \,| \\ \text{SYN} & — \\ \text{SEM} & [_{\textit{Event}}\ \lambda x,y.\textbf{drive}(x,y)] \\ \text{LI} & \textsc{drive} \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} \text{FORM} & \begin{bmatrix} \text{STEM0} & \ |\, \text{draɪv} \,| \\ \text{MORSIG} & \begin{bmatrix} \text{TNS} & \{\text{prs, pst}\} \\ \text{VFORM} & \{\text{ing-form, ed-form, base}\} \\ \text{AGR} & \{\text{3sg, none}\} \end{bmatrix} \end{bmatrix} \\ \text{SYN} & \begin{bmatrix} \text{SYNCAT} & \text{V} \\ \text{ARG-ST} & \langle \text{E}\langle x,y \rangle \rangle \\ \text{MORSIG} & \begin{bmatrix} \text{TNS} & \{\text{prs, pst}\} \\ \text{ASP} & \{\text{simple, prog, perf}\} \\ \text{AGR} & \{\text{3sg, none}\} \\ \ldots & \end{bmatrix} \end{bmatrix} \\ \text{SEM} & [_{\textit{Event}}\ \lambda x,y.\textbf{drive}(x,y)] \\ \text{LI} & \textsc{drive} \end{bmatrix}$$

Figure 6: Inflection in GPFM

Figure **??**: Illustration of the application of the GPF for standard derivation

DRIVE ⇒ DRIVER

Where δ = SubjNom, Σ = [$_{Event}$ λx,y.**drive**(x,y)], GPF(⟨DRIVE,δ⟩) =

$$\begin{bmatrix} \text{FORM} & \text{STEM0} \mid \text{draɪv} \mid \\ \text{SYN} & — \\ \text{SEM} & \Sigma \\ \text{LI} & \text{DRIVE} \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} \text{FORM} & [\text{STEM0} & \text{STEM0(DRIVE)}\oplus \mid \text{ə} \mid \\ \text{SYN} & — & \\ \text{SEM} & [_{Thing}\ \lambda x.\text{PERSON}(x) \wedge \Sigma & \\ \text{LI} & \delta(\text{DRIVE}) & \end{bmatrix}$$

Figure 7: Derivation in GPFM

The output of this GPF then undergoes specification of MORSIG by the Default Cascade, which defines the derived lexeme as a syntactic and morphological noun.

Abbreviations

AGRSUBJ	subject agreement
act	active
ARG-STR	argument structure
CONJ	conjugation
DECL	declension
GPF	Generalized Paradigm Function
GPFM	Generalized Paradigm Function Morphology
ipfv	imperfective
LI	lexemic index
L-PTCP, l-ptcp	l-participle
MORCAT	morpholexical category
MORCLASS	morphological class
MORSIG	morpholexical signature
MPS	morphosyntactic property set
pass	passive
PF	Paradigm Function
PFM	Paradigm Function Morphology
pfv	perfective
PHON	phonological form
REPR	REPRESENTATION
SEM	semantics
SYN	syntax

Acknowledgements

I am grateful to two anonymous referees and to Larry Horn for very helpful commentary and for correcting a number of typos and other errors. I would also like to express my gratitude to Steve Anderson for his pioneering work in getting morphology accepted as a legitimate subject for generative grammar, and also to thank him for his enlightening and entertaining contributions to the field, admirably limpid, often witty, and always thought-provoking.

Chapter A

Rules and blocks

Gregory Stump
University of Kentucky

In a series of publications, Stephen Anderson developed the idea that the definition of a language's inflectional morphology involves blocks of realization rules such that (i) realization rules' order of application follows from the ordering of the blocks to which they belong and (ii) realization rules belonging to the same block stand in a relation of paradigmatic opposition. A question that naturally arises from this conception of rule interaction is whether it is possible for the same rule to figure in the application of more than one block. I discuss two systems of verb inflection exploiting exactly this possibility – those of Limbu and Southern Sotho. In order to account for the special properties of such systems, I argue that in the definition of a language's inflectional morphology, one rule may be dependent upon another, and that in such cases, the dependent rule may figure in the application of more than one block precisely because the "carrier" rules on which it is dependent differ in their block membership. In formal terms, this means that the definition of a language's inflectional morphology may draw upon principles of rule conflation by which a dependent realization rule combines with its carrier rule to form a single, more complex rule, typically occupying the same block as the carrier rule. I further show that there is considerable independent motivation for the postulation of these principles.

1 Introduction

In a series[1] of articles culminating in his 1992 monograph *A-morphous Morphology*, Stephen Anderson developed a model for the precise inferential-realizational definition of complex inflectional patterns.[2]

Two central principles of this model are (1) and (2). According to (1), the definition of the Latin word form *laudā-ba-nt-ur* 'they were being praised' involves the realization of a morphosyntactic property set through the interaction of ordered rule blocks. One of these houses a rule realizing the imperfect indicative through the suffixation of *-bā*; this is followed by a block housing a rule realizing third-person plural subject agreement

[1] Key references include **Anderson1977a**; **Anderson1982**; **Anderson1984**; **Anderson1984b**; **Anderson1986**

[2] In the typology of morphological theories proposed by **Stump2001** a theory is inferential if it employs rules to infer the form of a language's words and stems from that of less complex stems; a theory is realizational if its definition of a language's morphology takes a word's content as logically antecedent to its form.

by the suffixation of *-nt*; this, in turn, is followed by a block containing a rule realizing passive voice through the suffixation of *-ur*. The successive application of these rule blocks infers the word form *laudābantur* from the stem *laudā-* as the realization of the property set {3 pl imperf ind pass} in the paradigm of the lexeme LAUDĀRE 'praise'.

(1) A language's inflectional rules are organized into ordered blocks such that two rules' order of application depends on the ordering of the blocks to which they belong.

(2) Rules belonging to the same block are disjunctive: at most one rule per block applies in the realization of a given word form. In general, competition between rules is resolved in favor of the rule with the narrower domain of application.

According to (2), the definition of a word form by a sequence of rule blocks involves the application of at most one rule per block. Two rules belonging to the same block may be defined so as to apply in disjoint contexts; for instance, the rule realizing third-person plural subject agreement through the suffixation of *-nt* realizes different property sets from the rule realizing first-person plural subject agreement throught the suffixation of *-mus*. But it can also happen that two rules belonging to the same block are both in principle applicable in the same context; for instance, *-ī* and *-ō* both realize first-person singlar subject agreement and might therefore be seen as entering into competition in the realization of certain forms. Given that the *-ī* rule applies only in the first-person singular perfect indicative active (e.g. *laudāvī* 'I have praised'), its domain of application is narrower than that of the *-ō* rule, which apparently applies as a default, surfacing in the present and future indicative active (*laudō* 'I praise', *laudābō* 'I will praise') and passive (*laudor* 'I am praised', *laudābor* 'I will be praised') as well as in the future perfect indicative active (*laudāverō* 'I will have praised'); accordingly, the *-ī* rule overrides the *-ō* rule in the realization of the first person singular perfect indicative active.

Anderson's model has afforded the most plausible existing accounts of a diverse range of inflectional systems (see, for example, the analyses of Potawatomi and Georgian in **Anderson1977a** 1984a, 1986 and that of German in **Zwicky1985a**), and it continues to raise important theoretical questions. One such question is whether the same rule[3] may figure in the application of more than one rule block. I argue here that in a particular class of cases, this is precisely what happens.

In the cases in question, there is always a relation of dependency among particular rules. **Harris2017** describes relations of this sort in affixal terms as involving a dependent affix that only appears in the presence of an available carrier affix. Adopting and extending her terminology, I describe such relations as involving a dependent rule that only applies in combination with an available carrier rule. As I show, a dependent rule may figure in the application of more than one block if the rules on which it is dependent differ in their block membership. Instances of this sort are of two kinds.

First, there are instances of multiple exponence in which the same rule of affixation apparently applies in more than one block in a word form's inflectional realization; an

[3] Two rule applications are seen as involving the "same rule" if they realize the same morphosyntactic content by means of the same exponent even if they introduce that exponent into different positions.

example from the Limbu language [Kiranti; Nepal] is the multiple exponence of certain agent concord properties in the inflection of transitive verbs. Second, there are instances of polyfunctionality involving a rule of affixation whose function varies systematically according to the block in which it applies; an example is the polyfunctionality of concordial affixes in the verbal inflection of the Bantu languages. In order to account for cases of these two sorts, it is desirable to supplement (1) and (2) with principles (3) and (4).

(3) A dependent rule may be conflated with a carrier rule to produce a more complex rule. Where a dependent rule R_1 realizes property set σ by means of exponent x and its carrier rule R_2 realizes property set τ by means of exponent y, the conflation of R_1 with R_2 is intuitively a rule realizing the property set $\sigma \cup \tau$ by means of the combined exponents x and y.

(4) A rule block may contain both simple and conflated rules.

As I shall show, these principles afford economical accounts of multiple exponence in Limbu verbs and of polyfunctional verbal concord markers in Southern Sotho [Bantu; Lesotho]. After describing the expression of agent properties in Limbu verbs (§2) and the Southern Sotho system of verbal concord (§3), I propose a formal framework for rule conflation in inflectional morphology (§4). I then present explicit theories of the observed pattern of multiple exponence in Limbu (§5) and that of polyfunctional concordial morphology in Southern Sotho (§6). I conclude with some observations about the wider importance of rule conflation in an adequate theory of morphology (§7).

2 Multiple exponence in the expression of agent inflection in Limbu verbs

In Limbu, two agent concord suffixes participate in relations of multiple exponence: *-ŋ*, an expression of first-person singular agent concord, and *-m*, an expression of nonthird-person plural agent concord. Table 1 exemplifies the distribution of these suffixes in positive forms of HUʔMAʔ 'teach'.[4] (In this table, parenthesized segments are superficially elided in prevocalic position by an ordinary phonological process.) Both suffixes appear in two different affix positions; **Driem1987** labels these positions **sf5** and **sf9**.[5]

[4] The structure of Table 1 should be carefully noted. Each row in the table is occupied by a different word form in the paradigm of the Limbu verb HUʔMAʔ 'teach'. Each word is in exploded form, with its parts arranged in columns corresponding to the affix position classes postulated by van Driem. (I follow him in labeling these classes **pf1** and **sf1**–**sf10**.) Thus, the word form in the 1s → 3ns row of the nonpreterite part of the table is *huʔr-u-ŋ-si-ŋ* 'I teach them'. This table does not comprise the complete paradigm of HUʔMAʔ 'teach', but encompasses those forms that involve the agent suffixes *-ŋ* and *-m* (as well as a few other pertinent forms in which the appearance of these suffixes is overridden). The claim that these suffixes appear in two different positions means that they appear in two different columns, since each column defines an affix position class in the traditional sense.

[5] Affix positions **sf3** and **sf6** are missing from Table 1 because the affixes that appear in these positions don't occur in forms having a first-person singular agent or a nonthird-person plural agent.

Table 1: The agent suffixes -ŋ and -m in positive forms of the Limbu verb HUʔ-MAʔ 'teach'

	agent	**pf1**			**sf**							
	→ patient	a	b	stem[1]	**1**	**2**	**4**	**5**	**7**	**8**	**9**	**10**
Nonpreterite	1s → 2s			*huʔ*	*nɛ*							
	1s → 2d			*huʔ*	*nɛ*					*ci*[3]	*ŋ*	
	1s → 2p			*huʔ*	*n(ɛ)*					*i*	*ŋ*	
	1s → 3s			*huʔr*			*u*	*ŋ*				
	1s → 3ns			*huʔr*			*u*	*ŋ*		*si*	*ŋ*	
	1pi → 3s	*a*		*huʔr*			*u*	*m*				
	1pi → 3ns	*a*		*huʔr*			*u*	*m*		*si*	*m*	
	1pe → 2			*huʔ*	*nɛ*				*ci*			*ge*
	1pe → 3s			*huʔr*			*u*	*m*				*be*[4]
	1pe → 3ns			*huʔr*			*u*	*m*		*si*	*m*	*be*[4]
	2 → 1	*a*	*gɛ*[2]	*huʔ*								
	2p → 3s		*kɛ*	*huʔr*			*u*	*m*				
	2p → 3ns		*kɛ*	*huʔr*			*u*	*m*		*si*	*m*	
Preterite	1s → 2s			*huʔ*	*n(ɛ)*	*ɛ*						
	1s → 2d			*huʔ*	*n(ɛ)*	*ɛ*				*ci*[3]	*ŋ*	
	1s → 2p			*huʔ*	*n(ɛ)*	*(ɛ)*				*i*	*ŋ*	
	1s → 3s			*huʔr*		*(ɛ)*	*u*	*ŋ*				
	1s → 3ns			*huʔr*		*(ɛ)*	*u*	*ŋ*		*si*	*ŋ*	
	1pi → 3s	*a*		*huʔr*		*(ɛ)*	*u*	*m*				
	1pi → 3ns	*a*		*huʔr*		*(ɛ)*	*u*	*m*		*si*	*m*	
	1pe → 2			*huʔ*	*n(ɛ)*	*ɛ*			*ci*			*ge*
	1pe → 3s			*huʔ*					*mʔna*			
	1pe → 3ns			*huʔ*					*mʔna*	*si*		
	2 → 1	*a*	*gɛ*[2]	*huʔr*		*ɛ*						
	2p → 3s		*kɛ*	*huʔr*			*u*	*m*				
	2p → 3ns		*kɛ*	*huʔr*			*u*	*m*		*si*	*m*	

1. *huʔr* is a prevocalic alternant of *huʔ*.
2. *gɛ* is an alternant of *kɛ* (**Driem1987**)
3. *s* becomes *c* after *ɛ* (**Driem1987**)
4. *be* is a phonologically conditioned alternant of *ge* (**Driem1987**)

The distribution of these suffixes is, in fact, doubly puzzling. Besides participating in relations of multiple exponence, they also exhibit gaps in their distribution. Consider first the suffix *-ŋ*. Because ten of the forms in Table 1 realize first-person singular agent properties, all ten would be compatible with the appearance of the *-ŋ* suffix in both the **sf5** and **sf9** positions. Yet, only two of the forms exhibit *-ŋ* in both positions; two exhibit it only in position **sf5**; four, only in postion **sf9**; and two lack *-ŋ* altogether. Consider likewise the distribution of the nonthird-person plural agent suffix *-m*. Among the fourteen forms in Table 1 that realize nonthird-person plural agent properties, only five exhibit *-m* in both the **sf5** and **sf9** positions; five have it in the **sf5** position only; and four lack *-m* altogether.

A cursory examination reveals the distributional generalization accounting for these results: *-ŋ* and *-m* appear in position **sf5** only if there is an overt affix in position **sf4**, and they appear in **sf9** only if there is an overt affix in **sf8**. In other words, the rules that introduce *-ŋ* and *-m* in Limbu are dependent rules whose application presumes that of a carrier rule filling position **sf4** or **sf8**. Because there are carrier rules in more than one rule block, the *-ŋ* and *-m* rules may both figure in the application of more than one block.

3 Polyfunctional concordial morphology in the verb inflection of Southern Sotho

Typically of Bantu languages, Southern Sotho has a rich noun-class system one of whose manifestations is the inflection of verbs for the noun class of their subject and object arguments. In the analysis proposed by **doke1985** this system exhibits seven noun classes; these have the effect of subclassifying the third person, so that like the first and second persons, each noun class subsumes both singular and plural forms. Table 2 presents the inventory of prefixes by which verbs inflect for the person, number and noun class of their subject. By a similar inventory of prefixes, transitive verbs may[6] inflect for the properties of their object; the examples in Table 3 illustrate. Table 4 presents the inventories of subject-coding and object-coding prefixes side by side; as this table shows, the two inventories are nearly identical; the only exceptions are in the singular of the first person and of class 1, where the exponents of subject properties differ from those of the corresponding object properties.

The principal difference between the two inventories is morphotactic: subject-coding prefixes occupy the position before that of tense prefixes (such as the future-tense prefix *tla-* in Tables 2 and 3) while object-coding prefixes occupy the position following that of tense prefixes. Thus, the general pattern is that the prefixes in Table 4 express properties of person, number and noun class, and that it is a prefix's position that determines whether the properties that it expresses are subject or object properties. Put another

[6] Unlike the subject concords, whose use is obligatory in finite forms, the object concords are optional, generally being use to express a pronominal object rather than to express agreement with an overt object phrase (**doke1985**).

Table 2: Future-tense forms of Southern Sotho BÒNA 'see': 'I / you / etc. will see' (**doke1985**)

Subject person	Subject class			Subject number	
	Doke & Mofokeng	Meinhof		Singular	Plural
		sg	pl		
1				*kē-tla-bòna*	*rē-tla-bòna*
2				*u-tla-bòna*	*lē-tla-bòna*
3	1	1	2	*ō-tla-bòna*	*ba-tla-bòna*
	2	3	4	*ō-tla-bòna*	*ē-tla-bòna*
	3	5	6/10	*lē-tla-bòna*	*a-tla-bòna, li-tla-bòna*
	4	7	8	*sē-tla-bòna*	*li-tla-bòna*
	5	9	10	*ē-tla-bòna*	*li-tla-bòna*
	6	14	6	*bō-tla-bòna*	*a-tla-bòna*
	7	15/17		*hō-tla-bòna*	

Table 3: Future-tense forms of Southern Sotho BÒNA 'see': 'they will see me / you / etc.' (**doke1985**)

Object person	Object class			Object number	
	Doke & Mofokeng	Meinhof		Singular	Plural
		sg	pl		
1				*ba-tla-m-pòna*	*ba-tla-rē-bòna*
2				*ba-tla-u-bòna*	*ba-tla-lē-bòna*
3	1	1	2	*ba-tla-mō-bòna*	*ba-tla-ba-bòna*
	2	3	4	*ba-tla-ō-bòna*	*ba-tla-ē-bòna*
	3	5	6/10	*ba-tla-lē-bòna*	*ba-tla-a-bòna, ba-tla-li-bòna*
	4	7	8	*ba-tla-sē-bòna*	*ba-tla-li-bòna*
	5	9	10	*ba-tla-ē-bòna*	*ba-tla-li-bòna*
	6	14	6	*ba-tla-bō-bòna*	*ba-tla-a-bòna*
	7	15/17		*ba-tla-hō-bòna*	

way, the rules introducing the noun-class concords in Table 4 generally figure in the application of more than one rule block, expressing subject properties in one block and object properties in another.

4 Rule conflation

It is clear from the foregoing evidence that in the definition of a language's inflectional morphology, the same realization rule may figure in the application of more than one rule block. I propose that this is an effect of the phenomenon of rule conflation; in particular, I propose that when rule R figures in the application of both Blocks A and B, it is because R may conflate both with certain Block A rules and with certain Block B

Table 4: Indicative verbal concords in Southern Sotho (**doke1985**)

Person	Class			Subject		Object	
	Doke & Mofokeng	Meinhof		sg	pl	sg	pl
		sg	pl				
1				*kē-*	*rē-*	*N-**	*rē-*
2				*u-*	*lē-*	*u-*	*lē-*
3	1	1	2	*ō-*	*ba-*	*mō-*	*ba-*
	2	3	4	*ō-*	*ē-*	*ō-*	*ē-*
	3	5	6/10	*lē-*	*a-, li-*	*lē-*	*a-, li-*
	4	7	8	*sē-*	*li-*	*sē-*	*li-*
	5	9	10	*ē-*	*li-*	*ē-*	*li-*
	6	14	6	*bō-*	*a-*	*bō-*	*a-*
	7	15/17		*hō-*		*hō-*	

**N* represents a homorganic nasal

rules. I represent the conflation of R_1 with R_2 as $[R_1 \copyright R_2]$.

I make six essential assumptions about the definition of rule conflation.

4.1 Rule-block membership

A conflated rule $[R_1 \copyright R_2]$ belongs to the same rule block as its carrier rule R_2.

4.2 Forms defined by conflated rules

Where R_1 is a rule that affixes *a* by means of operation F and R_2 is a rule that affixes *b* by means of operation G, the conflated rule $[R_1 \copyright R_2]$ affixes *b′* by means of operation G, where *b′* is the result of affixing *a* to *b* by means of operation F. According to this definition, there are four logically possible patterns of conflation for rules of affixation; these are represented schematically in part (A) of Table 5. The conflation of R_1 with R_2 is analogous to function composition when R_1 and R_2 both effect prefixation or when both effect suffixation. But when R_1 is prefixational and R_2 is suffixational, the application of $[R_1 \copyright R_2]$ to stem X is X*ab* rather than *a*X*b*; and when R_1 is suffixational and R_2 is prefixational, the application of $[R_1 \copyright R_2]$ to stem X is *ba*X rather than *b*X*a*. In these latter cases, the conflation of R_1 with R_2 cannot be likened to the mathematical notion of function composition.

4.3 A conflated rule's direction of affixation

Whether $[R_1 \copyright R_2]$ is a rule of prefixation or suffixation is uniquely determined by the properties of R_1 and R_2. If R_2 is a rule of affixation, then the direction of affixation of

Table 5: Six logical possibilities for the conflation [R_1 © R_2] of a dependent rule R_1 with a carrier rule R_2′

	Dependent rule R_1	Carrier rule R_2	Conflated rule [R_1 © R_2]	[R_1 © R_2] applied to stem X
(A)	*a*-prefixation	*b*-prefixation	*ab*-prefixation	*ab*X
	a-prefixation	*b*-suffixation	*ab*-suffixation	X*ab*
	a-suffixation	*b*-prefixation	*ba*-prefixation	*ba*X
	a-suffixation	*b*-suffixation	*ba*-suffixation	X*ba*
(B)	*a*-prefixation	identity function	*a*-prefixation	*a*X
	a-suffixation	identity function	*a*-suffixation	X*a*

[R_1 © R_2] is that of R_2, as indicated in (ii) above; but if R_2 is a rule of significative absence,[7] then the direction of affixation of [R_1 © R_2] is that of R_1, as in part (B) of Table 5.

4.4 Content realized by a conflated rule

If rule R_1 realizes the morphosyntactic property set α and rule R_2 realizes the property set β, then rule [R_1 © R_2] realizes the combination of α and β. In the simplest cases, the relevant mode of combination can simply be seen as set union: $\alpha \cup \beta$. But in the general case, it is preferable to regard the mode of set combination as unification;[8] for instance, the combination of {fut, {sbj 3 sg}} with {{sbj fem}} should be the unification {fut {sbj 3 sg fem}} rather than the union {fut, {sbj 3 sg}, {sbj fem}}. That is, if R_1 realizes α and R_2 realizes β, then [R_1 © R_2] realizes the unification $\alpha \sqcup \beta$.

4.5 Recursion

The definition of rule conflation does not exclude the possibility that a conflated rule might itself enter into the conflation of a still more complex rule; that is, rule conflation

[7] A rule of significative absence realizes a particular property set by means of an identity function. In a realizational theory of morphology, a rule of significative absence realizing a property set σ overrides the overt morphology of a competing rule realizing some property set of which σ is an extension. (Cf. the analysis of Bulgarian verb inflection proposed by **Stump2001**)

[8] The assumed definition of unification is as in (i); this definition depends on the assumed definition of extension in (ii). (Cf. Gazdar et al. 1985: 27; **Stump2001**)

(i) The ***unification*** of ρ and σ [i.e. $\rho \sqcup \sigma$] is the smallest well-formed extension of both ρ and σ .
Example: {{sbj 3 sg}, {obj pl}} ⊔ {prs, {obj 1}} = {{sbj 3 sg}, prs, {obj 1 pl}}

(ii) Given two sets σ, τ: σ is an ***extension*** of τ [i.e. $\tau \sqsubseteq \sigma$] iff for each property x $\in \tau$,
either (i) x is simple property and x $\in \sigma$
or (ii) x is a complex property (= a set of properties) such that y $\in \sigma$ and y is an extension of x.
Examples: {pl} ⊑ {1 pl}
{prs {obj 1}} ⊑ {prs {obj 1 pl}}

may be recursive.

4.6 Nonconcatenative rules and conflation

A priori, there is no reason why the morphological rules that enter into such conflations must necessarily be rules of affixation or of significative absence. The most convincing cases, however, do involve rules of these two sorts, and I shall focus exclusively on such cases here. Nevertheless, nothing that I say here should be seen as excluding the possibility that nonconcatenative rules might also enter into relations of rule conflation.

Rule conflation is an operation on rules rather than on affixes; nevertheless, if R_1 and R_2 are rules introducing the respective affixes *a* and *b*, one can, as a kind of shorthand, refer to the affix *ab* (or *ba*) introduced by the conflated rule $[R_1 \copyright R_2]$ as a CONFLATED AFFIX.

As I now show, this conception of rule conflation affords a straightforward account of multiple exponence in the expression of Limbu agent inflection (§5) and of polyfunctional concordial morphology in Southern Sotho (§6).

5 Rule conflation and multiple exponence in Limbu

Consider again the inflection of Limbu verbs for the properties of their agent argument. As was seen in §2, the suffix *-ŋ* (expressing first-person singular agent properties) and the suffix *-m* (expressing nonthird-person plural agent properties) may appear in either of two positions—or in both—in a verb form's inflectional morphotactics; but their appearance in either position is dependent on that of a suffix in the immediately preceding position. The following analysis of this distributional pattern is based on two key assumptions:[9]

- the agent-coding suffixes *-ŋ* and *-m* are introduced by dependent rules that only apply in conflation with another, "carrier" rule, and
- carrier rules for the *-ŋ* and *-m* rules exist in more than one block.

This analysis[10] employs independent realization rules that introduce the suffixes[11] in Table 1; these are organized into several rule blocks, each of which fills a particular affix position. These independent rules and their block membership are given in Table 6. There are also dependent realization rules; these introduce the agent-coding suffixes *-ŋ* and *-m*, as in (5). Rule conflation is defined by the conflation rules in (6). Rule (6a) conflates the

[9] The dependent rules at issue in the proposed analyses of Limbu and Southern Sotho are only manifested in conflation with a carrier rule. But one can also imagine that a rule might be able to function both as a dependent rule and as an independent rule; the rules introducing the Swahili relative affixes are argued to have this status in the analysis proposed by Stump (to appear a).

[10] The following abbreviations are employed for the morphosyntactic properties in this analysis: agt = agent, pat = patient, pret = preterite, 1/2/3 = first/second/third person, −3 = nonthird person, excl = exclusive, −incl = noninclusive, ns = nonsingular, sg = singular, pl = plural.

[11] Concerning the person prefixes *a-* and *kɛ-* in Table 1, see van **Driem1987** 77ff.

dependent rules with the three carrier rules identified in Table 6: **4**-a, **8**-a and **8**-b. The resulting conflated rules are listed (redundantly) in Table 7.

Because a conflated rule belongs to the same rule block as the carrier rule on which it is based, the conflated rule and the carrier rule compete to realize certain morphosyntactic property sets; in any such instance of competition, the conflated rule prevails by virtue of the fact that its domain of application is smaller than that of the carrier rule.[12]

Table 6: Some independent realization rules of Limbu verb inflection

Block	Rule label	Realization rules		Carrier
		Properties realized	Operation	rule?
sf1	**1**-a	{{agt 1}{pat 2}}	X → X*nɛ*	no
sf2	**2**-a	{pret}	X → X*ɛ*	
sf4	**4**-a	{{pat 3}}	X → X*u*	yes
sf7	**7**-a	{{agt 1 ns}{pat 2}}	X → X*ci*	no
sf8	**8**-a	{{pat ns}}	X → X*si*	yes
	8-b	{{pat–3 –incl pl}}	X → X*i*	
sf10	**10**-a	{{excl}}	X → X*ge*	no
sf11	**11**-a	{{agt 1 pl excl} pret {pat 3}}	X → X*m?na*	yes, for **8**-a

(5) Dependent realization rules lgfamsrc
- ŋ. {{agt1 sg}} : X→ X*ŋ* (van **Driem1987** 99)
- M. {{agt –3 pl}} : X → X*m* (van **Driem1987** 99f)

(6) Conflation rules
- a. Where R is a rule in Block α ($\alpha \in$ {**4**,**8**}), ŋ © R, [M © R] ∈ Block α.
- b. [**8**-a © **11**-a] ∈ Block **11**.[13]

This analysis correctly defines all of the forms in Table 1. In particular, it accounts for the superficially erratic distribution of the agent concords *-ŋ* and *-m*. Thus, Table 8 presents the manner in which the rules in Tables 6 and 7 define four words:

[12] Concerning each rule in Table 6, see **Driem1987** **1**-a, pp.88f; **2**-a, pp.89ff; **4**-a, p.82; **7**-a, p.100; **8**-a, pp.101f; **8**-b, pp.95f; **10**-a, pp.102f; **11**-a, 100f.

[13] Conflation rule (6b) helps to resolve a conundrum in Table 1. Notice first that the suffix *-m?na* introduced by rule **11**-a as an exponent of the property set {{agt 1 pl excl} pret {pat 3}} only combines with one other suffix, namely the suffix *-si* introduced by rule **8**-a as an exponent of {{pat ns}}; yet, it is featurally compatible with the suffixes introduced by **4**-a and **10**-a. Moreover, the suffix *-si* in the form *hu?-m?na-si* 'we (excl) taught them' does not carry *-m*, even though (a) it is a carrier elsewhere and (b) *-m* would be featurally appropriate for this word form. I therefore depart from van Driem in postulating Block **11** as a portmanteau rule block (**Stump2001**) that is paradigmatically opposed to and defaults to the sequence of other suffixal blocks. It houses exactly two rules: the simple rule **11**-a (which suffixes *-m?na*) and the conflated rule [**8**-a © **11**-a] (which suffixes *-m?na-si*). Because Block **11** is paradigmatically opposed to the sequence of rule blocks to which Block **8** belongs, the application of rule [**8**-a © **11**-a] excludes that of rule [M © **8**-a], effectively blocking the appearance of *-m* in forms such as *hu?-m?na-si* 'we (excl) taught them'.

Table 7: Some conflated realization rules of Limbu verb inflection

Block	Rule label	Realization rules	
		Properties realized	Operation
4	[ɲ © **4**-a]	{{agt1 sg}{pat 3}}	X → X*uŋ*
	[M © **4**-a]	{{agt −3 pl}{pat 3}}	X → X*um*
8	[ɲ © **8**-a]	{{agt1 sg}{pat ns}}:	X → X*siŋ*
	[ɲ © **8**-b]	{{agt1 sg}{pat−3 −incl pl }}	X → X*iŋ*
	[M © **8**-a]	{{agt −3 pl}{pat ns}}:	X → X*sim*
11	[**8**-a © **11**-a]	{{agt 1 pe} pret{pat 3 ns}}	X → X*m?nasi*

- *hu?r-u-ŋ-si-ŋ* 'I teach them', in which *-ŋ* appears twice—after *-u* and after *-si*;
- *hu?r-u-ŋ* 'I teach him', in which *-ŋ* appears after *-u* only;
- *hu?-nɛ-ci-ŋ* 'I teach you two', in which *-ŋ* appears after *-si* only; and
- *hu?-nɛ* 'I teach you (sg.)', in which *-ŋ* fails to appear.

As Table 8 shows, *-ŋ* only appears in conflation with an immediately preceding carrier: in one case, it appears twice because there are two carriers to conflate with; in another, only the carrier *-u* is available; in yet another, only the carrier *-si* is available; and sometimes, there is no carrier at all to conflate with. The proposed analysis provides a similar account of the comparable behavior of the suffix *-m*.

6 Rule conflation and polyfunctional concord in Southern Sotho

Return now to the morphology of verbal concord in Southern Sotho. As we saw in §3, this morphology is largely polyfunctional. Typically, a verbal concord may appear in either of two positions in a verb's inflectional morphotactics; but unlike the agent-coding suffixes in Limbu, which express the same content no matter where they appear, the Southern Sotho verbal concords express subject properties in one position but object properties in another. The notion of rule conflation makes it possible to account for this difference by assuming that in Southern Sotho, the rules expressing noun-class concord conflate with a general rule of subject concord in one block and with a general rule of object concord in a different block. Because the two general rules are formulated as identity functions (realizing subject concord and object concord, respectively), the conflated subject concords have the same phonological form as the conflated object concords.

Thus, consider the following definition of the Southern Sotho inflectional markings in Tables 2 and 3. In this analysis, there are three blocks of independent realization rules, as in Table 9. Block **a** houses the rules of object concord: these include the special object-concord rules for the first-person singular (**a**-i) and third singular class 1 (**a**-ii); in

Table 8: The definition of four Limbu verb forms in the proposed analysis

Property set:	{{agt 1 s}{pat 3 ns}}	{{agt 1 s}{pat 3 s}}
Stem:	*huʔ* 'teach' (prevocalically *huʔr*)	*huʔ* 'teach' (prevocalically *huʔr*)
Rule applying in Block:		
sf1:	(none)	(none)
sf2:	(none)	(none)
sf4:	[ɳ © **4**-a]: *huʔr-u-ŋ*	[ɳ © **4**-a]:*huʔr-u-ŋ*
sf7:	(none)	(none)
sf8:	[ɳ © **8**-a]: *huʔr-u-ŋ-si-ŋ*	(none)
sf10:	(none)	(none)
sf11:	(none)	(none)
	huʔr-u-ŋ-si-ŋ 'I teach them'	*huʔr-u-ŋ* 'I teach him'

Property set:	{{agt 1 s}{pat 2 de}}	{{agt 1 s}{pat 2 s}}
Stem:	*huʔ* 'teach'	*huʔ* 'teach'
Rule applying in Block:		
sf1:	**1**-a: *huʔ-nɛ*	**1**-a: *huʔ-nɛ*
sf2:	(none)	(none)
sf4:	(none)	(none)
sf7:	(none)	(none)
sf8:	[ɳ © **8**-a]: *huʔ-nɛ-ci-ŋ*	(none)
sf10:	(none)	(none)
sf11:	(none)	(none)
	huʔ-nɛ-ci-ŋ 'I teach you two'	*huʔ-nɛ* 'I teach you (sg.)'

addition, it includes a default rule (**a**-iii) realizing object concord by means of an identity operation. Block **b** houses rules realizing tense properties, here exemplified by the future tense. Block **c** houses rules of subject concord, including the special rule (**c**-i) of first-person singular subject concord and a default rule (**c**-ii) realizing subject concord by means of an identity operation. In addition to the independent realization rules in Table 9, the analysis requires the large inventory of dependent rules in Table 10. The conflation rule in (7) conflates each dependent rule with the default object-concord rule (**a**-iii) and with the default subject-concord rule (**c**-ii). The resulting conflated rules are listed (redundantly) in Table 11.

Table 9: Three blocks of independent realization rules in Southern Sotho

Block	Rule label	Realization rules: Properties realized	Realization rules: Operation	Carrier rule?
a	**a**-i	{{obj 1 sg}}	X ⟶ *N*X*	
	a-ii	{{obj 3 sg CL:1}}	X ⟶ *mō*X	
	a-iii	{{obj}}	X ⟶ X	yes
b	**b**-i	{fut}	X ⟶ *tla*X	
c	**c**-i	{{sbj 1 sg}}	X ⟶ *kē*X	
	c-ii	{{sbj}}	X ⟶ X	yes

*N represents a homorganic nasal.

Table 10: Dependent realization rules for verbal concord in Southern Sotho

Rule label	Realization rules: Properties realized	Realization rules: Operation
agr-i	{{2 sg}}	X ⟶ *u*X
agr-ii	{{3 sg}}	X ⟶ *ō*X
agr-iii	{{3 sg CL:3}}	X ⟶ *lē*X
agr-iv	{{3 sg CL:4}}	X ⟶ *sē*X
agr-v	{{3 sg CL:5}}	X ⟶ *ē*X
agr-vi	{{3 sg CL:6}}	X ⟶ *bō*X
agr-vii	{{3 CL:7}}	X ⟶ *hō*X
agr-viii	{{1 pl}}	X ⟶ *rē*X
agr-ix	{{2 pl}}	X ⟶ *lē*X
agr-x	{{3 pl CL:1}}	X ⟶ *ba*X
agr-xi	{{3 pl CL:1|2}}	X ⟶ *ē*X
agr-xii	{{3 pl CL:3}}	X ⟶ *a*X | *li*X
agr-xiii	{{3 pl CL:4|5}}	X ⟶ *li*X
agr-xiv	{{3 pl CL:6}}	X ⟶ *a*X

(7) Conflation rule
Where **agr**-*n* is a dependent realization rule and R is a carrier rule in Block α, [**agr**-*n* © R] ∈ Block α.

Table 11: Some conflated realization rules of Southern Sotho verb inflection

Block	Rule label	Realization rules	
		Properties realized	Operation
a	**[agr**-i © **a**-iii]	{{obj 2 sg}}	X → *u*X
	[agr-ii © **a**-iii]	{{obj 3 sg CL:1|2}}	X → *o͞*X
	[agr-iii © **a**-iii]	{{obj 3 sg CL:3}}	X → *le͞*X
	[agr-iv © **a**-iii]	{{obj 3 sg CL:4}}	X → *se͞*X
	[agr-v © **a**-iii]	{{obj 3 sg CL:5}}	X → *e͞*X
	[agr-vi © **a**-iii]	{{obj 3 sg CL:6}}	X → *bo͞*X
	[agr-vii © **a**-iii]	{{obj 3 CL:7}}	X → *ho͞*X
	[agr-viii © **a**-iii]	{{obj 1 pl}}	X → *re͞*X
	[agr-ix © **a**-iii]	{{obj 2 pl}}	X → *le͞*X
	[agr-x © **a**-iii]	{{obj 3 pl CL:1}}	X → *ba*X
	[agr-xi © **a**-iii]	{{obj 3 pl CL:1|2}}	X → *e͞*X
	[agr-xii © **a**-iii]	{{obj 3 pl CL:3}}	X → *a*X | *li*X
	[agr-xiii © **a**-iii]	{{obj 3 pl CL:4|5}}	X → *li*X
	[agr-xiv © **a**-iii]	{{obj 3 pl CL:6}}	X → *a*X
c	**[agr**-i © **c**-ii]	{{sbj 2 sg}}	X → *u*X
	[agr-ii © **c**-ii]	{{sbj 3 sg CL:1|2}}	X → *o͞*X
	[agr-iii © **c**-ii]	{{sbj 3 sg CL:3}}	X → *le͞*X
	[agr-iv © **c**-ii]	{{sbj 3 sg CL:4}}	X → *se͞*X
	[agr-v © **c**-ii]	{{sbj 3 sg CL:5}}	X → *e͞*X
	[agr-vi © **c**-ii]	{{sbj 3 sg CL:6}}	X → *bo͞*X
	[agr-vii © **c**-ii]	{{sbj 3 CL:7}}	X → *ho͞*X
	[agr-viii © **c**-ii]	{{sbj 1 pl}}	X → *re͞*X
	[agr-ix © **c**-ii]	{{sbj 2 pl}}	X → *le͞*X
	[agr-x © **c**-ii]	{{sbj 3 pl CL:1}}	X → *ba*X
	[agr-xi © **c**-ii]	{{sbj 3 pl CL:1|2}}	X → *e͞*X
	[agr-xii © **c**-ii]	{{sbj 3 pl CL:3}}	X → *a*X | *li*X
	[agr-xiii © **c**-ii]	{{sbj 3 pl CL:4|5}}	X → *li*X
	[agr-xiv © **c**-ii]	{{sbj 3 pl CL:6}}	X → *a*X

Each of the conflated rules in Table 11 belongs to the same rule block as the carrier rule on which it is based. As in the Limbu analysis proposed above, a conflated rule and the carrier rule on which it is based compete to realize certain morphosyntactic property sets, and being the narrower rule, the conflated rule prevails in each such case.

This analysis correctly defines all of the forms in Tables 2 and 3. In particular, it accounts for the fact that in all but a handful of cases, each subject concord has a corresponding object concord that expresses the same person, number and noun class by means of the same prefix. Thus, Table 12 presents the manner in which the rules in Tables 10 and 11 define two words:

- *ba-tla-bō-bòna* ‘they (CL:1) will see it (CL:6)’, in which *ba-* is a third-person plural class 1 subject concord and *bō-* is a singular class 6 object concord; and
- *bō-tla-ba-bòna* ‘it (CL:6) will see them (CL:1)’, in which *bō-* is a singular class 6 subject concord and *ba-* is a third-person plural class 1 object concord.

Table 12: The definition of two Southern Sotho verb forms in the proposed analysis

Property set:		{{sbj 3 pl CL:1} fut {obj 3 sg CL:6}}	
Stem:		*bòna* ‘see’	
Rule applying in			
Block	**a**:	**[agr**-vi © **a**-iii]:	*bō-bòna*
	b:	**b**-i:	*tla-bō-bòna*
	c:	**[agr**-x © **c**-ii]:	*ba-tla-bō-bòna*
		ba-tla-bō-bòna	
		‘they (CL:1) will see it (CL:6)’	
Property set:		{{sbj 3 sg CL:6} fut {obj 3 pl CL:1}}	
Stem:		*bòna* ‘see’	
Rule applying in			
Block	**a**:	**[agr**-x © **a**-iii]:	*ba-bòna*
	b:	**b**-i:	*tla-ba-bòna*
	c:	**[agr**-vi © **c**-ii]:	*bō-tla-ba-bòna*
		bō-tla-ba-bòna	
		‘it (CL:6) will see them (CL:1)’	

As Table 12 shows, the dependent rules introducing *bō-* (**agr**-vi in Table 10) and *-ba* (**agr**-x in Table 10) both conflate with the carrier rule **a**-iii (Table 9) to produce rules of object concord in Block **a** and both conflate with the carrier rule **c**-ii (Table 9) to produce a rule of subject concord in Block **c**.

7 Wider evidence for rule conflation

The analyses proposed here for multiple exponence in Limbu agent concord and for polyfunctional verbal concords in Southern Sotho both depend on the notion that morphological rules may conflate to produce more complex rules (= principle (3)) and the notion that conflated rules may compete with simple rules as members of the same rule block (= principle (4)).

These principles of rule conflation are motivated independently of the need to account for multiple exponence and polyfunctionality. First, they make it possible to account for apparent anomalies in the interaction of inflectional rule applications. For example, a rule’s order of application may seem to depend on whether or not another rule applies.

In Fula, a pronominal object suffix on a verb in the relative past tense ordinarily follows that verb's subject suffx, as in (8a,b); but in the particular case in which a verb has both a singular personal object suffix (2sg *-mA* or 3sg *-mO*) and the first-person singular subject suffix *-mi*, the expected order is reversed, as in (8c,d). Principles (3) and (4) allow one to say that the rules realizing the subject and object suffixes in the relative past tense belong to a single rule block; that the object rules ordinarily conflate with the subject rules; but that the *-mi* rule instead conflates with the *-mA* and *-mO* rules.

(8) a. *mball-u-mi-ɓe-'*
help-REL.PST.ACT-1SG.SBJ-3PL.CL.2.OBJ-FG
'I helped them'

b. *mball-u-ɗaa-mO-'*
help-REL.PST.ACT-2SG.SBJ-3SG.CL.1.OBJ-FG
'you (sg.) helped him'

c. *mball-u-mA-mi-'*
help-REL.PST.ACT-2SG.OBJ-1SG.SBJ-FG
'I helped you (sg.)'

d. *mball-u-mO-mi-'*
help-REL.PST.ACT-3SG.CL.1.OBJ-1SG.SBJ-FG
'I helped him' (**Arnott1970** Appendix 15)

In another apparently anomalous interaction of inflectional rules, an affix either precedes the stem with which it joins or follows it, with the choice of position being conditioned by the presence or absence of some other affix. In Swahili, the verbal concord coding the properties of a relative verb form's relativized argument appears postverbally in tenseless affirmative forms, but preverbally in forms that are prefixally marked for tense or negation; thus, the class 8 relative concord *vyo* is postverbal in (9a) but preverbal in (9b). The principles of rule conflation make it possible to say that the relative affix is suffixed to the verb stem by default, but is suffixed to an overt prefixal exponent of tense or negation (Stump to appear a).

(9) a. *a-vi-soma-vyo*
SBJ:CL.1-OBJ:CL.8-read-REL:CL.8
'(books) which he reads'

b. *a-si-vyo-vi-soma*
SBJ:CL.1-NEG-REL:CL.8-OBJ:CL.8-read]
'(books) which he doesn't read'

As Stump (to appear b) shows, the principles of rule conflation afford simple solutions to a number of other apparent anomalies in the interaction of inflection rules. These include the incidence of variable affix order (**Bickel2007**) and of Wackernagel affixes (**Nevis1992 Bonami2008**) as well as the superficially puzzling fact that affix sequences

may preserve the same internal order whether the sequence as a whole is prefixal or suffixal, as in European Portuguese verb inflection (**Luis2005**).

Second, the principles of rule conflation in (3) and (4) make it possible to account for nonmonotonic interactions among inflectional rules. The usual expectation is that a realization rule possesses the same intrinsic properties whether it applies alone or in combination with other rules. But there are anomalous cases in which this expectation is not met. Once the definition of a language's morphology includes a conflated rule $[R_1 \copyright R_2]$, this rule may evolve independently, taking on properties not directly stemming from either R_1 or R_2. In this way, the properties exhibited by a rule applying in isolation may not always be preserved when it is conflated with other rules. In view of this fact, the content attributed to conflated rules in §4.4 above should be seen as their *default* content, subject to modification by processes of grammaticalization. That is, the content expressed by rule $[R_1 \copyright R_2]$ is, in the default case, a monotonic function of the content expressed by rules R_1 and R_2; but this default is subject to override.

There are at least three ways in which the resulting nonmonotonicity may be manifested. One reflection of this fact is the phenomenon of 'potentiation' (**Williams1981**), by which an unproductive rule becomes productive when applying in combination with another rule (as the unproductive *-ity* rule becomes productive in combination with the *-able* rule; cf. **Aronoff1976 Bochner1992**).

Another reflection is the fact that the domain of rule R_1 apparently depends on whether a particular rule R_2 applies subsequently. By principles (3) and (4), such cases arise when a conflated rule $[R_1 \copyright R_2]$ evolves a domain application distinct from that of R_2. Thus, a stem may be in the domain of R_2 but not that of $[R_1 \copyright R_2]$, as in the case of *base* → *basic*, **basical*; at the same time, a stem may be in the domain of $[R_1 \copyright R_2]$ but not that of R_2, as in the case of *whimsy* → *whimsical*, **whimsic*. A third reflection arises in cases in which two rules apparently realize less content separately than they do together. In Latin *regēmus* 'we shall rule', the conflation of the rules that suffix *-ē* and *-mus* expresses the first-person plural future active even though neither rule by itself is an expression of future tense.[14] These nonmonotonic phenomena have never before been seen as manifestations of a single overarching principle; the principles of rule conflation, however, facilitate precisely such a perspective.

Third, the principles of rule conflation make it possible to account for parallelisms between the application of a single rule and that of a sequence of rules. A word form's inflectional morphology is sometimes informally conceived of as instantiating a sequence of "slots" each of which corresponds to a set of rules available to fill it. Andersonian rule blocks are a kind of formal reconstruction of this idea, whose simplest interpretation involves individual rules providing alternative ways of filling the same slot. There are, however, apparent deviations from this pattern, in which successive slots are ordinarily filled by successive rule applications but may in some instances be simultaneously filled by a single rule application introducing a "wide" affix that somehow straddles two or more slots. The Swahili portmanteau prefix *si-* is an example. In Swahili negative indicative verb forms, the usual pattern is for the negative *ha-* rule to fill slot 1 and a

[14] See Stump (to appear b) for discussion of a similar case from Old English.

subject-concord rule to fill slot 2, e.g. *ha-tu-ta-taka* [NEG-1PL-FUT-want] 'we will not want'. But in first-person singular negative verb forms, the application of the negative first-person singular *si-* rule seems to straddle slots 1 and 2. The principles of rule conflation resolve this conundrum by allowing a rule block to contain both conflated rules (e.g. the first-person plural negative *ha-tu-* rule) and simple rules (e.g. the *si-* rule) in paradigmatic opposition; in this way, the behavior of portmanteau rules is reconciled with the natural assumption that paradigmatic opposition is a relation between two rules rather than a relation between a rule and a sequence of rules.

The principles of rule conflation in (3) and (4) are a simple and natural extension of the principles of realization-rule interaction developed by Anderson (see again (1) and (2)). Rule conflation allows a variety of apparently recalcitrant phenomena to be reconciled with a general scheme of rule interaction based on ordered blocks of realization rules in which the members of a given block are mutually exclusive in their application.

Chapter B

u-umlaut in Icelandic and Faroese: Survival and death

Höskuldur Thráinsson
University of Iceland

Although Icelandic and Faroese are closely related and very similar in many respects, their vowel systems are quite different (see e.g. **anderson1969a**; **arnason2011**). This paper compares *u*-umlaut alternations in Icelandic and Faroese and shows that the Faroese umlaut has a number of properties that are to be expected if the relevant alternations are morphological (or analogical) rather than being due to a synchronic phonological process. In Icelandic, on the other hand, *u*-umlaut has none of these properties and arguably behaves like a living phonological process. This is theoretically interesting because the quality of the vowels involved (both the umlaut trigger and the target) has changed from Old to Modern Icelandic. In addition, *u*-umlaut in Modern Icelandic is more opaque (in the sense of **kiparsky1973**) than its Old Icelandic counterpart, i.e. it has more surface exceptions. An epenthesis rule inserting a (non-umlauting) /u/ into certain inflectional endings is the cause of many of these surface exceptions. Yet it seems that *u*-umlaut in Icelandic is still transparent enough to be acquired by children as a phonological process. In Faroese, on the other hand, *u*-umlaut became too opaque and died out as a phonological rule. It is argued that this has partly to do with certain changes in the Faroese vowel system and partly with the fact that the u-epenthesis rule was lost in Faroese.

1 Introduction

Anderson put the process of *u*-umlaut in Icelandic on the modern linguistic map with the analysis he proposed in his dissertation (**Anderson1969**) and several subsequent publications (**anderson1969b**; **anderson1972t**; **anderson1973**; **anderson1974**; **anderson1976t**). Because of changes in the vowel system from Old to Modern Icelandic, the nature of the umlaut process changed somewhat through the ages (see e.g. **benediktsson1959**). The most important part of *u*-umlaut, and the only part that is alive in the modern language, involves /a/ ~ /ǫ/ alternations in the old language (phonetically [a] ~ [ɔ], as shown in (2)), which show up as /a/ ~ /ö/ alternations in the modern language (phonetically [a] ~ [œ], cf. (2)). This is illustrated in (1) with the relevant vowel symbols highlighted:

(1)	Old Icelandic:	Modern Icelandic:
	saga 'saga', OBL *sǫgu*, PL *sǫgur*	*saga*, OBL *sögu*, PL *sögur*
	hvass 'sharp', DAT *hvǫssum*	*hvass*, DAT *hvössum*
	tala 'speak', 1.PL *tǫlum*	*tala*, 1PL *tölum*

As these examples suggest, the quality of the root vowel /a/ changes when a /u/ follows in the next syllable. The relevant proecesses can be illustrated schematically as in (2). For the sake of simplicity I use conventional orthographic symbols to represent the vowels and only give IPA-symbols for the vowels that are important for the understanding of the umlaut processes. The umlaut-triggering vowels are encircled:[1]

(2) a. *u*-umlaut in Old Icelandic and the system of short vowels:

	[-back]		[+back]	
	[-round]	[+round]	[-round]	[+round]
[+high]	i	y		ⓤ [u]
	e	ø		o
[+low]	ę		a [a]	ǫ [ɔ]

b. *u*-umlaut in Modern Icelandic and the system of monophthongs:[2]

	[-back]		[+back]	
	[-round]	[+round]	[-round]	[+round]
[+high]	í			ú [u]
	i	ⓤ [ʏ]		
[+low]	e	ö [œ]	a [a]	o [ɔ]

The gist of Anderson's analysis of *u*-umlaut can then be illustrated semi-formally as in the traditional generative phonological notation in (3), with the assimilating features highlighted (see also **roegnvaldsson1981 thrainsson2011**):[3]

(3) a. *u*-umlaut in Old Icelandic:

$$/a/ \rightarrow [\textbf{+round}] \; / \; \_C_0V \begin{bmatrix} \textbf{+round} \\ +high \\ +back \end{bmatrix}$$

b. *u*-umlaut in Modern Icelandic:

$$/a/ \rightarrow \begin{bmatrix} \textbf{+round} \\ \textbf{-back} \end{bmatrix} \; / \; \_C_0V \begin{bmatrix} \textbf{+round} \\ \textbf{-back} \\ -low \end{bmatrix}$$

[1] Note that in the representation of the Modern Icelandic vowel system, the accents over vowel symbols have nothing to do with quantity but simply denote separate vowel qualities. Thus /í/ is [i], /i/ is [ɪ], /ú/ is [u] and /u/ is [ʏ], as the schematic representation in (2) suggests.

[2] I am assuming here, like Thráinsson (1994) and Gíslason and Thráinsson (2000, p. 34), for instance, that Modern Icelandic only distinguishes between three three vowel heights and that /e/ [ɛ] and /ö/ [œ] are both phonologically [+low], like /a/ [a] and /o/ [ɔ]. For different assumptions see e.g. **arnason2011**

[3] Here, and elsewhere in this paper, I will use the kinds of formulations of rules and conditions familiar from classical generative phonology since much of the work on *u*-umlaut has been done in that kind of framework. For analyses employing more recent frameworks see **gibson2000 hansson2013** and **ingason2016** Most of the argumentation in this paper should be relatively framework-independent, however.

As the illustration in (3) shows, the modern version of the umlaut is somewhat more complex than the old one, assimilating two features rather than one. Nevertheless, it is still arguably a phonologically (or phonetically) natural assimilation process, assimilating rounding and backness.

Although the *u*-umlaut discussion was most lively on the international scene in the 1970s (see e.g. **iverson1978t**; **iverson1976**; **oresnik1975**; **oresnik1977** cf. also **valfells1967**), the topic keeps popping up to this day, e.g. in journals and conferences dedicated to Scandinavian linguistics (see e.g. **gibson2000**; **indridason2010**; **thrainsson2011**; **hansson2013**) and even in recent master's theses and doctoral dissertations (see **markusson2012**; **ingason2016**). The main reason is that while *u*-umlaut in Modern Icelandic is obviously very productive, being applied consistently to new words and loanwords, it shows a number of intriguing surface exceptions. These have been discussed extensively in the literature cited but here I will concentrate on the most common and widespread one, namely the lack of umlaut before a /u/ that has been inserted between the inflectional ending /r/ and a preceding consonant. This epenthesis did not exist in Old Icelandic as illustrated in (4):

(4)	Old Icelandic:	Modern Icelandic:
	dalr 'valley', *latr* 'lazy'	*dal**u**r*, *lat**u**r*

If *u*-umlaut is a phonological rule in the modern language, this *u*-epenthesis has to follow it, as it did historically. This is one of the properties of *u*-umlaut that have been used to argue for the necessity of relatively abstract phonological representations and derivations (e.g. **anderson1969a**; **anderson1974**; **roegnvaldsson1981**; **thrainsson2011**; **hansson2013**) while others have maintained that *u*-umlaut is not a phonological process anymore in Modern Icelandic and the relevant alternations are morphologized and purely analogical (see e.g. **markusson2012**) or at least "morpheme-specific", i.e. triggered by particular morphemes that may or may not contain a /u/ (**ingason2016**).[4]

In this paper I will compare *u*-umlaut alternations in Modern Icelandic and Modern Faroese. This comparison will show very clearly that *u*-umlaut in Modern Faroese has a number of properties (e.g. paradigm levelling, various kinds of exceptions, total absence

[4] **ingason2016** formulates his umlaut rule as follows:

> Realize an underlying /a/ as /ö/ in the syllable which precedes the morpheme which triggers the umlaut.

As can be seen here, no mention is made of a triggering /u/ in the rule. The reason is that Ingason wants to derive all all paradigmatic /a/ ~ /ö/ alternations the same way, including the ones where /u/ has been syncopated historically. Thus he argues that the NOM.SG. morpheme -ø in feminine nouns like *sök* 'guilt, case' and the NOM./ACC.PL. morpheme -ø in neuter nouns like *börn* 'children' triggers umlaut the same way that the DAT.PL. morpheme *-um* does in *sökum* and *börnum*. But many researchers have wanted to distinguish between morphologically conditioned umlaut, where there is no triggering /u/, and phonologically conditioned umlaut triggered by /u/, e.g. **roegnvaldsson1981** One reason for doing so comes from the behavior of loanwords like the adjective *smart* 'smart, chic'. Here the NOM.SG.F and the NOM/ACC.PL.N can either be *smart* or *smört*, i.e. a morphologically conditioned umlaut may or may not apply. But once an umlauting inflectional ending containing /u/ is added to the loanword *smart*, the *u*-umlaut becomes obligatory. Thus DAT.PL can only be *smört-um* and not **smart-**u**m* and the NOM.PL.WK form has to be *smört-**u*** and not **smart-**u***. This suggests that the morphologically conditioned umlaut is more prone to exceptions than the phonologically conditioned one, which is actually to be expected. Thanks to Eiríkur Rögnvaldsson for pointing this out to me.

from certain paradigms, inapplicability to loanwords ...) that are to be expected if the relevant alternations are no longer due to a synchronic process. In Modern Icelandic, on the other hand, *u*-umlaut has none of these properties and behaves more like a phonological rule. This is of general theoretical interest since it illustrates how phonological rules can survive (in the case of Icelandic) despite reduced transparency (in the sense of **kiparsky1973**) and how changes in the phonological system can cause the death of a phonological rule (in the case of Faroese) and what the consequences can be.

The remainder of the paper is organized as follows: In section 2 I first illustrate how the *u*-epenthesis works in Modern Icelandic and then present a couple of arguments for the phonological (as opposed to morphological) nature of Modern Icelandic *u*-umlaut. Section 3 first describes some facts about the Faroese vowel system that must have been important for the development of *u*-umlaut and then shows that *u*-epenthesis does not exist anymore as a phonological process in Modern Faroese. It is then argued that these developments led to the death of *u*-umlaut as a phonological process in Faroese. Section 4 then contains a systematic comparison of *u*-umlaut alternations in Modern Icelandic and Faroese, concluding that the Faroese ones must be analogical (and morphological) in nature as they do not exhibit any of the crucial phonological properties that Modern Icelandic *u*-umlaut alternations show. In Icelandic, on the other hand, *u*-umlaut does not show the non-phonological properties listed for its Faroese counterpart. Section 5 concludes the paper.

2 *u*-epenthesis and *u*-umlaut in Modern Icelandic

2.1 The epenthesis rule

The phoneme /r/ frequently occurs in Old Icelandic (Old Norse) as a marker of various morphological categories, including NOM.SG of strong masculine nouns and adjectives as illustrated in (5). It sometimes assimilated to a preceding consonant, e.g. /s, l, n/ (cf. (5c)),[5] but it was deleted after certain consonant clusters, such as /gl, gn, ss/ (cf. (5d)):

(5) a. *stór-r* 'big', *mó-r* 'peat', *há-r* 'high'
b. *dal-r* 'valley', *lat-r* 'lazy', *tóm-r* 'empty', *harð-r* 'hard'
c. *ís-s* 'ice', *laus-s* 'loose', *stól-l* 'chair', *fín-n* 'fine'
d. *fugl* 'bird', *vagn* 'wagon', *foss* 'waterfall' (stem *foss-*)

It is likely that the /r/ in words of type (5b) was syllabic in Old Icelandic. There are no syllabic consonants in Modern Icelandic, on the other hand. Instead a /u/ appears between the /r/ and the preceding consonant in the modern version of words of type (5b). There is historical evidence for this *u*-insertion from the late thirteenth century and

[5] Assimilation to stem-final /l, n/ only happened in Old Icelandic if these consonants were preceded by long vowels, i.e. Old Icelandic diphthongs and vowels that are standardly represented by accented vowel symbols in Old Icelandic orthography, cf. *stól-**l*** 'chair' vs. *dal-r* 'valley', *fín-**n*** 'fine' vs. *lin-r* 'soft, limp', *heil-**l*** 'whole' vs. *hol-r* 'hollow'.

onwards (see e.g. **kristinsson1992** and references cited there) and many linguists have argued that *u*-epenthesis still a productive phonological process in Modern Icelandic (e.g. **anderson1969a**; **anderson1969b**; **oresnik1972**; **roegnvaldsson1981**; **kiparsky1984**).[6] This implies that speakers distinguish between a *-ur*-ending where the underlying morpheme is #-r# and the /u/ is epenthetic (and does not trigger *u*-umlaut) and a *-ur*-ending where the /u/ is not epenthetic and the underlying morpheme is #-ur# (and the /u/ triggers *u*-umlaut). This contrast is illustrated in (6a) vs. (6b) (see also the examples in (1) and (4) above):

(6)	a.	#dal+r#	'valley'	NOM.SG.M →	***dal-ur*** by epenthesis	no umlaut
		#lat+r#	'lazy'	NOM.SG.M →	***lat-ur*** by epenthesis	no umlaut
	b.	#sag+ur#	'sagas'	NOM.PL.F →	***sög-ur***	*u*-umlaut
		#tal+ur#	'numbers'	NOM.PL.F →	***töl-ur***	*u*-umlaut

Thus the NOM.SG ending #-r#, which is both found in strong masculine nouns like *dalur* 'valley' and in the strong masculine form of adjectives like *latur* 'lazy', does not have the same properties as the NOM.PL ending #-ur# which is found in feminine nouns like *sögur* 'sagas' and *tölur* 'numbers'. Despite their surface similarities in certain environments, speakers can clearly distinguish these endings. A part of the reason must be that the NOM.SG.M ending #-r# only shows up as *-ur* in phonologically definable environments, i.e. the modern version of words with stems of type (5b), whereas the NOM.PL.F ending #-ur# is not so restricted and always shows up as *-ur*. This is illustrated in Table **??** (compare the examples in (5)).

Comparison of Table **??** and the Old Icelandic examples in (5) reveals a slight extension of *r*-deletion: The /r/ of the morphological ending #-r# is now deleted after /r/ (compare line d of the table to (5a)) and after all instances of /s/, not just /ss/ (compare line d of the table to (5c,d)). The *u*-epenthesis illustrated in line b of Table **??** is an innovation, of course. Otherwise the NOM.SG.M ending behaves in much the same way as in Old Icelandic. The different behavior of the morphemes compared in Table **??** can be seen as an argument for distinguishing them in the underlying form, e.g. for not analyzing the N.SG.M ending as #-ur#.

2.2 Some phonological properties of Modern Icelandic *u*-umlaut

In this section I will mention two sets of facts which show that *u*-umlaut still has certain properties in Modern Icelandic that are to be expected if it is a phonologically conditioned process.

[6] Orešnik later maintained that *u*-epenthesis could not be a synchronic rule in Modern Icelandic because of the existence of exceptional word forms like *klifr* 'climbing' (from the verb *klifra* 'climb'), *sötr* 'slurping' (from the verb *sötra* 'slurp'), *pukr* 'secretiveness' from the verb *pukra(st)* 'be secretive about', etc. (**oresnik1978** see also the discussion in **kjartansson1984**). In words of this kind one would have expected *u*-epenthesis to apply. The importance of these exceptions is not very clear since this is a very special class of words (all derived from verbs ending in *-ra*) and it is typically possible or even preferred to apply the epenthesis rule to these forms, giving *klif**u**r*, *söt**u**r*, *puk**u**r*, etc. For the sake of completeness it should be noted that the final *-r* in word forms like *sötr*, *pukr* has to be voiceless and this may be related to the fact that there are no syllabic consonants in Modern Icelandic, as stated above.

Table 1: Phonological realization of the inflectional endings #-r# and #-ur# in Modern Icelandic.

	type of stem	phonological realization of the NOM.SG.M ending #-r#	phonological realization of the NOM.PL.F ending #-ur#
a.	ending in a vowel	*-r* (*mó-r* 'peat', *há-r* 'high')	***-ur*** (*ló-ur* 'golden plovers')
b.	ending in a single consonant (but see c)	***-ur*** (*dal-ur* 'valley', *lat-ur* 'lazy')	***-ur*** (*sög-ur* 'sagas', *töl-ur* 'numbers')
c.	ending in a high vowel + /l,n/	assimilation (*stól-l* 'chair', *fín-n* 'fine')	***-ur*** (*spús-ur* 'wives', *súl-ur* 'columns', *dýn-ur* 'mattresses')
d.	ending in /s, r/ or consonant clusters ending in /l, n/ such as /gl, gn/	deletion (*ís* 'ice', *laus* 'loose', *foss* 'waterfall' *bjór* 'beer', *stór* 'big', *fugl* 'bird', *vagn* 'wagon')	***-ur*** (*ýs-ur* 'haddocks', *aus-ur* 'scoops', *hór-ur* 'whores', *ugl-ur* 'owls', *hrygn-ur* 'spawning fish', *byss-ur* 'guns')

First, if *u*-umlaut was morphologically conditioned and not phonologically, we would expect it to be restricted to certain morphological categories or parts of speech. It is not. It applies in the paradigms of nouns, adjectives and verbs when a /u/ follows in the inflectional ending (with the exception of the epenthetic /u/ already mentioned). This is illustrated in (7):

(7) a. *saga* 'saga', OBL.SG *s**ö**g-**u***, NOM/ACC.PL *s**ö**g-**u**r*, DAT.PL *s**ö**g-**u**m*
b. *snjall* 'smart', DAT.SG.M *snj**ö**ll-**u**m*, NOM.PL.WK *snj**ö**ll-**u***
c. *kalla* 'call', 1.PL.PRS *k**ö**ll-**u**m*, 3.PL.PST *k**ö**ll**u**ð-**u***

The so-called *i*-umlaut is very different in this respect. It is clearly not alive as a phonological rule anymore but its effects can still be observed in the modern language in certain morphologically definable environments. As a result we can find near-minimal pairs of word forms where *i*-umlaut has applied in one member but not the other although the phonological conditions seem identical. Some examples are given in (8):

(8) a. *háttur* 'mode', DAT.SG *hætt-i*/**hátt-i*, NOM.PL *hætt-ir*/**hátt-ir*
b. *sáttur* 'satisfied', NOM.SG.M.WK **sætt-i/sátt-i*, NOM.PL.M **sætt-ir/sátt-ir*

In (8a) we see examples of the paradigmatic alternation /á ~ æ/ (phonetically [au] ~ [ai] in the modern language, probably [a:] ~ [ɛ:] in Old Icelandic) originally caused by *i*-umlaut.

In the NOM.SG we have /á/ in the stem but in the DAT.SG the only acceptable form is *hætti* and the "non-umlauted" version **hátti* is unacceptable. Similarly, in the NOM.PL only *hættir* is acceptable and **háttir* is not. At a first glance we might think that an /i/ in the inflectional ending is still causing this "umlaut" but a comparison with the adjectival forms in (8b) indicates that this cannot be the case. Here the only acceptable weak NOM.SG.M form is *sátti* and not **sætti* and the only NOM.PL.M form is *sáttir* and not **sættir*. So the *i*-umlaut alternations in Modern Icelandic are clearly morphologically conditioned and not phonological anymore (see also **thrainsson2011** for further examples of this kind).

Second, recall that standard generative phonology formulations of *u*-umlaut in Icelandic of the kind illustrated in (3b) above state explicitly that /u/ only triggers umlaut of /a/ in the immediately preceding syllable. This is illustrated by examples like the following:

(9) a. *bakki* 'bank' DAT.PL *bökk-**u**m/*ba**a**kk-**u**m*
b. *akkeri* 'anchor' DAT.PL **ökker-**u**m/**a**kker-**u**m*

In (9a) the *u*-umlaut obligatorily applies to the root vowel /a/ in the immediately preceding syllable. In (9b), on the other hand, the /u/ in the (same) inflectional ending cannot apply to the root vowel /a/ because there is a syllable intervening. An interesting and much discussed case, e.g. by Anderson in several of the publications cited above, involves trisyllabic words with two instances of /a/ in the stem. Consider the examples in (10):

(10) a. *k**a**ll**a*** 'call'
1.SG.PST *kalla-ð-i*, 1.PL.PST **k**a**llö-ð-um/köll**u**-ð-um/*k**a**ll**u**-ð- um/*kölla-ð-**u**m*
b. *b**a**n**a**n-i* 'banana'
DAT.PL *b**a**nön-um/bön**u**n-um/*b**a**n**u**n-um/*bön**a**n-**u**m*

Consider first the conceivable 1.PL.PST forms of the verb *kalla* 'call'. Based on the formulation (3b) of the *u*-umlaut rule, one might have expected the form **kallöðum*, where the /u/ in the inflectional ending triggers *u*-umlaut of the /a/ in the preceding syllable. This is not an acceptable form, however. The reason is that in forms of this sort a "weakening" of unstressed /ö/ to /u/ is obligatory. This weakening is found in in many words, e.g. the plural of the word *hérað* 'district', plural *héröð* or (preferred) *hér**u**ð, meðal* 'medicine', plural *meðöl* or (preferred) *með**u**l*. It is not always obligatory but it seems that in the past tense of verbs of this sort it is. But once the (umlauted) /ö/ in **kallöðum* has been weakened to /u/ it obligatorily triggers *u*-umlaut of the preceding /a/ so *köll**u**ðum* is acceptable but **k**a**ll**u**ðum* is not. Finally, the form **köll**a**ðum* is not acceptable either, since there *u*-umlaut would be applied across an intervening syllable, which is not possible, as we have seen (cf. (9b)). The *u*-umlaut works in a similar fashion in the word *banani*, except that here the weakening of the second (and unstressed) syllable from /ö/ to /u/ is not obligatory. Hence *banön**u**m* is an acceptable form, with the /u/ in the final syllable triggering *u*-umlaut of the preceding /a/ to /ö/. But if this /ö/ is further weakened to

/u/, then *u*-umlaut of the first /a/ is obligatory and ***bönunum*** is an acceptable form but **banunum* is not.[7] As predicted by the formulation of the *u*-umlaut rule in (3b) a form like **bönanum* is unacceptable because there the *u*-umlaut would have applied across an intervening syllable. Facts of this sort have been interpreted as showing that *u*-umlaut in Modern Icelandic is of a phonological nature since it depends on syllabic structure (no syllables can intervene between the umlaut trigger and the target) and it can be applied iteratively (a /u/ which itself is derived by *u*-umlaut and subsequent independently needed weakening can trigger *u*-umlaut).

3 The conditions for *u*-umlaut in Modern Faroese

3.1 *u*-umlaut and the Modern Faroese vowel system

Modern Faroese has preserved some *u*-umlaut-like vowel alternations. A couple of examples are given in (11) (see also **thrainsson2012**):

(11) *dag-ur* 'day', DAT.PL *døg-**u**m*; *spak-ur* 'calm', NOM.PL.WK *spøk-**u***

At first glance, these alternation seem very similar to the Icelandic ones described in the preceding sections. But while the *u*-umlaut alternations are arguably phonologically (or phonetically) natural in Modern Icelandic (see the diagram in (2b) and the formulation in (3b)), it will be claimed below that this is not the case in Faroese. To demonstrate this, it is necessary to look closely at the Faroese vowel system. Consider first the following schematic representation of Faroese *u*-umlaut of the type just illustrated, where the alleged umlaut trigger is encircled (cf. **thrainsson2011 thrainsson2012** compare **arnason2011**):[8]

(12) *u*-umlaut in Modern Faroese and the system of monophthongs:

	[-back]		[+back]	
	[-round]	[+round]	[-round]	[+round]
[+high]	i	y		ⓤ [u:/ʊ]
	e	ø [ø:/œ]		o
[+low]	æ [ɛa:/a]		a	ɔ

Something like (13) would seem to be a possible formulation of a process of this kind in traditional generative phonology terms (compare (3b)):

[7] It is sometimes claimed that ***bönönum*** is also an acceptable form for some speakers. If this is so, it is possible that the /ö/ in the next-to-last syllable triggers *u*-umlaut (i.e. *ö*-umlaut!) of the /a/ in the first syllable. That would simply mean that the feature [–low] in the definition of the environment of the *u*-umlaut in (3b) would be omitted. But since there are no derivational (nor inflectional) morphemes containing an underlying /ö/ (i.e. an /ö/ that cannot have been derived by *u*-umlaut), this proposal cannot be tested independently of the iterative rule application, as pointed out by a reviewer.

[8] Vowel length is predictable in Faroese, as it is in Icelandic: Vowels are long in stressed open syllables, otherwise short. As illustrated in the brackets in (12), there is often a considerable difference in the phonetic realization of the long and short variants. This will be illustrated below. — It should be noted that **arnason2011** assumes a different analysis of Faroese monophthongs.

(13) Possible phonological formulation of *u*-umlaut in Modern Faroese:

$$/æ/ \rightarrow \begin{bmatrix} \textbf{+round} \\ \textbf{-low} \end{bmatrix} / \_C_0V \begin{bmatrix} \textbf{+round} \\ +back \\ \textbf{-low} \end{bmatrix}$$

Presented this way, *u*-umlaut in Faroese looks like a fairly natural assimilation rule at a first glance.[9] But the facts are somewhat more complicated.

First, the alleged trigger /u/ is not too stable in Modern Faroese. The reason is that unstressed /i,u/ are not distinguished in all Faroese dialects. In some dialects they merge into an [ɪ]-like sound, in others into an [ʊ]-like sound but some dialects distinguish them as [ɪ] and [ʊ] (see **thrainsson2012** and references cited there). This situation has clearly added to the phonological opacity of *u*-umlaut alternations for speakers acquiring Faroese.

Second, the target of the *u*-umlaut in Faroese is arguably a "moving" one. As indicated in (12), the umlaut affects the phoneme represented there as /æ/. As the orthography suggests, it is a descendant of Old Norse /a/ in words like *da**a**gur*, *sp**a**kur* (see (11)). It is realized phonetically as [ɛa:] when long and [a] when short, as shown in (12), cf. *spakur* [spɛa:^{h}kʊɪ] 'calm', SG.N *spakt* [spakt] (see **thrainsson2012** passim). But in the history of Faroese Old Norse /a/ [a] and /æ/ [ɛ:] merged so the phoneme represented here as /æ/ can also be a descendant of Old Norse /æ/ and then it is represented in the spelling as 'æ', cf. *trælur* [t^{h}ɪɛa:lʊɪ] 'slave', *æða* [ɛa:va] 'eider duck'. Words written with 'æ' show the same alternation between long [ɛa:] and short [a] as demonstrated for *spakur* and *spakt* above (e.g. *vænur* [vɛa:nʊɪ] 'beautiful' SG.M vs. *vænt* [vaɳt], cf. Thráinsson et al. 2012, p. 34). Yet it seems that *u*-umlaut is rarely if ever found in the 'æ'-words. Thus the DAT.PL of *trælur* is *træl**u**m* and not **tr**ø**l**u**m* (compare DAT.PL *d**ø**l**u**m* of *dalur* 'valley') and although the words *æða* 'eider duck' and *aða* '(big) mussel' sound the same, i.e. as [ɛa:va], the DAT.PL of the former has to be *æðum* [ɛa:vʊn] and ***ø**vum* [ø:vʊn] can only be DAT.PL of *aða*.[10]

To further complicate matters, the development of Old Norse /a/ in Faroese has left "room" for a "regular /a/" in the Faroese vowel system, as shown in the diagram in (12). It occurs in loanwords and is realized as [a:] when long and [a] when short, cf. *Japan* ['ja:^{h}pan], *japanskur* [ja'p^{h}anskʊɪ] 'Japanese'.[11] It does not seem that this vowel ever undergoes *u*-umlaut in Faroese (for further discussion see section 4).

[9] A reviewer suggests, however, that a process changing rounding and height as formulated for Faroese in (13), might be less natural from the point of view of acoustic phonetics than a process changing rounding and backness the way the *u*-umlaut rule in Modern Icelandic does according to (3): The former affects both F1 (for the height difference) and F2 (for rounding) whereas the latter affects F2 in opposite directions (raising it for fronting but lowering it for rounding). Thus the Modern Icelandic *u*-umlaut rule would "generate more similar input-output mappings", which may be preferred to less similar ones.

[10] A reviewer points out that the fact that *u*-umlaut does not apply do 'æ'-words in Faroese suggests that "*u*-umlaut had already taken on a morphological character before /a/ and /æ/ merged." But since there are no written records of Faroese from 1400–1800, the historical development of the language is very murky.

[11] In the noun *Japan* the stress falls on the first syllable, in the adjective *japanskur* it falls on the second one as indicated. Hence the quantity alternation in the first vowel.

Finally, there is no *u*-epenthesis in Modern Faroese to "explain away" apparent exceptions to *u*-umlaut as will be shown in the next section.

3.2 The lack of *u*-epenthesis in Modern Faroese

Now recall that the most obvious surface exception to *u*-umlaut in Modern Icelandic is due to the *u*-epenthesis described above. This rule creates *-ur*-endings that do not trigger *u*-umlaut. It was argued that this epenthesis rule is still productive in Icelandic, witness the fact that it only applies in phonologically definable environments. Hence there is a clear distributional difference between *-ur*-endings produced by the epenthesis rule (and not triggering *u*-umlaut) and *-ur*-endings where the /u/ is a part of the underlying form (and triggers umlaut). This is not the case in Faroese, where the ending *-ur* as a marker of the NOM.SG of strong masculine nouns and adjectives, with a /u/ that was historically inserted by epenthesis, has been generalized to all environments. Hence it has become distributionally indistinguishable from other *-ur*-endings. Table **??** compares the phonological realization of the NOM.SG.M #-r#-ending in Modern Icelandic to its Modern Faroese counterpart (see also **thrainsson2011**):

Table 2: Phonological realization of a strong NOM.SG.M-ending in Modern Icelandic and Modern Faroese.

	type of stem	phonological realization of a strong NOM.SG.M ending in Modern Icelandic	phonological realization of a strong NOM.SG.M ending in Modern Faroese
a.	ending in a vowel	*-r* (*mó-r* 'peat', *há-r* 'high')	***-ur*** (*mó-ur/mógv-ur* 'peat', *há-ur* 'high')
b.	ending in a single consonant (but see c)	***-ur*** (*dal-ur* 'valley', *lat-ur* 'lazy')	***-ur*** (*dal-ur* 'valley', *lat-ur* 'lazy')
c.	ending in a high vowel + /l,n/	assimilation (*stól-l* 'chair', *fín-n* 'fine')	***-ur*** (*stól-ur* 'chair', *fín-ur* 'fine')
d.	ending in /s, r/ or consonant clusters like /gl, gn/	deletion (*ís* 'ice', *laus* 'loose', *foss* 'waterfall', *stór* 'big', *fugl* 'bird', *vagn* 'wagon')	***-ur*** (*ís-ur* 'ice', *leys-ur* 'loose', *foss-ur* 'waterfall', *stór-ur* 'big', *fugl-ur* 'bird', *vagn-ur* 'wagon')

This has clearly made the *u*-umlaut rule in Faroese more opaque since now the non-umlauting and umlauting *ur*-endings occur in the same phonological environments. It seems very likely that this has contributed to the death of *u*-umlaut as a phonological process in Faroese.

4 Testing the predictions

In the preceding discussions I have described /a ~ ö/ alternations in Modern Icelandic and their Modern Faroese counterparts. I have argued that the Icelandic alternations are still governed by a synchronic phonological process. Although these alternations are still found in Modern Faroese, I have argued that they cannot be governed by a phonological rule. Instead they must be morphologically governed or analogical. This analysis makes several testable predictions (see **thrainsson2011**).

First, we do not a priori expect phonologically conditioned alternations to be restricted to particular morphological categories whereas morphologically conditioned alternations obviously are, by definition. As we have already seen, the Icelandic *u*-umlaut occurs in the inflectional paradigms of nouns, adjectives and verbs and in various grammatical categories (cases, numbers, tenses, persons …). Its Faroese counterpart behaves differently. It is found in the inflectional paradigms of nouns and adjectives, as we have seen (cf. (11)), but not in the past tense forms of verbs, where it would be expected on phonological grounds. Thus we have *við kölluðum* in Icelandic vs. *vit kallaðu* in Faroese for 1.PL.PST 'we called', and *við frömdum* vs. *vit framdu* in Faroese for 1.PL.PST 'we did, made'.

Second, a phonological rule should not allow analogical extensions to forms that do not fit its structural conditions. Such extensions are not found for Icelandic *u*-umlaut but in Faroese they are very common. Thus the /ø/ of the oblique cases *søg**u*** 'saga' has been analogically extended to the NOM.SG form *søga* and many other words of a similar type. The corresponding form **söga* is unacceptable in Icelandic.[12]

Third, a phonologically conditioned rule should apply whenever its structural conditions are met. Thus we would not expect to find inflectional forms in Icelandic where *u*-umlaut fails to apply in an appropriate environment. Such examples are very common in Faroese, on the other hand. Thus the DAT.PL of the noun *rakstur* 'shave' in Faroese is *r**a**kstr**u**m* and not the expected **r**ø**kstr**u**m*, the DAT.PL of *spakur* 'calm' can either be *spø**k**um* or *spa**k**um*, etc. (see **thrainsson2012**). Corresponding unumlauted forms are unacceptable in Icelandic.

Fourth, there is evidence for "iterative" application of *u*-umlaut in Icelandic, with one application of the *u*-umlaut rule feeding another. This was discussed above (second part of section 2.2) in connection with forms like 1.PL.PST *köll**u**ð**u**m* '(we) called' and DAT.PL *bön**u**n**u**m* 'bananas'. No such evidence is found in Faroese, where the corresponding forms are *k**a**ll**a**ðum* and *b**a**n**a**num*.[13]

Finally, Icelandic *u*-umlaut is so productive that it is naturally applied in loanwords,

[12] As a reviewer reminds me, the Icelandic neologism for *computer* is interesting in this connection. It was supposed to be *tölva* (related to the word *tala* 'number' — this was when computers were mainly used for computing) in NOM.SG, oblique singular cases *tölvu*. In Proto-Nordic time /v/ could trigger umlaut of /a/ to /ǫ/ so we have Old Norse words like *vǫlva* 'sooth-sayer, witch'. But since /v/ is not a trigger of umlaut in Modern Icelandic (witness loanwords like *salvi* 'salve, cream'), speakers tend to use the form *talva* for NOM.SG, thus in a way undoing the underlying /ö/ in the nominative as if they are "assuming" that the /ö/ in the oblique cases is derived by a synchronic *u*-umlaut from /a/, as in words like *saga* 'saga', oblique *sögu* (for some discussion see **thrainsson1982**).

[13] The latter form may be related to the fact that *banan* 'banana' is a loanword and contains the vowel /a/ (long variant [a:]) and not /æ/, cf. the discussion in section 3.1. See also the next paragraph.

as we have seen. This is not so in Faroese. Thus the word *app* (for a small program) has been adopted into both languages. In Icelandic the DAT.PL has to be ***öpp**u**m*** whereas the natural form is ***app**u**m*** in Faroese. This can easily be verified by searching for the word combinations *með öppum* and *við appum* 'with apps' on Google. For the first variant one finds a number of Icelandic hits, for the second Faroese ones.

The general conclusion, then, is that *u*-umlaut in Modern Icelandic has a number of properties that are to be expected if it is a phonological process but none of the properties one might expect of morphologically conditioned or analogical alternations.

5 Concluding remarks

While it has often been argued that phonology need not be "natural" (see e.g. **anderson1981**), there must obviously be limits to the "unnaturalness" and opacity of phonological processes. Once they become too unnatural and opaque, they can no longer be acquired as such and the phonological alternations originally created by them will be relegated to morphology. Then their productivity will be limited and it will at best survive to some extent by analogy, but analogical processes are known to be irregular and unpredictable. The fate of *i*-umlaut in Icelandic is a case in point, as described above (see the discussion of the examples in (8)). But whereas we do not have detailed information about how *i*-umlaut died as a phonological process, comparison of the development of *u*-umlaut in Icelandic and Faroese sheds an interesting light on how a phonological rule can die and how it can survive despite changing conditions.

Acknowledgements

Many thanks to Steve Anderson for introducing me to the wonders of synchronic *u*-umlaut way back when. Thanks are also due to the editors and to two anonymous reviewers for help, useful comments, suggestions and corrections.

Chapter C

Saussure's Dilemma: *Parole* and its potential

Alan Timberlake
Columbia University

Saussure's account of the transformation of Latin to French stress leads to the unintended conclusion that *parole* has a life of its own: parole persists even after it is no longer dictated by *langue*; *parole* can prevent change or, conversely, presage potential change. Saussure's example is paralleled by intrusive *r* (*Cuba[r]against your friends*, not **day[r]and*) and the competition of *napron* and its near twin, and eventual successor, *apron*. *Parole* lives.

1 Saussure's Dilemma

It seems only fitting to begin this tribute to Steve Anderson, friend and erstwhile UCLA colleague, historian of linguistics and confirmed *helvétophile*, with Ferdinand de Saussure and his discussion of the history of stress from Latin to French (Chapter III.4–6 of the *Course*). That history presents a dilemma for Saussure's separation of synchrony from diachrony and linguistic activity (*parole*) from system (*système*, *langue*).

Here is Saussure's account of the transition from (ante)penultimate stress in Latin to final stress in French:

> In French, the accent always falls on the last syllable unless this syllable contains a mute *e* (*ə*). This is a synchronic fact, a relation between the whole set of French words and accent. What is its source? A previous state. Latin had a different and more complicated system of accentuation: the accent was on the penultimate syllable when the latter was long; when short, the accent fell back on the antepenult (cf. *amícus, ánĭma*). The Latin law suggests relations that are in no way analogous to the French law. Doubtless the accent is the same in the sense that it remained in the same position; in French words it always falls on the syllable that had it in Latin: *amícum* → *ami, ánimam* →*âme*. But the two formulas are different for the two moments because the forms of the words changed. We know that everything after the accent either disappeared or was reduced to mute *e*. As a result of the alteration of the word, the position of the accent with respect to the whole was

> no longer the same; subsequently speakers, conscious of the new relation, instinctively put the accent on the last syllable, even in borrowed words introduced in their written forms (*facile, consul, ticket, burgrave*, etc.). Speakers obviously did not try to change systems, to apply a new formula, since in words like *amícum* → *ami*, the accent always remained on the same syllable; but speakers changed the position of the accent without having a hand in it. A law of accentuation, like everything that pertains to the linguistic system, is an arrangement of terms, a fortuitous and involuntary result of evolution. **(DESb)**

To explain modern French final stress, Saussure goes back to its source in Latin, or rather to a stage of Romance subsequent to classical Latin but much earlier than modern French. He first defines the stress rules for French (final) and Latin ((ante)penultimate, depending on the quantity of the penult). That done, Saussure mentions in passing that the position of stress in French preserves the original position of stress from Latin, and illustrates the claim with examples of the two subcases of Latin stress, on a long penultimate (*amícum* → *ami*) and on the antepenultimate when the penult is short (*ánimam* →*âme*). It is worth mentioning that his formulation "the accent is the same in the sense that it remained in the same position" is not original; it repeats a standard observation from French philology in the middle of the nineteenth century. Thus in 1862 Gaston Paris (and seven other scholars he mentions, p. 11) stated the observation that stress falls on the same syllable in French as it did in Latin; for that phenomenon Paris in particular uses the apt term "persistence" (*persistance*) (**PA**). Saussure does not use that term, but his formulation echoes this earlier tradition. Saussure then mentions the familiar fact that syllables after the syllable with the "persistent" stress are subject to reduction. At this point one might think he is preparing to explain how final stress in French arose, for example, perhaps by generalizing the word-final stress of words like *amícum* → *ami* that had undergone apocope. Saussure does not go in that direction. Instead, he declines to give a linguistic explanation for the development of final stress and places the burden on speakers acting "instinctively" and then on the vague assertion that "a diachronic fact was interposed." So although Saussure at the outset seemed prepared to explain modern French stress in terms of its source in Latin, the source is not relevant to Saussure's interpretation. He ends with a summary blaming the chaotic nature of change: "A law of accentuation, like everything that pertains to the linguistic system, is an arrangement of terms, a fortuitous and involuntary result of evolution."

Saussure seemed to recognize that the Latin rule of (ante)penultimate stress was lost as a rule. Further, it cannot have been a rule of the *système*; had it been rule of the *système*, it would have been maintained. Saussure's response was this:

> The synchronic law is general but not imperative. Doubtless it is imposed on individuals by the weight of collective usage…, but here I do not have in mind an obligation on the part of speakers. I mean that in language no force guarantees the maintenance of a regularity when established on some point. Being a simple expression of an existing arrangement, the synchronic law reports a state of affairs; it is like a law that states that trees in a certain orchard are arranged in the shape

> of a quincunx. And the arrangement that the law defines is precarious precisely because it is not imperative. Nothing is more regular than the synchronic law that governs Latin accentuation…; but the accentual rule did not resist the forces of alteration and gave way to a new law, the one of French… In short, if one speaks of law in synchrony, it is in the sense of an arrangement, a principle of regularity. (**DESb**)

This analysis, developed in connection with a discussion of six facts of Indo-European, is here applied to the Latin stress rule, which was downgraded to a descriptive observation about behavior maintained by social convention, analogous to the rule stating that "the trees in a certain orchard are arranged in the shape of a quincunx."

The phenomenon of Latin stress, with its properties of persistence, precariousness, and regularity presented a dilemma for Saussure. The phenomenon of persistence implies that usage (*parole*) has a life of its own; *parole* is maintained as parole, as habit, transmitted by imitation from one generation to the next. To invoke a metaphor, substituting "language" for "body" in Newton's first law, we could say: "a language at rest remains at rest unless it is acted on by an external force." Usage, such as the Latin stress rule, can exhibit coherent patterns (such as the elegant parallelism of length of the penultimate and antepenultimate position). Moreover, the patterns of *parole* are capable of defining the conditions for change (such as the "persistent" location of stress after Latin which conditions post-tonic apocope), and in this way patterns of *parole* act like elements of *système*.

Saussure's dilemma was that the more he insisted on the dominant, special, pure exclusionary status of *système*, the more ethereal and abstract system became, and that had the paradoxical effect of elevating *parole* to be the object of investigation.

2 *Law and Order* and Sandhi Doublets

To document the fate of /r/ in weak position (after a vowel, not before a vowel), **KUR** divided the eastern seaboard into four discontinuous zones, two northern and two southern, in which /r/ becomes [ə̯] in weak position: northern – including New England (Connecticut River east) and metropolitan New York; southern – including Upper South (Virginia, into northern North Carolina) and Lower South (South Carolina, Georgia). The four zones are separated by transitional belts which retain some form of *r* (presumably American [ɹ] or "velarized constricted" *r* [ɚ]). The largest of the transitional belts is Pennsylvania, from which rhoticism spread to Midwest and Midland dialects.

The discussion here focuses on the northern zones, which **KUR** treated as a single zone. As shown in Table 1, in the north /r/ is reflected in weak position as [ə̯] after mid and high vowels ([i, u, e, o]). After low vowels ([a/ɐ, ɒ/ɔ] and here [ə] as well) what must have been the earlier reflex [ə̯] was lost or absorbed by the vowel.

Like northern dialects, southern dialects also absorbed [ə̯] after low-mid [a/ɐ, ɒ/ɔ, ə]. Furthermore: "In Southern folk speech, /ə̯/ is often lost, *door*, *four*, *poor* /doə̯, foə̯, poə̯/

Table 1: Postvocalic Rhotic Reflexes, North

high/mid V	low V
[ir], [ur], [er], [or] >	[ar/ɐr], [ɒr/ɔr], [ɘr] >
[iə̯], [uə̯], [eə̯], [oə̯]	[aə̯], [ɒə̯/ɔə̯], [ɘ] >
ear [iə̯]	[a/ɐ], [ɒ/ɔ], [ɘ]
poor [puə̯]	*far* [fa/fɐ]
care [keə̯]	*for* [fɒ/fɔ]
four [foə̯]	*father* [faðə]

thus becoming /do, fo, po/" (**KUR**).[1]

Kurath and McDavid devoted special attention to sandhi contexts – contexts in which the word-final vowel which once had /r/ is used in a phrase with a following word. Then the word with original /r/ can be said to have two "sandhi doublets," depending on whether the second word begins with a consonant (when the original /r/ would have been in weak position) or with a vowel (when the original /r/ would have been prevocalic). They stated: "…*ear, poor, care, four* have… the positional allomorphs /iə̯ ~ iə̯r, puə̯ ~ puə̯r, kæə̯ (keə̯) ~ kæə̯r (keə̯r), foə̯ (fɔə̯) ~ foə̯r (fɔə̯r)/, and *car, for* (stressed), *father* the allomorphs /ka (kɐ) ~ kar (kɐr), fɒ (fə) ~ fɒr (fɔr), fað ~ faðər/." Note the difference between non-low vowels, in which the reflex is [Və̯(r)], and low vowels, in which the reflex [V(r)] lacks [ə̯], since [ə̯] had been absorbed by the preceding low vowel. The idea of calling these doublets (and they provide a notation for doublets) suggests a model of the lexicon in which a lexical item is composed of multiple subunits, which could be written as an ordered pair such as {[iə̯]/sandhi before consonant; [iə̯r] / sandhi before vowel}. One might, for example, write a doublet for the noun *Cuba* as pronounced by John F. Kennedy in his speech "in the Cuban Missile Crisis" generally as [kubə], as in *and then shall Cuba[ə] be welcomed* (16:07), but [kubər] in phrases such as *Soviet assistance to Cuba[r] and I quote* (4:32) and *turned Cuba[r] against your friends* (15:05).[2]

Examples constructed in the spirit of **KUR** are given in Table 2, top.

It is worth drawing attention to the fact that prevocalic sandhi examples with non-low vowels have the sequence [ə̯r] (p. 171b); as in [iə̯r, puə̯r, kæə̯r (keə̯r), foə̯r (fɔə̯r)] from the list above. The sandhi sequence [Və̯rV] has in effect two segments – [ə̯] and [r] – which reflect earlier /r/. Both cannot be original. There must have been an antecedent stage of [*irV, *purV, *kærV (*kerV), *forV (*fɔrV)] in sandhi position before a vowel. The [ə̯] we see now in [iə̯r], etc., had to have been introduced by analogy from other forms to the sandhi forms before vowel.

Analogy is relevant to history in another respect (**SOS**). Not uncommonly, words that ended originally in low vowels without /r/ acquired a non-etymological, or "intrusive,"

[1] The lexeme *poor* is treated once as having a mid vowel (171a) and otherwise as a high vowel (170b, 171b, 172a, 172b). (One, also 171b, is ambiguous.)

[2] http://www.historyplace.com/speeches/jfk-cuban.htm.

Table 2: Sandhi /r/ and Intrusive /r/, North

context	high, mid V	low V
sandhi	[iɻ], [uɻ], [eɻ], [oɻ]	[a/ɐ], [ɒ/ɔ], [ɘ]
[V(ɻ)r̂V]	*ear and* [iɻrænd]	*far and* [farænd]
	poor and [puɻrænd]	*for all* [fɒral]
	care and [keɻrænd]	*father and* [faðərænd]
	four and [foɻrænd]	
intrusive	[i], [u], [e], [o]	[a/ɐ], [ɒ/ɔ], [ɘ]
[Vr̂V]	*three and* *[θrirænd]	*ma and* [maræ̲nd]
	two and *[turænd]	*law and* [lɒræ̲nd]
	day and *[derænd]	*Martha and* [maθərænd]
	know it *[norɪt]	

/r/ in sandhi, as in the familiar *law and order* [lɒrəndɒdə] and other examples in Table 2. As **KUR** state,

> On the analogy of such doublets as *for* /fɒ (fɔ) ~ fɒr (fɔr)/, *car* /ka (kɐ) ~ kar (kɐr/, and *father* /faðə ~ faðər/, positional allomorphs ending in /r/ are often created in Eastern New England and Metropolitan New York for words that historically end in the vowels /ɒ ɔ, a ~ ɐ, ə/, as *law*, *ma*, *Martha.* Thus one hears *law and order* /lɒr ənd ɒdə, lɔr ənd ɔdə/, *ma and pa* /mar ən(d) pa, mɐ ən(d) pɐ/, *Martha and I* /maθər (mɐθər) ənd ai/.

The examples of intrusive /r/ just cited involved only words which end in a low vowel – that is, they have the same vocalism in the non-sandhi environments as words that originally ended in /r/ but which absorbed the [ɻ] reflex of /r/; thus *law* /lɒ (lɔ)/ has the same vocalism as originally rhotic words like *for* /fɒ (fɔ)/. But words like *three*, *two*, *day*, *know*, which end in mid and high vowels, differ. **KUR** state:

> It is worth noting that after the normally upgliding free vowels /i, u, e, o/, as in *three*, *two*, *day*, *know*, an analogical "intrusive" /r/ never occurs. The reason for this is clear: since /θri, tu, de, no/ do not end like the phrase-final /r/-less allomorphs of *ear*, *poor*, *care*, *four* /iɻ, puɻ, keɻ, foɻ/, the basis for creating allomorphs ending in /r/ is lacking.

Thus according to **KUR** the development of intrusive /r/ involves the comparison of stem shapes, for example [lɒ] with [fɒ], which are similar and permit analogy, as opposed to *three* [θri] with [iɻ], which are dissimilar and do not permit analogy.

This distribution is interesting. What determines whether analogical intrusive /r/ develops is an arbitrary division of vowels inherited from the previous history of derhoticism; that is to say, a distinction in vowels involved in the earlier history of reflexes of /r/ in weak position continues to have an effect on later developments. Thus *parole* has

the property of inertia (*persistance*), so that later changes (such as the analogical development of intrusive /r/) can be sensitive to properties of *parole* that persist. At the same time as *parole* is inertial and conservative, *parole* nevertheless carries with it the possibility of change. Thus original *r*-less words ending in low vowels have the potential to develop an intrusive sandhi /r/, as happened in northern dialects. Conversely, original *r*-full words had the potential to eliminate the second member of the "doublet" in which /r/ reappears in sandhi before a vowel; this is what happened in southern dialects (especially Upper South but even in the Lower South sandhi forms with /r/ are "only half as frequent as the variants without /r/" (**KUR**)).

Parole, then, is inertial but carries the potential for change. This example is similar to what Saussure said about Latin stress, that it remained on the syllable where it had always been – by convention, or memory, or inertia – but eventually the stress was repositioned.

It might be objected that it would be easy to state a rule inserting /r/ that is sensitive to vowel height; insertion would happen only in position after low vowels. But why low vowels? Low vowels are not universally more likely than other vowels to adopt a phonotactic sequence [VCV] that other vowels. Intrusive /r/ develops only after low vowels because it is only low vowels that offered a model for analogical extension, and that is a distribution that goes back to a prior change; the restriction to low vowels can only be understood by viewing it as the hangover from a previous stage. Moreover, it is not just any consonant that reappears; it is just the one sound /r/. The /r/ can participate in "intrusive" analogy because the /r/, and only the /r/, was carried over from earlier history. The fact that /r/ is involved in analogy at all is a further instance of persistence of *parole*.

3 *Watergate* and its Ilk

Against this background I want to discuss how innovations can arise directly out of speech. The word *Watergate* and its derivatives can serve as an illustration. As is familiar, *Watergate* is the name of a complex of five buildings built in Washington, D.C., over the period 1963–1971. An office building in this complex was used by the Democratic National Committee as headquarters leading up to the 1972 election. The Democratic offices suffered a break-in, for which staff members of the Republican administration were later discovered to be responsible. The break-in triggered an embarrassing scandal and, because of the attempt to cover up the original crime, led to the resignation of President Richard Nixon.

A modification of the name for this location keeps being applied to more events, which, like the original Watergate, include at least two events, layers of agency, times, places. The core is the pairing of two events: first, an event carried out in secret and, second, the fallout, including the embarrassment caused by the event for the participants and perhaps further developments (cover-up, disclosure). The whole scenario is a rich instance of the familiar trope of metonymy, which points to one event – here, the original transgression – which can invoke associated events (here, the fall-out) and the constituents

of those events (locus, agents, patients). The name for this complex of events and constituents, which occurred in 1972, is of course *Watergate* – the name for the place is applied to the whole package of events, by the trope of *pars (locus) pro toto* (complex of events – crime, scandal, cover-up, further fall-out). The semantic operations involved in *Watergate* are familiar, banal tropes.

Event complexes similar to the original *Watergate* scenario can be named by the new compound {*x+gate*}, where {*gate*} refers to the existence of a scandalous event (and fall-out) and *x* refers to a focus – a constituent that is central to the events – such as the agent (*Billygate*) or causal entity (*nannygate*) or the patient (*contragate*).

The morphological structure and semantics of the new compound {*x* 'focus'+*gate* 'event(s) leading to scandal'} seem clear, and it seems clear that the compound is related to the origin *Watergate*. How? Given the apparent overlap of {*gate*} in both, one might imagine that the word *Watergate* was decomposed into two morphemes, {*water*} and {*gate*}, and that reanalysis provided the model for neologisms. But this cannot be: by itself "*water*" does not mean anything in this context; it is not the focus. And for that matter, *gate* doesn't mean scandal here in the compound *Watergate*. In the original word *Watergate*, there is no division; *Watergate* is the name for the complex as a whole, not for any of its constituents.

And yet *Watergate* was self-evidently the source for the formula {*x+gate*} and novel applications of the formula. What this means is that "Watergate" – the name for a whole complex of agents and events – allowed speakers to imagine a new structure {*x+gate*} whose semantics give overall semantics analogous to the meaning of *Watergate* (secret event and subsequent scandal, specific place or agents, etc.) but in which the event complex is broken into two constituents; one of them, {*x*}, refers to the focus of events, and the other part, {*gate*}, establishes the existence of a secret event and its attendant scandal involving the focus {*x*}, whereas in the source *Watergate*, the whole included all the components.

Two aspects of this change are significant. First, the new structure is motivated by the inherited word, but it is not a copy; it cannot be generated by a proportional analogy. Instead, what the example shows is that speech has the potential of providing motivation for creating new speech directly. To say it another way, speech is not just speech; speech invites modal possibilities. The second point is that the source here really is speech that actually occurred in real time: *Watergate* started as a single event complex that occurred at some time in history; it did not start as a pattern. That is, a singular event and the accompanying speech give rise to an innovation; speech creates speech. This new {*x+gate*} is a virtual structure which might exist indefinitely. We cannot verify its existence until it is acted on. Therein is a property of language that has eluded description: the fact that speech happens, that activity matters, it happens when a novel formation is used, and it happens to the extent that neologisms are created and used in speech.

This example, then, suggests a more active role for *parole* (performance, speech) than has usually been assumed. In this instance actual speech from a very specific historical time (1972) provided the model and created the potential for new speech, and that is what resulted. It is worth stating that speech is not just blind activity; speech comes with

implicit patterns, whether firmly established or – as in this case – potential, possible, modal speech.

It could be mentioned that this formation, along with similar neologisms motivated by *alcoholic*, have distinctive stylistic overtones and spheres of usage – in the personal sphere, gentle mockery (*shopaholic*, *chocoholic*) and not-so-gentle journalistic irony for the former (*Camillagate*). The News History Gallery at the Newseum (http://www.newseum.org/) in Washington, D.C., devoted to the history of journalism, has an exhibit called "The 'Gate' Syndrome," illustrated by five examples, starting with *Koreagate* (1976). [3]

4 *(N)apron* as Dynamic Doublet

A somewhat similar change is the change from *napron* to *apron* in Middle English. As is familiar, a dozen or so nouns which had once begun with an initial consonant *n* lost the *n* and came to begin with the vowel of the first syllable. According to the standard analysis, this happened because when such nouns were used with the indefinite article *a(n)*, a sequence of [*anV*] would result, and then it is unclear whether the intervocalic [n] belongs to the stem of the noun or to the article. The ambiguity opened up the possibility that the [n] could be attributed to the article and the noun could be reanalyzed as beginning with a vowel. Subsequently the stem shape without the vowel could be extended to all contexts; thus {a+napron} > [anapron] was analyzed as {an+apron} > [anapron], leading to the use of {apron} elsewhere. As is well known, the converse also occurs, where nouns beginning with an initial vowel (*an ewt*) acquired an initial *n* from the indefinite article (> *a newt*). It is not clear why the change of metanalysis should be able to go in either direction.

This standard analysis discusses only the end-points of this change – prior to metanalysis, after metanalysis – but does not describe how the change progressed. To get a sense of how this change actually proceeded, I attempted to trace the history of spellings *(n)apron* in Middle English with an eye to variation in the choice of the word form in different contexts. The task was rather more challenging than I had expected. The word *(n)apron* is quite specific. It occurs infrequently, primarily in wills and inventories of good to be bequeathed. (And also, as will be noted below, in a description of the rules of the household of Edward IV.) The item is mentioned only in a minority of the wills or inventories available, and usually when the deceased is a woman. For example, the extensive *Wills and Inventories of Bury St. Edmunds* has approximately 150 printed pages of wills from the beginning of the fifteenth century (one will from 1370, then 1418, etc.) to the late sixteenth century (1570), and has no instances of the word in either variant, *napron* or *apron*. That, despite instances such as the will of one Agas Herte (a. 1522, pp. 114–18), who bequeathed about 50 distinct household objects to her son, including "*ij tabyll clothes, vj napkyns, iiij pleyne and to of diap, a salte saler of pewter…*" and about the same number to her daughter, including "*ij tabell clothes, vi napkyns, iiij pleyn and ij of*

[3] Arnold Zwicky calls {gate} a "libfix" – "lib" in the sense of "liberated" – which captures the idea that a mental operation extracts a new affix (https://arnoldzwicky.org/2010/01/23/libfixes/). The author wishes to thank the editors for this and many other valuable and droll comments and corrections.

diap, and a pleyn towel…" Among all the items she bequeathed, including the items made of cloth just mentioned, no *(n)apron* was mentioned. This might because this household, and other households as well, did not use *(n)aprons*; it might be they were considered too insignificant to be mentioned in bequests (though towels and napkins and sheets are recorded regularly). In any event, the frequency with which *(n)apron* appears is modest. In short, it has proven difficult to find document sets in which *(n)apron* is mentioned multiple times; examples are isolated. To maximize the range of texts examined, I used Hathitrust/Google scans subjected to OCR. I searched for both *napron* and *apron*, both singular and plural, in variant spellings.

We can first take a quick look at chronology, using a ledger (*Fabric rolls*) kept by the York Minster which recorded miscellaneous expenses annually. The entries are written in Latin, though names for some items specific to the contemporary realia appear in English. Half a dozen times the rolls record payment for the costs of masonry, both for wages and equipment – aprons and gloves for masons (called "setters"). The earliest record from 1371 surprisingly has *n*-less *aprons* (*ij aprons et cirotecis* 'two aprons and gloves', 1371). Then at the beginning of the fifteenth century come two instances of *naprons: In remuneracione data cementariis vocatis setters ad parietes* ***cum naprons*** *et cirotecis, per annum 9s. 10d.* 'as compensation given to the masons known as setters at the wall with aprons and gloves, annually 9s. 10d.' (1404); *In ij pellibus emptis et datis eisdem* ***pro naporons***, 'two hides were bought and given to them to serve as aprons' (1423). At the end of the fifteenth century there are two examples of Latin *limas* (***duobus limatibus,*** 1497–98; ***Pro ij limatibus***, 1499), and shortly thereafter, *aprons* (***pro ij le aprons*** *de correo pro les setters per spacium ij mensium, 12d.* 'for two aprons of hide for the setters for the period of two months, 12 shillings' (1504). The use of *aprons* in 1371 seems anomalously early (could it be an error in transcribing the text?). This anomaly aside, the examples suggest a chronology: *napron* was used in the fifteenth century (1404, 1423) and shifted to *apron* the beginning of the fifteenth century (1504). Other texts suggest there was still some variation in the sixteenth century. By 1600 *apron* had taken over.

Against the background of generally skimpy attestation of *(n)apron* in the fifteenth and sixteenth century, there are two texts which offer enough examples to allow us to say something about usage. One is a single text, the so-called Liber Niger Domus Regis, which specifies the duties and compensation of the staff of King Edward IV's household in the last quarter of the fifteenth century (c. 1480). A modern edition compiles three manuscripts (discussion, **myers**). The oldest is a manuscript from the end of the fifteenth century, which served as the basis for the famous **XEDWARD** publication by the Society of Antiquaries (abbreviated "A"); however, text A is now defective, and it also appears that the 1790 edition took some liberties, so the printed 1790 edition cannot be trusted to represented the oldest text A. In the accompanying Table 3 I've cited the location of readings from the 1790 reading edition in ||. The next oldest is a sixteenth-century manuscript (preserved in the Public Records Office, the Exchequer, abbreviated E), from the era of Henry VIII, is similar to A but fuller. Third, the youngest of the three manuscripts, known as Harleian 642 (here H), is a seventeenth-century copy made by Sir Simonds d'Ewers. In fact, differences recorded in footnotes by Myers in his edition

are minimal and affect the analysis here in only one respect, mentioned below.

There are basically two contexts (with one additional outlier). Examples repeat over the descriptions of many different servants. The twelve examples of *(n)apron* are given in abbreviated form Table 3.

Table 3: *(n)apron* in Liber Niger of Edward IV

	type	text [variants]
§55 |49.7d|	MOD	they have part of the $^{\alpha}$yeftes$^{\alpha}$ geuvn to the houshold… but ***none aprons*** [1790: $^{\alpha}$gyftes$^{\alpha}$]
§62 |52.28d|	DO	[takith] at euery of the iiij festes of the yere, ***naprons*** of the great spycery
§62 |52.28d|	DO	take ***naprons*** also at euerych of the iiij festes
§80 |71.6|	DO	etithe in the halle; taking for wages… and nyȝt lyuereye, ***naporons***, and parte of the generall giftes
§80	DO	taking for his wynter clothing chaunces, ***napors***, parte of the giftes generall [extended passage, absent in 1790]
§74 |61.26d|	1DO	Eche of them takethe… ***j napron*** of lynyn cloth of ij ellez [1790: ***a naperon***]
§77 |65.19d|	1DO	At euery of the iiij festes, ***j napron*** of j elle, price vjd. [1790: ***one napron*** of one elle]
§33 |36.10u|	PRP	[he takith]… ij elles of lynen clothe ***for aprons***, price the elle, xijd.
§77 |64.25d|	PRP	ij ellez of lynyn cloth… ***for naprons***
§77 |65.8d|	PRP	j elle of lynnyn clothe ***for naprons***
§77 |64.41|	PRP	j elle ***for naprons*** of lynyn cloth
§80 |71.26d|	ORO	$^{\alpha}$and for chaunces iiijs. viijd.$^{\alpha}$ of ***napors*** at euerye [H$^{\alpha\alpha}$; H ***aprons***]

§ = section in Myers, || page in 1790 edition

In one isolated instance, the noun is preceded by *none*; the nasal might have elicited the following *apron*. The other eleven tokens are split between two contexts. In six instances the noun is the direct object of the verb 'take' (listed as "DO"). The entities taken have already been formed into garments; all have *napron*. Within this group of six

examples, in two of these six, marked here as "1DO," the word *napron* is preceded in texts E and H by *j*, that is, Roman numeral 'one'. (The published 1790 version has the indefinite article *a napron* in one instance and the written word *one napron* in the other, rather than the numeral.) Both E and H use numerals consistently in discussions of compensation; prices of elles of linen cloth are cited with numerals, such as *j elle*. The numeral here must be original, thus *j napron* in both examples. I will return to these in a moment.

The second group of six examples involves the statement that servants receive compensation in the form of linen cloth which is supposed to be turned into *(n)aprons*, expressed by a preposition, usually *for*, once *of*. In this context the entity referred to as *(n)apron* does not yet exist; the noun has a future attributive sense: "the speaker wishes to assert something about whatever or whoever fits that description" of being (or becoming) an apron (**DON**). An example is: *at euerychе of the iiij festes of the yere, of the clerk of greete spycery, ij elles of lynen clothe for* ***aprons****, price the elle, xij*[d]. This one sentence has *aprons*, which seems to suggest that an attributive reading implies *aprons*. But this sentence is the only instance with *aprons* among the five tokens of this attributive context, so an attributive reading by itself can't explain *aprons* in this specific example. I return to this token below.

Let us turn to the second text, namely *Durham wills and inventories*. The next text is not, strictly speaking, a single text but a series of wills; still, they are all from one locale and one tradition over a short interval, from 1562 to 1570, and can be treated as a single text. (In fact, there is a string of four tokens of *(n)apron* in a row.) The tokens are given in Table 4.

Only two contexts occur in this small corpus. One context is represented by two tokens, in which the indefinite article and noun are separated by a modifier (*a linn Apron*, *a blewe apron*). The modifier makes these constructions novel. This pair of examples suggests two thoughts: that the innovative form *apron* is favored to the extent the context in which the noun is used is non-idiomatic, novel; and the concept of idiomaticity is a gradated (not discretely binary) parameter. To continue down this path, both in the example from 1562 and the 1570 example (from volume III of these documents), *an apron* occurs in the middle of a miniature list of three bequeathed items. Lists by their nature hint that a set of entities could be extended, so they promise a modal, possible, open-endedness. Thus it appears that open-endedness favors the innovative form *an apron*. In contrast, in the will of cook William Hawkesley (presented in Table 4 as a block of four tokens), the second through fourth tokens have a fixed phrase *a napron*, and the whole construction, 'I give to *x a napron*', is a standard idiom of bequests. Thus fixed idioms use the inherited older form *a napron*. In 1569 Alice Barnes receives two worsted items; possibly the parallel in material is the critical information, and the fact that one is a *(n)apron* is incidental.

The most interesting example is the first example of the 1570 set. Throughout this will of William Hawkesley, the recipients are identified in an unambiguous but not expansive fashion; the recipients are listed by name alone (22 xx) or with name and geographical location (3 xx) or name and relationship (9 xx), such as mother-in-law or midwife. That is, the recipients are presumed to be known by name with the briefest of descriptions,

Table 4: *(n)apron* in *Durham Wills and Inventories*

type	text
MOD	It' I bequith to Agnes Carter ***a linn Apron***. (I.277, 1567)
MOD	It'm I gyve to Helenor Huntley iiij^{or} blake patletts iiij^{or} cherches ***a blewe apron*** & ij^{o} velvett pattletts (I.343, 1570)
ART	to Thomas Burdon a busshell of wheat – to Jane Brantinga' a line kyrcheff ***an apron*** & a pair of hoose (I.198–99, 1562)
ART	I geve unto Elizabeth Hackforth a kerchif, a raill, a smock, ***an apron*** and all my workday rayment and in mony 3*s.* 4*d.* (III.56–57, 1570)
ART	It'm I gyve to katheryn barnes ij^{s} vj^{d}. It'm I gyve to thomeis hynde y^{t} was my p'ntice ***an apron*** & a new
ART	fyshe knyffe. | It'm I gyve to thomas capstone ***a***
ART	***napron***. | It'm I gyve to thomas boswell ***a napron***.|
ART	It'm I gyve to luke hanynge ***a napron*** & a fyshe borde. (I.327, 1570)
ART	And to alles Barnes a gowne of worsted & ***a napron of worsted*** (I.305, 1569)

almost titles. Against that background, the description of the first recipient of aprons, *thomeis hynde y^{t} was my p'ntice,* stands out; given its relative clause *y^{t}* ('that'), the identification of Thomas is relatively elaborate. Indirectly, this means that the bequest – an apron and a fish knife – is out of the ordinary, atypical. In contrast, in the three bequests that follow immediately thereafter are idiomatic. It appears, then, that novel or unexpected bequests of aprons – the bequest itself or the recipient – are expressed by the innovative *(an) apron*, while less novel scenarios are expressed by the older form *a napron.*

This takes us back to the one example in the Liber Niger Domus Regis Edward IV which had *aprons* (other than *none aprons*): *[he taketh] at eueryche of the iiij festes of the yere, of the clerk of greete spycery, ij elles of lynen clothe* ***for aprons***, *price the elle, xij*^{d}. In and of itself, the sentence is unremarkable and indistinguishable from the other examples with prepositions which had *naprons.* What might be atypical is the office described here, which is that of *sewar*, the highest ranking and first mentioned of the king's servants: *A SEWAR FOR THE KYNG, wich owith to be full cunyng, diligent, and attendaunt. He receueth the metes by sayez and saufly so conueyeth hit to the kinges bourde with saucez according therto, and all that commith to that bourde he settith and dyrectith* (§33, p. 112). In this instance, although the act of taking aprons is not exceptional, the recipient – the *sewar* – is unique. This is then similar to *thomeis hynde y^{t} was my p'ntice* from *Durham wills*

and inventories, in the sense that the non-idiomatic character of the example derives from the recipient, not the *(n)apron* phrase. That should not be surprising, since the act of bequeathing includes a recipient as well as the item bequeathed. The innovative *aprons* here acts effectively as an honorific to draw attention to the unusual status of the recipient, as it did with *thomeis hynde y^{t} was my p'ntice*.

In general, it appears the innovative form is favored if the transfer of *apron* is novel, not typical, and this extends to the recipient of the transfer (relative to other recipients). This principle applies to both example sets from different stages of the change. This distribution – unidiomatic context prefers the novel form – turns out to match other instances of the competition between equivalent morphological forms. Thus in contemporary Czech the locative singular (used with certain prepositions) can be either the traditional ending {-e} or a new ending {-u}. (That ending is original with nouns of the Indo-European *u*-stem declension, but its use with *o*-stem masculine nouns is new.) The parallel is that the traditional {-e} is used with "typical" combinations while innovative {-u} is used with atypical contexts (**BERM**).

There is another regularity of some interest that applies to both texts. We saw above that in Liber Niger there were two instances in which *(n)apron* followed the numeral *j* ('one'), and in both the older form *napron* was used. The examples are more or less equivalent in meaning to a true indefinite article as in *a napron*; the two examples of *j napron* (with *napron*) with the numeral invite the suspicion that true indefinite articles at this time might have *napron*, if they were attested. Conveniently, there is a contemporaneous will that has two tokens of an indefinite article one after the other: *Also I gyve to Margarete Holton my best kyrtill & **a napron**. Also I gyve to Elisabeth Wike a smok & **a napron*** (will of Jone Montor, 1489, *Surrey Wills*, p. 95). A slightly later example is consistent: *A jak & a salet, a gorget, ij gussettes, **a napron**, and iij gauntlettes* (*York wills*, p. 35, 1512). These examples at least suggest that the context with an indefinite article used the more conservative form.

To return to the other text under discussion here, *Durham wills and inventories*, there were two recognizable contexts. One involves an indefinite article split from the noun; it does show that the change had progressed to novel (unidiomatic) contexts. The other context had seven tokens with an indefinite article (not separated from the noun); the older form *napron* was used 4 times, the novel form time 3 times. That is to say, the novel form *apron* was slow to appear in the context in which the indefinite article immediately preceded the noun. For both periods (late 1400s, third quarter of 1500s) it appears that a construction with an indefinite article uses *apron* less (or at least not more) than other contexts.

Now the standard analysis is that the ambiguous combination of indefinite article and *napron* led to a reanalysis of {a+napron} to {an+apron}. If so, it would be natural to expect that *apron* would be used first in the context of reanalysis and only later in other contexts – that is, it should appear earliest with the indefinite article. But we just saw that in both texts, *apron* was used in other contexts when *an apron* was not yet used (the first text plus the auxiliary wills) or not used as frequently (the second text).

This suggests a revision of the account of reanalysis. Since the appearance of *apron*

is in fact not tied to the indefinite article, the unit *apron* appears to have some degree of autonomy. The reanalysis consists not of replacing the underlying shape of the noun, but it consists of imagining the possibility of an alternate word form {apron}, which coexists, for a time, with an alternate sublexeme {napron}. Imagined {apron} becomes real only when it is actually used. Following the general principle that innovative forms appear first in novel contexts, {napron} was maintained with the indefinite article – in fact, the most conventional and idiomatized construction – while the sublexeme {apron} was used in novel contexts cited above, such as *[takith]… ij elles of lynen clothe* ***for aprons*** and … *to thomeis hynde y*^{t} *was my p'ntice* ***an apron***. Over time, {apron} and {napron} compete; {apron} keeps on increasing, in a fashion that could be understood as the other half of Newton's first law: once in motion, a body, or linguistic subsystem, will remain in motion.

Semantically the new demilexeme {apron} must be basically similar to traditional {napron}. For example, both demilexemes refer to protective coverings, usually of cloth, though in artisanry, aprons could be sheepskins. In the York fabric rolls – the record of expenses of the York expenses, including irregular expenses of masons and their equipment – we observe naprons used at the beginning of the fifteenth century – *In ij pellibus emptis et datis eisdem pro naporons* 'two hides were bought and given to them to serve as aprons' (1423) – and then the form is *aprons* at the end of the fifteenth century]: ***pro ij le aprons*** *de correo pro les setters per spacium ij mensium, 12*d. 'for two aprons of hide for the setters for the period of two months, 12 shillings' (1504). That is only to say that *napron* and *apron* seem to have the same extension.

Still, despite the overlap in extension, there are indications that the two sublexemes began to develop slightly different connotations.[4] Two facts argue for this.

The first, perhaps unexpectedly, has to do with translations of the Bible. As is well-known, John Wyfcliffe translated much of the Bible from the Vulgate, around 1382. (His translation was finished by his followers after his death.) A passage of interest is Genesis 3:7 – the famous story of the nakedness of Adam and Eve – for which Wycliffe (or his followers) translated Vulgate *…cognovissent esse se nudos consuerunt folia ficus et fecerunt sibi perizomata* as *…and when they knew that they were naked, they sewed the leaves of a fig tree, and made breeches to themselves.*

Wycliffe's Bible became the model for an extended tradition of English translations thereafter, but with a difference in this passage. Starting with Tynsdale (1534), the subsequent translations have a different noun in Genesis 3:7: *…vnderstode how that they were naked. Than they sowed fygge leves togedder and made them apurns.* The translation with *aprons* continues through Cloverdale (1535), the Great Bible (1540), Matthew's Bible (1549), the Catholic Bishops' Bible (1568), the Geneva Bible (1587), and finally the King James (1604–1611). All have *aprons* (variant spellings) except for the Geneva Bible, a retrograde Protestant Bible which returned to Wycliffe's *breeches*.

The improvised fig-leaf garment of Genesis 3:7 wasn't exactly an apron in the sense of linen or hide aprons, but it was somewhat similar. Why was *apron* used instead of *napron*? One reason might be that *apron* was the innovative form, and innovative forms

[4] In a fashion consistent with Bréal's "law of differentiation" of synonyms (1900: ch. 2

are more appropriate than conventional forms for encoding semantic extensions. There is another possibility. As we saw in the Liber Niger, *naprons* were something that would result from linen, and their value was defined by the price of the linen used to make them. In earlier wills *naprons* were classified with other items of cloth with different functions; *naprons* belonged to the *naperie* (the collection of similar cloths) along with *napkins* (same sense as modern) and *borde clothes* or *table cloths*. So the sublexeme {napron} emphasized the origin of the entity in cloth or hide; secondarily, such a flat piece of material could be donned for protection. With the sublexeme {apron} the dominant feature is not that it was made from material (or hide); the dominant feature is that it is a garment worn to provide protective covering. This difference in the ranking of features – MATERIAL as opposed FUNCTION – might be why the corrections to Wycliff's Bible used *apron*.

A second indication is the way items are grouped in *Chesterfield wills and inventories of household goods*. For example, from Derbyshire #177 (Margaret Capper, 1588) here is a partial list of items (omitting tools and animals). Items are listed in the inventory in natural classes. Categories are added here:

> <*furniture*> / 4 bedstids in the Chamber / 2 bedstids in the parlar with pented Cloates about them / 1 bed teaster of Cloathe / 3 bed stids in the nether Chamber / <*bedding*> / 8 pillobears / 8 hand towels / 4 shietes / <*garment*> / 1 smock / *2 aperns* / 1 bruse and a grater / 1 Coat and a pear of house / 1 Gone and a for kertle [?] / 1 buckrame savgard / <*utensils*> / 2 Chamber pots / 1 morter and a Cresset / 3 Chaffindishes / 1 Skomar and a ladle of brase / 9 bear potes and 2 black potes / 1 falling bord in the house / 3 pans 2 ketles / 1 basson brase / 4 brase potes

Note that *aprons* is listed next to smock and other garments. (The listing of a "bruse and grater" below *aperns* seems out of place.) By this time, in the late sixteenth century, an {apron} was classified as a garment.

Does this explain why {apron} continued to displace {napron}? Possibly. The demilexeme {apron} removes aprons from the domain of the *naperie* (the collection of pieces of fabric) to the domain of garments. The extension may be the same, but the intension changes, by the re-ranking of the semantic features of the two demilexemes: {napron} ranks the material over the function, whereas {apron}, while it does not that fabric may be involved, ranks garment and its function of covering as more important.

There are several conclusions here. The two demilexemes have a certain autonomy: they have overlapping but not identical semantics; they have different preferred contexts in which they appear. From this it follows that the change of *napron* to *apron* is not a simple substitution of one form for the other. Next, the newer form *apron* seems not to appear in the context of an indefinite article ahead of other contexts, as might be expected if *apron* merely replaced *napron*. This again implies that *apron* and *napron* are somewhat separate entities. Third, the ambiguity of [anapron] made the change possible, but the change was the creation of two demilexemes here, {napron} and {apron}, not a reparsing.

5 Conclusion

The examples above suggest that *parole* exists, that it has a role in language. Saussure, as we saw, did his best to hide *parole* from view, but it ended up that *parole* has a life of its own: it required its own set of rules and it was maintained (the accent stayed on the same syllable as in Latin) without justification in the system.[5] All the action of stress in Romance was in *parole*, not *système*. In other examples, we saw that *parole* is maintained from one generation to the next, not because it is motivated by higher principles, but because it was the usage and it was then transmitted as usage. *Parole* can also shape other changes (such as the apocope of post-tonic vowels in the transition from Latin to French and the restriction on intrusive /r/ to low vowels). Variants of lexemes, such as {napron} and {apron}, have partially separate existences and properties, including semantics.

Parole is not always static and it is not one-dimensional. *Parole* is, after all, activity, and human activity implies the possibility of more activity and other paths of activity, which may differ from inherited activity. *Parole* is habit infused with potential.

[5] Boris Gasparov (**gasparov2012**) has argued that Saussure's thinking was more complex (and less rigidly categorial and structuralist) than the subsequent reception would have it (especially chapter 4, pp. 111–37).

Abbreviations and texts cited

Abbreviation	Explication of abbreviation
Surrey wills	*Surrey wills.* (Archdeaconry Court, Spage Register). 1922. Vol. 5 (Surrey Record Society). Surrey: Roworth & Co., for the Surrey Record Society.
York wills	*Testamenta eboracensia, or Wills registered at York: illustrative of the history, manners, language, statistics, &c., of the province of York, from the year MCCC downwards.* London: J. B. Nichols & Son, 1836-1902.
Wills and inventories of Bury St. Edmonds	*Wills and inventories from the registers of the commissary of Bury St. Edmunds and the archdeacon of Sudbury.* 1850. London: Printed for the Camden Society.
Fabric rolls	*The fabric rolls of York Minster with an appendix of illustrative documents.* 1859. Vol. 35. (Publications of the Surtees Society). Durham: Published for the Society by G. Andrews.
Chesterfield wills and inventories	Bestall, J. M.; Fowkes, D. V., ed., a glossary by Rosemary Milward with an introduction by David Hey & an index by Barbara Bestall. 1977. *Chesterfield wills and inventories 1521–1603.* Vol. 1 (Derbyshire Record Society). Derbyshire: Derbyshire Record Society.
Durham wills and inventories	*Wills and inventories illustrative of the history, manners, language, statistics, &c., of the northern counties of England, from the eleventh century downwards.* 1835. Vol. I (Publications of the Surtees Society, vol. 2). London: J. B. Nichols & Son; 1906. Vol. III (Publications of the Surtees Society, 112.). London: J. B. Nichols & Son.

a community proofreader or typesetter
at http://www.langsci-press.org/register.

Change your backtitle in localmetadata.tex

Change your blurb in localmetadata.tex