The Phonetics-Phonology Interface and the Acquisition of Perseverant Underspecification*

Mark Hale
Department of Classics, Modern Languages and Linguistics
1455 de Maisonneuve Boulevard West
Concordia University
Montréal, Québec, H3G 1M8, Canada
Tel: (514) 848-2424, ext. 2758
Fax: (514) 848-8679
Email: hale1@alcor.concordia.ca

Madelyn Kissock
Linguistics Department
329 O'Dowd Hall
Oakland University
Rochester, MI, 48309 USA
Tel: (248) 370-2174
Fax: (248) 370-3144
Email: kissock@oakland.edu

*The authors would like to thank the editors of this volume, especially Charles Reiss, and the audiences of the *6e Journées Internationales du Réseau Français de Phonologie* (June 2004) and the *13th Manchester Phonology Conference* (May 2005) for their helpful comments and suggestions.

1 Introduction

This paper explores the phonetics-phonology interface, both in the broadest sense of 'phonetics' which addresses aspects of articulation and acoustics and in the narrow sense of 'phonetics' which is concerned with abstract surface/phonetic representations which are the output of phonological computations. We examine here a particular type of underspecification, which we will call 'perseverant,' acquisition of which is motivated primarily by exposure during the learning process to speech sounds which are not consistent in their physical properties with assignment to any fully-specified set of features. The more common types of underspecification are motivated by data from alternations or by a desire for economy of representation which results in underspecification at the level of UR but not at the phonetic level. By contrast, the type of underspecification we will be discussing is motivated by a particular well-defined property of the acoustic signal and *perseveres* from the underlying forms *through* the surface or output forms. Because of its perseveration, the underspecification detailed here provides an especially fertile ground for the exploration of the phonetics-phonology interface. Data from the Marshallese vowel system will provide the empirical basis for our discussion. We will use this data to explore what sorts of physical phonetic information can trigger an acquirer to store underspecified representations as well as how both the acquisition and ultimate grammar can be modelled using optimality theoretic assumptions. This discussion will reveal various problems with the notion of 'markedness' as currently used, as well as with the acquisition path to an OT grammar. Much of the discussion will have implications for the more global question of how and why an acquirer assigns certain features to an abstract representation based on information extracted from an acoustic event.

The second section of the paper outlines our basic theoretical assumptions and definitions of terms. Section three presents very briefly an example of a more common type of underspecification for purposes of contrast and then outlines two cases of what we will call 'perseverant' underspecification, a Russian example discussed in Keating (1988) and the case of Marshallese vowels. In Section four, we present the details of a constraint ranking that would capture this underspecification in an adult grammar. In Section five, we discuss how an adult grammar of this type can be acquired. Section 6 presents some additional supporting data from Marshallese loanwords and Section 7 presents our conclusions.

2 Background Assumptions

Since our primary interest is in modelling aspects of the phonological computational system (including the primitives over which such computation takes place), which are the result of phonological acquisition (and some innate properties of UG), we begin by presenting our assumptions about the nature of this computational system. This necessitates a discussion of the division between phonology and phonetics and between linguistic and non-linguistic processing, as well as explicit definitions of the terminology we use. Explicitness is particularly important in the present context since there is not general agreement among researchers on these matters and we believe that there has been considerable confusion introduced by the use of vague or ambiguous terminology. The one point on which there *is* general agreement, at least among phonologists, is that 'phonology,' in its narrowest sense, consists of the mapping from underlying representation to surface representation. Keating (1988) explicitly addresses this issue and we follow her in both our use of terms and in our definition of 'phonology.' For Keating and for us, phonology involves only a feature-to-feature mapping and nothing else. Other researchers have defined phonology much more broadly and have extended it well beyond the feature-to-feature mapping. Hammarberg (1976), for example, defines any aspect of pronunciation that involves cognition (for example, anticipatory co-articulation) as part of 'phonology.' For Hammarberg, then, a mapping of a set of features to a gestural score[1] or of an acoustic score to a set of features, because both are considered 'cognition,' would be included within the scope of his 'phonology.'[2]

We propose that mappings between dissimilar representational formats, such as from features to a gestural score, are performed by *transducers*. Transduction, in general, is a function which converts a form in one representational 'alphabet' to a form in a different representational alphabet.[3] Phonetic transduction is thus to be distinguished from phonological computation by the fact that it incorporates some type of *conversion* process – it changes one type of representation (featural, for example) into another type of representation (gestural score, for example). In our model phonological computations, unlike transduction, operate on only a single type of symbolic representation – features. Features are both the input to and the output of the phonology – phonological computations cannot convert features into other types of representations. Under any analysis, however, the incorporation of a transduction process of some type into the model of speech production seems inescapable, since there are no features actually present in the acoustic output of speech.

A further logical necessity is the presence of two distinct transducers – one for processing representations concerning audition and one for processing representations concerning articulation. Positing two transducers is suggested by the fundamentally distinct nature of auditory vs. articulatory processing. For example, while both involve 'unidirectional' processing, the *direction* of processing is not the same in the two cases. Articulation demands transduction of features (the input) to some gestural score (the output), whereas audition requires transduction of a percept (the input) to features (the output).[4] In addition, we assume that there is an actual, physical difference between the mechanisms involved in audition and those involved in articulation and that dedicated transducers reflect this difference. We assume that these two transducers are innate and invariant – they are identical in all humans (barring some specific neurological impairment) and do not change over time or experience (i.e., they do not 'learn').[5]

Our model assumes strict modularity—the modules are 'cognitively encapsulated' such that

[1]We borrow the term 'gestural score' from Browman and Goldstein (1990) and extend it to the acoustic domain.

[2]While Hammarberg defines phonology more broadly than we do, he maintains distinctions between processes within this broader domain.

[3]We borrow this terminology from Pylyshyn (1984). Our use of the term is related to his but not identical with it. Pylyshyn's primary concern is the more general area of computation and cognition.

[4]We do not intend by this that there are *only* two transducers. We assume that transduction is a complex process which involves many different transducers. However, for our purposes here, only the highest level of transduction is immediately relevant – that of features to gestural score or acoustic score to features.

[5]This claim regarding innate, unchanging transduction is made solely for these two transducers which take one type of symbolic representation and convert it to another type of symbolic representation. It is not a claim that motor skills, for example, are not learned or do not develop or mature over time – it is not a claim about motor skills at all.

no component can see ‘inside’ another component. Only the output of one module may be fed to another module and then only in the case of particular modules. So, for example, the output of the phonology is the input to the auditory transducer, but the acoustic transducer does not feed its output to the auditory transducer, nor vice-versa. Thus the two transducers operate independently of one another and have no interaction. Following Hale and Reiss (2000ab), our model necessarily divorces (phonological) features from both articulation and audition. Features are simply symbolic, ‘substance-free,’ primitives which are manipulated by the phonology and the transducers. The very fact that two separate transducers are required – one for articulation and one for audition – forces the separation of features from any physical substance. Since what is mapped onto a single feature comes from two very different sources, this separation from the physical substance is a logical necessity – a single feature cannot, for example, be *both* derived from the muscle commands involved in raising the tongue body *and* from a neural impulse triggered by some portion of an acoustic wave (let alone some actual property of the wave itself). Thus, we can consider the transduction process, too, as invariant in that the *relationship* or *mapping* between a particular feature bundle and a particular gestural score is a deterministic (and thus consistent) conversion process and, similarly, that the relationship or mapping of a particular auditory input to a feature bundle is deterministic. Crucially, the features and their transduced output forms are different from one another.

Finally, we assume that phonological features, themselves, are universal in the sense of Universal Grammar (UG). A universal feature is not one that is found ‘universally’ but rather a feature which is drawn from a universally available but finite inventory.[6] Any UG feature *may* be present in the actual mental representations of a particular instantiation of natural language but it is not clear that every feature *must* be present. Since the symbolic representations of natural language segments appear invariably to be feature *bundles*, it is actually the possible featural combinations that are responsible for the wide variety of sounds found in language, not the sheer number of features.[7]

The relevant mappings in our model of phonology and transduction are shown in (1) below.

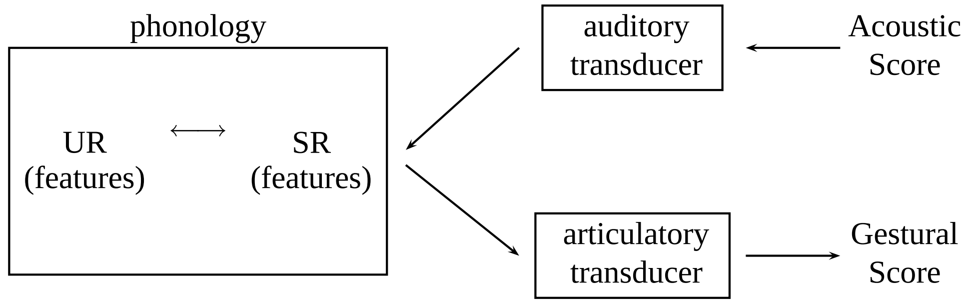
(1) Phonology: UR $\leftrightarrow$ SR

Transducer$_{auditory}$: SR $\leftarrow$ Acoustic Score

Transducer$_{articulatory}$: SR $\rightarrow$ Gestural Score

[6] For a view opposing the universality of features, see Pulleyblank (2001).

[7] We use the term ‘bundles’ loosely to represent groupings of features. We take no position here on the best way to represent these groupings.



This model and the accompanying discussion have so far presented only what we believe to be computations and processes *specific to language*. As such, they represent only a portion of what makes up an individual's actual behavioral output or *performance*. We assume that between, for example, the gestural score and the point of actual physical output, there can be input or modification from many other *non-linguistic* facets of cognition that determine amplitude, speech rate, affect (e.g., tone of voice), and other situational effects. Crucially, none of these post-gestural-score additions contained in the physical output appears to be utilized by the linguistic computational or processing systems and therefore a sharp distinction between the two types of processing, linguistic vs. non-linguistic, is indicated.[8]

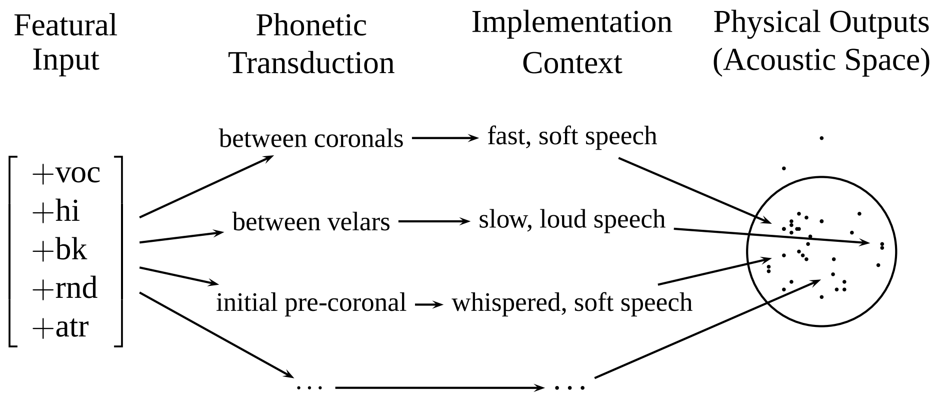
While we believe that features are purely formal representations and that the transducers are invariant across speakers and deterministic in the way they perform conversions, we still predict that there will be differences in the measurable, physical instantiations of any *single* feature bundle. These differences will have at least one of three possible sources (and will probably be due to more than one): 1) the articulatory transducer, although mechanical, implements features and feature bundles in a context-sensitive manner[9]; 2) there are within- and cross-speaker differences in physical attributes (e.g., subglottal pressure, size and shape of oral cavity); and 3) there are external, physical forces which may interfere with production (e.g., external air density). As a result, a particular featural representation will correspond to multiple, physical instantiations (on separate occasions) which nevertheless typically fall within some reasonably well-defined area – the *acoustic space.*[10] We will see that, in the Marshallese cases, it is specifically the *absence* of a well-defined area along certain dimensions that is crucial evidence for the acquirer of Marshallese. This potential chain of events leading to physical output can be schematically represented as in (2) below.

(2) Transduction and the acoustic space.

[8]As is well known, phonological acquisition routinely and successfully takes place in environments where there is great variability in the physical production of featurally-identical speech sounds. For a summary of some of the research on this topic, see Jusczyk and Luce (2001).

[9]This includes whatever effects the implementation of one feature may have upon the implementation of another feature within a single feature bundle as well as co-articulation effects.

[10]We avoid using the term 'target' since it incorrectly suggests both that there exists a single, 'correct' physical target and that the transducer has frequent 'misses.'



The above figure only schematizes articulation. However, the concept of multiple physical instantiations being related to a single featural representation is, of course, the same for audition. In the case of audition, however, physically distinct inputs (from different occasions) will be stripped of their context-dependent and idiosyncratic physical properties, and reduced to a single identical set of features. In both articulation and audition, the acoustic space is a label for some physically definable area. We hypothesize that only a change in features will produce any significant change in acoustic space. This is because the differences incurred through changes in context or physical attributes will remain, we assume, relatively constant independent of the make-up of the particular feature bundle. So, roughly, the particular features determine the *locus* of activity and non-featural attributes determine the cluster pattern around the locus. In the following section we turn to a consideration of some of the implications of underspecification, both 'underlying' and so-called 'phonetic', for this model of phonology.

3 Phonological Underspecification with and without Phonetic Underspecification

3.1 Phonological Underspecification with Full Phonetic Specification

Most commonly, discussions of underspecification have been directed toward featural underspecification at the underlying, lexical level of representation and not toward underspecification in derived phonetic representations. The motivations for such underlying underspecification have been varied but generally are based upon arguments from markedness or economy (see, for example, Steriade 1995). Underspecification has also been proposed within Optimality Theory to account for alternations of the type discussed in Inkelas (1994). Some of the data Inkelas provides to support underspecification concerns the case of Turkish plosives, an example of which is given below.

(3) Turkish Plosive Underspecification (Inkelas 1994)

	Nom	Acc	gloss
Non-alternating	sanat	sanat-ɨ	‘art’
Non-alternating	etüd	etüd-ü	‘etude’
Alternating	kanat	kanad-ɨ	‘wing’

Inkelas argues that, in the case of predictable alternating forms such as the ‘wing’ forms above, the process of Lexicon Optimization will force an underspecified representation. The plosive will be underspecified for voice in the input and the voice value in the winning output candidate will be determined by the ranking of structure-filling constraints.

Underspecification of this type is crucially different both in the type of data which motivate the need for underspecification (largely data from alternations) and the extent of the underspecification itself (underlying representation only) from the cases we focus on in the rest of this chapter.

3.2 Underspecification in both Phonological and Phonetic Representations – Perseverant Underspecification

The focus of this paper is on underspecification which persists from underlying representation through phonetic representation, resulting in forms which are *never* fully specified featurally. We will call this phenomenon ‘perseverant underspecification.’ The general characteristics of perseverant underspecification are: 1) absence of an articulatory or acoustic target for one or more features; 2) an articulation which is determined in some relevant respect by context (as opposed to by some feature value or values); and 3) the existence of alternating but entirely predictable articulations and corresponding acoustic products.

Keating (1988) convincingly argues that Russian [x], in the case where no context rules apply, consists of a *phonetic* feature bundle with no specification for the feature [back].[11] Therefore, the following two instances of [x] are crucially different with respect to phonetic feature specification in the output of the phonology: the (a) case has a fully-specified feature bundle and the (b) case is underspecified, having no [back] feature.

(4) Fully-specified and Underspecified Russian /x/ (Keating 1988)

(a) /axi/ → fully fronted fricative; context rule filled in [-back] in the course of phonological computation

(b) /ixa/ → transient fricative, gradual transition throughout its duration from acoustic correlates appropriate to the preceding [-back] segment to those appropriate to the following [+back]) segment; fricative remains underspecified throughout phonological computation

Keating (1988:285) notes as part of this discussion that:

[11] Note that this means that feature bundle will *never* have a feature specification for [back] because the transducer does not ‘fill in’ feature values.

> . . . if a segment acquires a feature value from an adjacent segment, it will share a phonetic property with that segment across most or all of its duration; if a contour is built through a segment it will have a more or less continuously changing, transitional, quality from beginning to end that will depend on context on either side.

As described above, the outcome of the velar fricative in /ixa/ shows the continuous, transitional features characteristic of those sounds which are entirely dependent in some aspect of their articulation upon a feature or features of adjacent sounds.

The case of Marshallese vowels is a rather more elaborate instance of perseverant underspecification and one which will highlight the questions which such underspecification cases raise for acquisition. The Marshallese vowel system, discussed in Hale (2000) inter alia, is quite striking. The 'surface' vowels are given below, where the 'tie' symbol (as in i͜u) represents a smooth transition from one vowel to another, e.g. in this case, *i* to *u*.

(5) Marshallese ‘Surface’ Vowel Inventory

i	ɯ	u	i͜ɯ	i͜u	ɯ͜i	ɯ͜u	u͜i	u͜ɯ
ɪ	ɤ	ʊ	ɪ͜ɤ	ɪ͜ʊ	ɤ͜ɪ	ɤ͜ʊ	ʊ͜ɪ	ʊ͜ɤ
e	ə	o	e͜ə	e͜o	ə͜e	ə͜o	o͜e	o͜ə
æ	a	ɒ	æ͜a	æ͜ɒ	a͜æ	a͜ɒ	ɒ͜æ	ɒ͜a

To understand how these surface vowels come into being, we need to explore the *phonology* of Marshallese. We will begin with the consonants which are rather unusual. The features of these underlying consonants play a key role in developing an account of the Marshallese surface vowels given above. A chart of the consonants is given below.

(6) Marshallese Underlying Consonant Inventory

	Oral Stops			Nasals			liquids & glides		
	Labial	Dental	Velar	Labial	Dental	Velar			
‘light’ [-bk,-rnd]	pʲ	tʲ		mʲ	nʲ		lʲ	rʲ	j
‘heavy’ [+bk,-rnd]	bᵚ	tᵚ	k	mᵚ	nᵚ	ŋ	lᵚ	rᵚ	ɰ
‘round’ [+bk,+rnd]			kʷ		nʷ	ŋʷ	lʷ	rʷ	w

As Choi (1992) demonstrates, in CVC sequences, there is a steady transition during the vowel articulations between the [back] and [round] features of the preceding consonant to the [back] and [round] features of the following consonant in every instance. These transitions are phonetically distinct from diphthongs which have a relatively long-duration nucleus and a brief on- or off-glide. As Bender (1968) showed, the most coherent phonological analysis of the Marshallese vowel inventory is one in which the vowels themselves bear no features along the dimensions back and round. That is, they differ from one another *only* along the height and ATR dimensions. We will use Cʲ to represent ‘front’ (i.e., palatalized) consonants, Cᵚ to represent back non-round (i.e., velarized) consonants and Cʷ to represent back round (i.e., labialized) consonants. For the vowels not specified along the back or round dimensions, we introduce the new symbols [V_{HI}] for a [+high,-low,+ATR] underspecified vowel, [V_{MID}] for a [-high,+low,+ATR] underspecified vowel, and [V_{LO}] for a [-high,-low,-ATR] vowel, likewise underspecified.[12] One can readily see how the ‘surface inventory’ of Marshallese vowels arises from these consonant-vowel interactions in the table below. The phonological representations for vowels are listed on the left with only the vowels’ featural representations. The corresponding articulations are listed on the right. The notation for the surrounding consonants is as follows: F=nonback, nonround consonants; B=back, nonround consonants, R=back, round consonants).

(7) The Underlying Sources of the Marshallese ‘Surface’ Vowels

[12] Hale (2000) uses the somewhat less common symbols ☕, ☯, and ⚽, respectively.

Input V-features		consonantal environment								
hi	ATR	F—F	B—B	R—R	F—B	F—R	B—F	B—R	R—F	R—B
+	+	i	ɯ	u	i͜ɯ	i͜u	ɯ͜i	ɯ͜u	u͜i	u͜ɯ
+	-	ɪ	ɤ	ʊ	ɪ͜ɤ	ɪ͜ʊ	ɤ͜ɪ	ɤ͜ʊ	ʊ͜ɪ	ʊ͜ɤ
-	+	e	ə	o	e͜ə	e͜o	ə͜e	ə͜o	o͜e	o͜ə
-	-	æ	a	ɒ	æ͜a	æ͜ɒ	a͜æ	a͜ɒ	ɒ͜æ	ɒ͜a

Finally, we introduce the *necessary* distinction between the output of phonological computation (which we place between traditional square brackets) and the articulatory-acoustic output of the body (which we place between 'body' brackets). We can then represent schematically the treatment of the underspecified Marshallese vowel segments as follows:

(8) ¿From UR to 'Surface' Realization

a. $C^jV_{HI}C^w$: /$n^jV_{HI}k^wn^jV_{HI}k^w$/ > [$n^jV_{HI}k^wn^jV_{HI}k^w$] > 🚹n^{j}i͜uk^{w}n^{j}i͜uk^w🚹 'clothing'

b. $C^jV_{MID}C^ɯ$: /$n^jV_{MID}t^ɯ$/ > [$n^jV_{MID}t^ɯ$] > 🚹n^{j}e͜ʌt^ɯ🚹 'squid'

c. $C^jV_{LO}C^j$: /ṯ^jV_{LO}ṯ^j/ > [ṯ^jV_{LO}ṯ^j] > 🚹ṯ^{j}æṯ^j🚹 '*Lutjanus Flavipes*'

Note that this gives Marshallese what appears superficially to be a large and rather unique vowel inventory with a grand total of 36 vowels whereas, in fact, for all grammatical (i.e., featural) purposes, the Marshallese inventory is quite small, having only four featurally distinct vowels at both UR and SR levels.

The case of /ṯ^jV_{LO}ṯ^j/ is particularly interesting. This vowel will show apparent steady-state realization in the 🚹æ🚹 space, much like English 🚹æ🚹. Further analysis, however, reveals that this identity is purely superficial and that it actually obscures a significant difference between Marshallese 🚹æ🚹 and English 🚹æ🚹. The two are quite distinct, representationally, with the front and non-round properties of Marshallese 🚹æ🚹 determined entirely by the [back] and [round] features of the adjacent consonants and the front and non-round properties of English 🚹æ🚹 the result of a feature bundle that includes specific values for [back] and [round]. Cases like this of superficial similarity are especially relevant to acquisition and will be discussed in more detail in section 5.

4 Constraint-based approach to perseverant (phonetic) underspecification

At first glance, the Russian and Marshallese cases appear to be made for an Optimality Theoretic approach.[13] The obligatory underspecification of Marshallese output vowels for the features [back] and [round] can be made to follow trivially in OT by simply ranking the Markedness constraints against values for these features (abbreviated here as *[±round] and *[±back]) higher than the

[13]Note, however, that this is true only if one assumes the formal approach to phonology sketched above – one where features are simply symbolic representations manipulated by the grammar.

Faith requirements which would necessitate respecting such values. Such a ranking will force winning candidates to be underspecified, as can be seen from the sample tableau for /p^{ɯ}V$_{HI}$p^{ɯ}/ ‘black triggerfish’, which is articulated as [p^{ɯ}ɯp^{ɯ}], is given below.[14]

(9) OT and Perseverant Underspecification

/p^{ɯ}V$_{HI}$p^{ɯ}/	*V[±round]	*V[±back]	MAX-IO	DEP-IO	*[+ATR]	*[+hi]
a. [p^{ɯ}up^{ɯ}]	*!	*		**	*	*
b. [p^{ɯ}ip^{ɯ}]	*!	*		**	*	*
c. [p^{ɯ}ap^{ɯ}]	*!	*		**		
d. [p^{ɯ}V[+hi]p^{ɯ}]			*!			*
☞ e. [p^{ɯ}V$_{HI}$p^{ɯ}]					*	*

The ranking in this tableau results in candidates (a)-(c) above being eliminated because of their violations of the relevant Markedness constraints. The even more underspecified, and thus seemingly less marked, candidate in (d) is unfaithful to the [+ATR] specification on the input vowel. Since this vowel is *too* underspecified for Marshallese, the Faithfulness constraint MAX-IO (which requires that the [+ATR] specification in the input be respected in the output) must outrank the Markedness constraint *[+ATR] (which would require the elimination of the underlying [+ATR] specification) in Marshallese. This leaves as the optimal candidate the V[+hi,+ATR] vowel of (e), underspecified in the output along the backness and roundness dimensions.

We see a number of interesting and unusual properties in the way these underspecified vowels interact with constraints. For example, insofar as the features of the vowel itself go, the optimal candidate violates neither Faithfulness constraints (MAX-IO and DEP-IO) nor the relevant Markedness constraints e.g., *V[±+round], *V[±-back]. In addition, the ‘missing’ features of the input candidate’s vowel are treated, for purposes of IO constraints, in the way that epenthetic vowels (or consonants) are treated. “Given the fact that an epenthetic segment has no input features to be faithful to, their feature content is delegated to markedness constraints.” [Kager 1999:125] However, unlike in the epenthesis cases, there is no violation of DEP-IO by the winning candidate since *no features are present in the output candidate which were not present in the input representation*. Finally, if we consider only those features relevant to vowels, there will be no metric for determining the relative ranking of Faithfulness and Markedness constraints, except with respect to those Markedness constraints regarding height and ATR features for vowels (both of which must be ranked below the IO-Faith constraints).

The tableau also raises some issues which are potentially more serious, however. One of these concerns the definition of Markedness in OT. While formally defined as ‘having violation marks’ (i.e., a more marked form is one which has incurred more violation marks for its output structure than some other competing form), markedness is, in OT practice, closely associated with the notion of ‘typological markedness’ and issues such as cross-linguistic frequency of occurrence. By formal

[14] Since the input representation in this case is underspecified, MAX and DEP constraints alone—regardless of the ranking of the Markedness constraints against specification along the [back] and [round] dimensions—would suffice to get the appropriate output. As pointed out above, the requirement that the Markedness constraints be ranked high is necessitated by the attempt to develop an account of the across-the-board underspecification of Marshallese vowels.

markedness definitions, Marshallese vowels will be highly unmarked, with a mid vowel which is less marked than a fully-specified /ə/.[15] However, in terms of typology or frequency of occurrence, the inventory of Marshallese vowels is highly marked in two ways:

(10) Marshallese Vowels as ‘Marked’ Segments

1. at the phonological level. From a featural standpoint, the vowel inventory is extremely small, a total of 4 contrastive vowels. Due to the lack of features, there can be no real notion of distribution throughout the vowel space – arguably the same as having a ‘poor’ distribution.

2. at the articulatory/acoustic level. The majority of the articulations appear to fall into the ‘not a natural language sound’ category. From a more formal standpoint, it appears that the articulations of the vowels give physical outputs that virtually cover the acoustic space, such that many of the vowels would be minimally contrastive acoustically. [16]

Before we turn to the next section, it should be noted that the contextual determination of vowel realizations along the [back] and [round] dimensions is *not* due to Marshallese speakers somehow ‘storing’ the complex transitions for each lexical entry.[17] We see this from an examination of affixes such as the productive agent noun prefix /rʲV_{HI} -/ whose vowel alternates between 🚹rʲi-🚹, 🚹rʲi͜ɯ-🚹, and 🚹rʲi͜u-🚹 depending upon the features of the initial consonant of the root to which it attaches.

(11) Variant Realizations of Marshallese /rʲV_{HI} -/

- 🚹rʲi-tʲerʲpᵚalᵚ🚹 ‘worker’ (cf. 🚹tʲerʲpᵚalᵚ🚹 ‘to work’)
- 🚹rʲi͜ɯ-pᵚɯ͜itʲpᵚɯ͜itʲ🚹 ‘one who kicks’ (cf. 🚹pᵚɯ͜itʲpᵚɯ͜itʲ🚹 ‘to kick’)
- 🚹rʲi͜u-ŋʷo͜erʲtᵚak🚹 ‘snorer’ (cf. 🚹ŋʷo͜erʲtᵚak🚹 ‘to snore’)

Further evidence of this is provided from loanwords into Marshallese in section six.

5 Acquiring perseverant underspecification in an OT grammar

We turn now to the question of the acquisition of underspecified forms in an Optimality Theoretic framework. Acquiring an adult-state constraint ranking in OT is done through the process of Constraint Demotion (Tesar and Smolensky, 2000, *inter alia*). The majority position holds that, at the initial state, all Markedness constraints are ranked above all Faithfulness Constraints (but see

[15] As indicated earlier, Faith constraints are not relevant to this determination.

[16] This is actually the ‘correct’ result from the point of view of the lexicon, where there are only 4 contrastive vowels, i.e., the *number* of contrasts is minimal.

[17] It would be hard to imagine what form such transitions could, in fact, be stored in, in any event.

Hale and Reiss (1998) for the alternative view). Certain Markedness constraints are then demoted based upon positive evidence – evidence which indicates that some more marked candidate wins over some lesser-marked candidate.

As Tesar et al. (2003) note, the awkwardness in this process lies in the fact that the acquirer must determine both lexical representations and a constraint ranking, each of which is dependent to a greater or lesser extent upon the other. They propose that the acquirer approaches the problem with a bias toward changing the ranking as a first step (with a further bias toward keeping Markedness constraints high whenever possible) and later, only if their ranking fails, modifying the lexical representation.

The target grammar for Marshallese is going to require a ranking where markedness constraints on round and back features outrank the relevant faithfulness constraints, so *V[$\pm$round], *V[$\pm$back] $>>$ MAX-IO,DEP-IO. The data that we assume is available to the acquirer includes the following:

a. both alternating and non-alternating forms.

b. forms, for example the one originating from Marshallese /t^{j}V$_{LO}$t^j/, that are indistinguishable (acoustically) from what could be fully-specified forms such as [t^{j}æt^j].

c. forms which, as far as we know, cannot be represented by any combination or geometry of our current feature set – e.g., the steady transition from the [ɯ]-space to the [u]-space in ⟨ɯ͜u⟩ – and therefore cannot be outputs of any human grammar (nor inputs for a human grammar).

If we consider an initial state where Markedness constraints are ranked above Faithfulness constraints (again, this is the majority view), then the path to acquisition of the grammar will proceed as follows. Suppose the acquirer first gets a form such as ⟨t^{j}æt^j⟩. The default hypothesis should be that the vowel in this form is a fully-specified vowel [æ] featurally identical to English [æ]. Such a hypothesis should result in the acquirer changing the initial ranking by moving Faithfulness constraints MAX-IO and DEP-IO (at least for the features [back] and [round]) above Markedness constraints *V[$\pm$round] and *V[$\pm$back].

Now consider the effect of a second possible piece of evidence for the acquirer – a form such as ⟨n^{j}e͜ət^ɯ⟩. The transitional quality of the vowel in combination with the non-back and round features of the consonants will lead the acquirer to posit *only* height and ATR features for the vowel but leave the vowel unspecified for [back] and [round]. The effect on the ranking of those constraints involving the features [back] and [round] will be null. The form [n^{j}V$_{MID}$t^ɯ], realized as ⟨n^{j}e͜ət^ɯ⟩, will still be the winning candidate with no change of ranking, since any competitors which have back and round features specified will incur gratuitous DEP-IO violations.[18]

A further step in the grammar acquisition process, Lexicon Optimization, will presumably lead to the acquirer positing additional underspecified vowels, namely those which are involved in morphological alternations. A vowel which is invariably realized as ⟨æ⟩, however, will maintain

[18]Note that if we reverse the order in which the acquirer receives the two pieces of data above, this does not alter the ranking outcome.

its fully-specified underlying representation and its fully-specified phonetic output form, since the grammar will produce an acceptable input-output mapping for such cases (i.e., it will not ‘fail’). This grammar, once hypothesized by the learner, will also be supported by (1) the fact that the *V[±round] and *V[±back] constraints have *already been demoted* relative to MAX-IO and DEP-IO, and thus their ‘re-elevation’ would be disallowed under OT learning-theoretic assumptions, and (2) by the fact that Richness of the Base requires that no constraints on input forms (such as one requiring all inputs to be underspecified) be imposed.

This ranking is, however, the *wrong* result. First, it holds that Marshallese has a fully-specified vowel for the 🕴æ🕴 of 🕴tʲætʲ🕴 as well as underspecified vowels such as the [-hi,+ATR] vowel of /nʲV_{MIT}tᵚ/. Second, it fails to account for the complete absence in Marshallese of words of the shape [kʷækʷ] with a fully-specified [æ] between [+back] and [+round] consonants.

Now let us consider the alternative view, one where the initial ranking places Faithfulness constraints above Markedness constraints. Keeping other things equal, the first data for the acquirer is the form 🕴tʲætʲ🕴. As before, the acquirer’s hypothesis at this point should be that the vowel is a fully-specified [æ]. In this case, however, nothing about the constraint ranking needs to be changed since the winning candidate will be the one that is *most faithful* to fully-specified [æ].

Upon examination of the second piece of data, 🕴nʲe͜ətᵚ🕴, the acquirer’s hypothesis should once again be the same as in the first case, that is that the vowel has height and ATR features but is underspecified on the back and round dimensions. And, as in the first case, no action with respect to the constraint ranking will be taken. As in the earlier scenario, these forms will be correctly handled by the current grammar (because Faith constraints outrank Markedness constraints, and maximally faithful outputs in this case will be underspecified).

Once again, the process of Lexicon Optimization will lead to the acquirer positing underspecification for vowels which show alternations, which reveals to a learner that even seemingly fully-specified vowels (like the 🕴i🕴 of 🕴rʲitʲerʲpᵚalᵚ🕴 ‘worker’) may be only apparently fully-specified, and can in fact be derived from an assumption of underspecified inputs. This will only be successful if the learner has not already (mistakenly) reranked Markedness constraints. Further Lexicon Optimization on the part of the acquirer uncover the fact that this is true of *all* apparently fully-specified vowels (even the non-alternating ones). This process also reveals that *all* outputs can be treated as underspecified along the back and round dimensions (even those which are apparently [+back,+round]), leading to the reranking of *V[±back] and *V[±round] relative to the MAX-IO and DEP-IO constraints. Note that this is the first change in constraint ranking under the assumption of high-ranking initial Faith.

An initial ranking of Faithfulness above Markedness constraints achieves the correct result. It allows the construction of a grammar of Marshallese in which vowels are underspecified phonetically for [back] and [round]. It also makes the correct prediction—unlike the assumption of highly ranked Markedness constraints at the initial state—about the behavior of vowels in loanwords into Marshallese, as shown in the next section.

6 Evidence from loanwords

As we have seen from the discussion above, the optimality theoretic learner, under standard assumptions, ends up positing fully-specified vowels for Marshallese in just those cases in which the surrounding consonants have identical specifications for [back] and [round]. The inevitability of this development is rooted in the articulatory/acoustic identity between the realization of Marshallese's underspecified vowels in this context (e.g., /V_{LO}/ realized as [æ] between non-back, non-round consonants) and fully-specified vowels in other languages (e.g., English /æ/ realized as [æ], regardless of surrounding consonants). Since the assumption of Richness of the Base *precludes* restrictions on the inventory of underlying segments, the learner, having (in our view, mistakenly) posited, e.g., underlying /æ/ for a given Marshallese segment, is under no pressure to modify this assumption. We mentioned one argument for why this is an undesirable result above—that the widespread and predictable distributional regularity of Marshallese vowels such as [æ] (which occurs only between 'light' consonants) would be completely unexpected.

This aspect of the problem can also be examined using a standard OT technique for exploring Richness of the Base—the study of the phonological properties of 'loanwords'. Since there is no restriction on underlying forms under the assumption of Richness of the Base, loanwords are often assumed in the literature to have the *same* underlying representation in the borrowing language as in the source language. Loanword 'adaptation' is then simply a function of playing the source language's lexical items through the constraint ranking of the borrowing language. If Marshallese speakers allow fully-specified underlying representations for vowels such as /æ/, and allow the winning candidates for such fully-specified underlying vowels to also contain fully-specified vowels (as would seem to be required by the 'standard' OT acquisition account, as sketched above), then loanwords containing such vowels in the source language should require no 'adaptation'.

In CVC loanwords with English vowels which 'match' (roughly) the pronunciation of Marshallese vowels and where the flanking consonants have the same values for the features [back] and [round] as the English vowel, the Marshallese pronunciation is, as expected, relatively similar to the English source vowel in articulatory/acoustic terms.[19]

(12) Marshallese Loanwards with Relatively Faithful Realization of 'Target' Vowels

-back,-round both	+back,-round both	+back,+round both
/$\mathrm{jV}_{\mathrm{HI}}\mathrm{t}^{\mathrm{j}}$/ [$\mathrm{jit}^{\mathrm{j}}$] 'yeast'	/$\mathrm{l}^{\mathrm{ɯ}}\mathrm{V}_{\mathrm{LO}}\mathrm{k}$/ [$\mathrm{l}^{\mathrm{ɯ}}\mathrm{ak}$] 'lock'	/$\mathrm{k}^{\mathrm{w}}\mathrm{V}_{\mathrm{HI}}\mathrm{k}^{\mathrm{w}}$/ [$\mathrm{k}^{\mathrm{w}}\mathrm{uk}^{\mathrm{w}}$] 'cook'

This is of course expected under any analysis: it is impossible to tell whether such vowels are fully specified or underspecified along the back and round dimensions in Marshallese without consideration of the full range of data.

When the flanking consonants have different values for the features [back] and [round], or when these consonants conflict with the values for the English vowel along these dimensions, a Marshallese speaker who, as in the 'standard' OT account, allows fully specified vowels in underlying and output representations, and is confronted by an English vowel which (roughly) 'matches' that found in his own output forms, would be predicted to borrow such words intact, and pronounce

[19] The loanword data cited in this section of the paper has been taken from Abo et al. (1976).

them accurately. Marshallese shows a variety of treatments of such English CVC loanwords, all of which, however, involve glide epenthesis, contrary to expectation on the ‘standard’ OT account.

Marshallese has three glides – a palatal (traditionally, /y/, here /j/), a back non-round (traditionally, /h/, here /ɰ/) and a back round (traditionally, /w/) glide, each with the expected effect on adjacent vowels. The glides themselves are said to be ‘weakly articulated’ and noticable principally through their effects on adjacent vowels. The examples below are grouped by type of flanking consonant in the loanword and by the ‘target’ vowel in the source language.

(13) Marshallese Loanwords with Seriously Divergent Realizations of the ‘Target’ Vowel

a. flanking back consonants

front V target	round vowel target
/kV$_{MID}$jV$_{MID}$k/ 🚹kə͡ejə͡ək🚹 ‘cake’	/tᵚV$_{HI}$wV$_{HI}$lᵚ/ 🚹tᵚɯ͡uwu͡ɯlᵚ🚹 ‘tool’
/kV$_{LO}$jV$_{LO}$ŋ/ 🚹ka͡æjæ͡aŋ🚹 ‘gang’	/kV$_{MID}$wV$_{MID}$tᵚ/ 🚹kə͡owo͡ətᵚ🚹 ‘goat’

b. flanking front consonants

back non-round V target	round V target
	/tʲV$_{HI}$wV$_{HI}$nʲ/ 🚹tʲi͡uwu͡inʲ🚹 ‘June’
	/tʲV$_{HI}$wV$_{HI}$tʲ/ 🚹tʲi͡uwu͡itʲ🚹 ‘shoes’

c. mixed flanking consonants

front V target	back non-round V target
/kʷV$_{HI}$jV$_{HI}$nʲ/ 🚹kʷu͡ijinʲ🚹 ‘queen’	/tʲV$_{LO}$ɰV$_{LO}$tᵚ/ 🚹tʲæ͡aɰatᵚ🚹 ‘shot’
/tʲV$_{LO}$jV$_{LO}$k/ 🚹tʲæjæ͡ak🚹 ‘check’	
/pʲV$_{LO}$jV$_{LO}$k/ 🚹pʲæjæ͡ak🚹 ‘back, bag’	
/pᵚV$_{MID}$ɰV$_{MID}$tʲ/ 🚹pᵚəɰə͡etʲ🚹 ‘base’	

round V target
/tʲV$_{LO}$wV$_{LO}$pᵚ/ 🚹tʲæ͡ɒwɒ͡apᵚ🚹 ‘soap’
/tʲV$_{LO}$wV$_{LO}$k/ 🚹tʲæ͡ɒwɒ͡ak🚹 ‘chalk’

Two facts emerge from a consideration of this loanword data from Marshallese. First, the glide-insertion process appears to take place just in the case where an underspecified Marshallese output vowel would be realized—because of the effects of the surrounding consonants—as a vowel quite distinct from that of the source language. For example, English /ʃuz/ without glide-insertion would be, under our assumptions, stored as Marshallese /tʲV$_{HI}$tʲ/ and pronounced as 🚹tʲitʲ🚹. Note that under ‘traditional’ OT assumptions, the word would presumably be stored as Marshallese /tʲutʲ/ and pronounced as 🚹tʲutʲ🚹. Neither of these reflect the observed data: apparently the Marshallese speaker *cannot* avail himself of the ‘standard’ OT analysis, and *does not* allow an analysis which leads to such a divergent realization as in the non-glide insertion analysis. It would appear that the most plausible analysis of the observed loanword phenomenon involves not the automatic positing of a Marshallese underlying form which matches precisely that of the English source (as is commonly assumed in approaches to these questions), but instead some consideration of how a Marshallese speaker might parse—i.e., assign a linguistic analysis to—a particular English word. In the 'shoes' case, for example, what the Marshallese speaker appears to do is analyze the voiceless coronal fricatives as realizations of his/her /ţʲ/.[20] The [+back] and [+round] properties of the English vowel in [ʃuz] are of course very salient, and the Marshallese speaker needs to posit a parse which accounts for the presence of these properties in the string. Since the non-nasal coronal stops of Marshallese do not include a [+back], [+round] phoneme, and since [back] and [round] specifications arise in Marshallese only through the presence of consonants in the representation, the speaker posits the existence of a back, round glide (/w/) in his/her parse to account for the vowel realization.[21] By inserting a glide which triggers backing and rounding of the high vowel in 'shoes', the result is a vowel a significant portion of the duration of which is accurate given the English target. This leads us to our second point—it appears that mechanically adopting underlying representations from the source language, the process generally assumed in much of the OT discussion of loanword phenomena, is insufficient to capture the somewhat more subtle factors which shape contact-related phonological phenomena.

At first blush, it may appear that this story about how Marshallese loanwords come to have glide-insertion is in direct violation of the principle of Richness of the Base. After all, why doesn't the borrower simply posit either an underlying /u/, or underlying back, round coronal stops, or, for that matter, /ʃ/ and /z/, in his/her representation of English /ʃuz/, as the standard OT approach to these matters appears to assume? But we would like to maintain an in our view important contrast between what the borrower actually does in a particular instance of borrowing—which is a function of the nature and intensity of contact, the sociolinguistic situation, and potentially many other factors—and what the *representational capacity* of the borrower is. Under the right set of (essentially extralinguistic) circumstances, it may be quite possible for Marshallese speakers to posit representations containing /u/'s, or /ʃ/'s, or whatever. Our claim is merely that those circumstances did not prevail at the time of the borrowing of /ʃuz/ (or any other loanwords in Marshallese from English known to us).

7 Conclusions

With these Marshallese facts, we hope first to have contributed to the literature on the need for underspecification that persists through to phonetic form. It is already well-known that an abstract, feature-based representation system requires that *some* aspects of articulation not be specified (e.g., the transitions from [g] to [ɛ] in 'get'), because the grammar provides no appropriate representational apparatus. In parallel fashion, we assume that one of the cues to the acquirer in the Marshallese case is precisely the absence of any UG-given feature available to be assigned to some input form, resulting in *no* feature being assigned along the relevant dimension. Specifically, as we discussed earlier, one source of variation in the realizations of identical feature bundles that is stripped away for purposes of phonological representation is the transitional effect determined

[20] This is a consistent pattern, perhaps aided by the fact that Marshallese /ţʲ/ often has affricated realizations.

[21] Presumably the glide is posited, as opposed to some other 'heavy' consonant (e.g., /kʷ/) because it has, aside from its effects on adjacent vowels, the least salient acoustic cues, i.e., it is the segment most likely to be diffi cult to perceive beyond its effect on vowels.

by context (co-articulation effects). Crucially, in Marshallese, the transitions last *throughout the duration of the vowel* for backness and roundness in the appropriate consonantal environment and are treated, for purposes of featural representation, just as all other transitions are, with no features for back and round being assigned. More interestingly, perhaps, Marshallese teaches us that the grammar may similarly leave some aspects of articulation not specified (e.g., the backness of the tongue and roundedness of the lips during the vocalic segments of a string) for which the representational apparatus *is*, in principle, available. This lends further support for the idea that the symbolic representations manipulated by the grammar should be divorced from the articulatory and acoustic dimensions. The Marshallese case illustrates that no necessary connection between a particular articulatory/acoustic event and a given feature holds.

The Marshallese case also reveals several shortcomings in Optimality Theory, as traditionally practiced. The first of these is that formal notions of markedness and those based on typology, or cross-linguistic frequency, do not converge on the same set of 'Markedness constraints' — the latter should therefore not be used in arguing in support of the former (and vice-versa). Second, the most reasonable account of the acquisition path for Marshallese appears to require an initial state in which Faithfulness constraints are ranked above Markedness constraints, contrary to widespread OT practice. Finally, we have presented some reasons to believe that the current approach to loanword phonology within OT, particularly the invocation of loanword phenomena in support of the concept of Richness of the Base, is overly optimistic about the assumed simplicity of the mechanisms involved in 'language contact' events.

References

Abo, Takaji, Byron W. Bender, Alfred Capelle, and Tony DeBrum. 1976. *Marshallese-English Dictionary*. University of Hawaii Press, Honolulu.

Bender, B. 1968. 'Marshallese Phonology.' *Oceanic Linguistics* 7:16-35. Browman, C. and L. Goldstein. 1990. Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics* 18:299-320.

Choi, J. 1992. 'Phonetic Underspecification and Target Interpolation: An Acoustic Study of Marshallese Vowel Allophony.' *UCLA Working Papers in Phonetics* 82.

Hale, M. 2000. 'Marshallese Phonology, the Phonetics-Phonology Interface and Historical Linguistics.' *The Linguistic Review*. 17:241-257.

Hale, M. and M. Kissock. 'Underspecification: Neutralization, and Acquisition' paper presented at the 13th Manchester Phonology Meeting, May 26-28, 2005.

Hale, M., M. Kissock and C. Reiss. (to appear) 'Microvariation, Variation, and the Features of Universal Grammar,' *Lingua*.

Hale, M., M. Kissock and C. Reiss. 'Underspecification, Markedness, and the Acquisition Process,' paper presented at the 6e Journées Internationales du Réseau Français de Phonologie, Orléans, June 3-4, 2004.

Hale, M., M. Kissock and C. Reiss. 2002. 'Phonological Accounts of Microvariation in Pronunciation,' paper presented at the 25th GLOW Colloquium, Amsterdam, April 9-11.

Hale, M. and C. Reiss. 2000 'Phonology as Cognition.' in Burton-Roberts, N., P. Carr, and G. Docherty (eds.) *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: Oxford University Press.

______. 1998. Formal and Empirical Arguments Concerning Phonological Acquisition. *Linguistic Inquiry* 29:656-683.

Hammarberg, R. 1976. The Metaphysics of Coarticulation. *Journal of Phonetics* 4, 353-363.

Inkelas, S. 1994. 'The Consequences of Optimization for Underspecification,' in Beckman, J. (ed.) *Proceedings of the North East Linguisitc Society 25* pp. 287-302.

Inkelas, S., O. Orgun, and C. Zoll. 1997. 'The Implications of Lexical Exceptions for the Nature of Grammar,' in Roca, I. (ed.) *Derivations and Constraints in Phonology*. New York: Oxford University Press. 393-418.

Ito, J., A. Mester, and J. Padgett. 1995. 'Licensing and Underspecification in Optimality Theory,' *Linguistic Inquiry* 26, 4: 571-613.

Jusczyk, P., and P. Luce 2001. Speech perception and spoken word recognition: past and present. *Ear and Hearing*, to appear.

Kager, R. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.

Kager, R., J. Pater, W. Zonnenveld (eds.). 2004. *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press.

Keating, P. A. 1988. 'Underspecification in Phonetics.' *Phonology* 5:275-292.

McCarthy, J. 2002. *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.

Pulleyblank, D. 2001. Defining features in terms of complex systems: Is UG too complex?. Paper presented at the Workshop on Early Phonological Acquisition, Carry-le-Rouet, France.

Pylyshyn. S. 1984. *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, MA: MIT Press.

Steriade, D. 1995. 'Underspecification and Markedness,' in Goldsmith, J. (ed.) *Handbook of Phonological Theory*, Oxford: Blackwell, 114-174.

Stevens, K. N. 1998. *Acoustic Phonetics.* Cambridge, MA: MIT Press.

______. 1972. 'The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data.' in Denes, P. and E. David Jr. (eds.) *Human Communication: A Unified View.* New York: McGraw Hill.

Tesar, B. and P. Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge, MA: The MIT Press.

Tesar, B., J. Alderete, G. Horwood, N. Merchant, K. Nishitani, & A. Prince. 2003. 'Surgery in Language Learning,' in Garding, G. and M. Tsujimura (eds.) *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, 477-490.