
ELSEVIER
Volume 38
Issue 1
January 2010
ISSN 0095-4470
Journal of
Phonetics
Special Issue: Phonetic Bases of Distinctive Features
Guest Editors:
Editor
Stefan Frisch
Associate Editors
Ocke-Schwen Bohn
Jennifer Hay
Kenneth de Jong
Alice Turk

Journal of Phonetics 38 (2010) 44–59

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/phonetics

Distinctive features: Phonological underspecification in representation and processing

Aditi Lahiri [a,*], Henning Reetz [b]

[a] University of Oxford, Faculty of Linguistics, Philology and Phonetics, Centre for Linguistics and Philology, Walton Street, Oxford OX1 2HG, UK
[b] Goethe University Frankfurt, Institute for Linguistics, Georg-Voigt-Str. 6/II, 60325 Frankfurt am Main, Germany

A R T I C L E I N F O

Article history:
Received 27 April 2008
Received in revised form
4 January 2010
Accepted 6 January 2010

A B S T R A C T

Phonological variation of any sort (determined by speech styles, phrasing, or morphophonological rules) affecting the shapes of words and morphemes are a matter of concern for theories of speech perception and language comprehension. To come to grips with parsing the speech stream, accessing the lexicon and ultimately recognizing words, both representational as well as processing issues must be considered. The central questions in the research presented here are: WHAT is represented in the mental lexicon? HOW is it represented? HOW is the speech signal parsed and information mapped onto the mental lexicon? In this paper we will address four issues within the framework of our *Featurally Underspecified Lexicon* model (FUL): (a) our assumptions concerning distinctive feature organization defined by phonological, perceptual and acoustic constraints; (b) specification of features in the mental lexicon (based on universal and language specific requirements); (c) extracting distinctive features from the signal; (d) mapping features from the signal to the lexicon. We claim that phonological features are extracted from the variable acoustic signal based on broad acoustic properties. A three-way matching algorithm maps these features onto highly abstract phonological mental representations. We provide evidence from synchronic phonological analyses, language change, psycholinguistic and neurolinguistic data.

1. Introduction

Phonologists generally assume single underlying morphemes when morphophonological alternation is governed by phonological rules and constraints. The surface output is produced by the application of rules or chosen by the interaction of constraints. Thus, alternations like German *Hund* [t]~*Hunde* [d] 'dog SG/PL' would be accounted for by assuming a single representation /hʊnd/ with a final devoicing rule applying word finally. If we assume the same representation for comprehension, we need a model which would show us how the two surface forms [hʊnt] and [hʊndə] have to tap the underlying form. Other than morphophonological alternations, surface phonetic variation occurs all the time. Here as well, if one assumes unique underlying forms, we need an algorithm for matching the surface to the underlying forms. In this paper we would like to explore a model of language comprehension, *Featurally Underspecified Lexicon* (FUL), which makes precise predictions about the nature of phonological representations and mapping algorithms from surface to underlying mental representations. We begin by providing an in-depth discussion of phonological features and feature representational hypotheses, and consider the consequences of underspecified representations. We then briefly discuss our earlier experimental results in the context of other models of language comprehension and move on to provide further evidence for our model based on a perception experiment involving morphophonological alternations of vowels in German.

Phonological contrasts and distinctive features have been the hallmark of phonological analyses for at least half a century. Jakobson, Fant, and Halle (1952) began with 21 distinctive binary features which were intended to capture all the phonological systems of natural languages. The features had well defined perceptual, acoustic and articulatory correlates and, crucially, the same features classified vowels and consonants. For instance, *acute* consonants (dentals, palatoalveolars and palatals) were characterized as having high frequency energy, which was also true for *acute* front vowels like [i, e, æ]. The focus changed to articulatorily oriented features a decade later in the *Sound Pattern of English* (Chomsky & Halle, 1968), the argument being that these features were better suited to describe phonological patterns particularly with [±back], which was introduced in addition to [coronal]. A crucial divergence was the establishment of separate

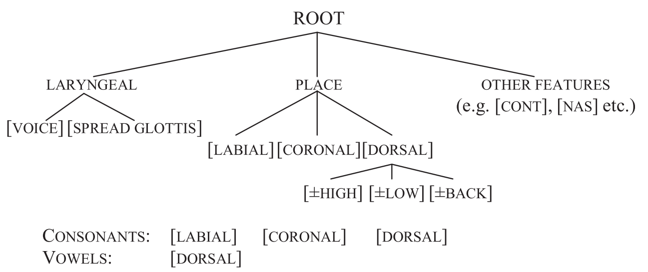
* Corresponding author. Tel.: +44 1865 280401; fax: +44 1865 280412.
E-mail addresses: aditi.lahiri@ling-phil.ox.ac.uk (A. Lahiri), reetz@em.uni-frankfurt.de (H. Reetz).

0095-4470/$ - see front matter © 2010 Published by Elsevier Ltd.
doi:10.1016/j.wocn.2010.01.002

place features for vowels and consonants. Vowels, for example, were all [– anterior] while consonants could be both [± anterior]; vowels were characterized by [± back] and were always [– coronal]. Thus, there was no correlation between the [coronal] consonants like dentals and palatoalveolars and [– back] vowels. Notwithstanding the move towards articulatorily oriented features in phonology, research continued on the acoustics of features by Stevens, Blumstein and colleagues (cf. Blumstein & Stevens, 1980; Lahiri, Gewirth, & Blumstein, 1984; Stevens & Blumstein, 1978), the goal being to locate invariant acoustic cues for distinctive features rather than for segments, which had proved to be impossible. Spectral cues for velars, for instance, were inseparable from the neighbouring vowel context, making it impossible to define unique cues for these consonants independent of the vowels (cf. Jusczyk, 1986). This lack of invariance for place of articulation cues for consonantal phonemes was one of the critical reasons that led to the foundation of the influential motor theory of speech perception (Liberman & Mattingly, 1985). The invariance theory, in comparison, was linked to features (i.e. cues for groups of sounds) rather than to individual segments (cf. Lahiri et al., 1984, for cues to distinguish CORONAL and LABIAL *diffuse* stops).

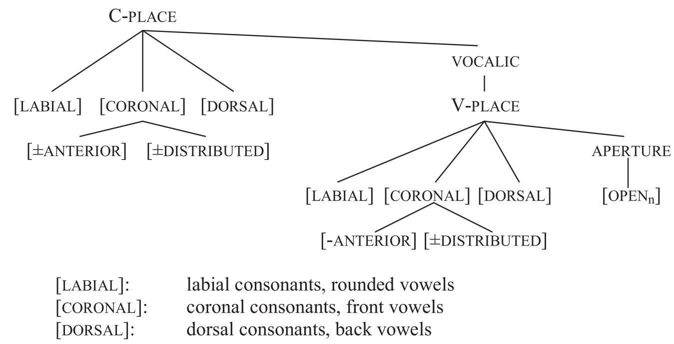
Within phonological theoretical analyses, the eighties led the way to grouping features into natural classes rather than listing them indiscriminately (Clements, 1985, 1989; McCarthy, 1988; Sagey, 1986). Hypotheses concerning feature-groups were made largely on the basis of phonological rules. For example, if features spread as a cluster, there was clear motivation for grouping them under one node, as the PLACE node or LARYNGEAL node (McCarthy, 1988). Although controversies raged over the precise grouping of features, one analysis remained dominant, viz. vowels and consonants were no longer grouped under the same place features (cf. Sagey, 1986). Vowels came under the DORSAL node and were distinguished by [± back]. That is, front vowels, which fell together with 'front' consonants under the feature *acute* in Jakobson et al. (1952), were now fully segregated from the dentals, alveolars and palatoalveolars, which were grouped under [coronal]. Additionally, however, the height features [± high], [± low] were dominated by [dorsal]. The feature tree in (1) provides an approximation of the features modified from the Halle–Sagey model (combining Sagey, 1986, and Halle, 1995, as in Halle, Vaux, & Wolfe, 2000).[1]

(1) Established class nodes (Halle–Sagey)



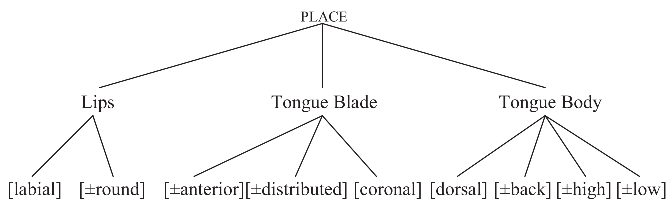
Clements (1989) brought forth a novel proposal arguing that descriptions of vowels and consonants should and could be unified together if one seriously considered the notion of constriction of the vocal tract with the parameters degree and location. Under this proposal, the constriction of consonants would be represented by the oral cavity node, the degree being characterized by [± continuant], and the location by the place node. The constriction location of vowels would be similar to that of consonantal place node, but the constriction would be represented by the vocalic node, and the degree by an aperture node dominating [± open], which can be arrayed on different tiers to convey various degrees of height. The place nodes for vowels and consonants would be on separate tiers usually designated as V-place and C-place respectively.

(2) Feature tree based on Clements and Hume (1995)



A fundamental difference between Clements' model and Halle–Sagey's model is that [coronal] entirely replaced [± back].[2] In response to Clements' unified theory, Halle et al. (2000) revised the Halle–Sagey model to dispense with dependencies such that [back] [high] [low] were no longer dependents of DORSAL.[3] Thus, any fronting that would spread [dorsal] would not necessarily spread [– back]. Nevertheless, vowels and consonants still do not share the same features.

(3) Feature organization based on Halle et al. (2000)



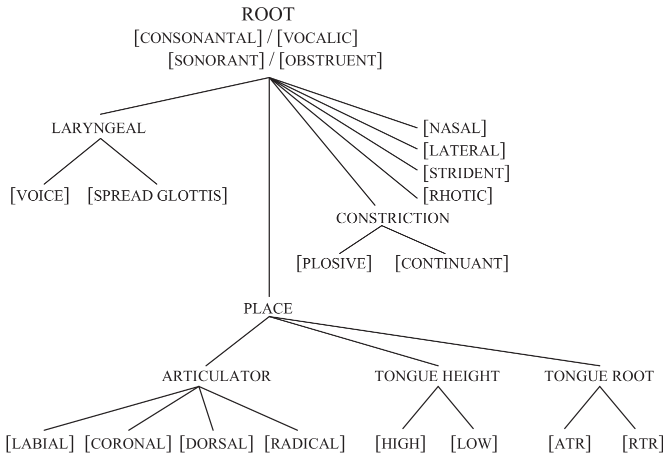
Inspired by the work of Clements, and reverting back to Jakobson et al.'s view of combining all consonantal and vocalic features, Lahiri and Evers (1991) and later Lahiri and Reetz (2002) in their model of the Featuraly Underspecified Lexicon (FUL, cf. also Ghini, 2001a, Lahiri, 2000) took it one step further and argued that there was no necessity to duplicate the V-place node for vowels and secondary articulations, and that the aperture node was not only relevant for vowels but also for consonants. The idea was that the constriction relevant on the horizontal dimension along the vocal tract was determined by the ARTICULATORS, and on the vertical dimension was characterized by the height of the tongue; the output of both could be defined by broad acoustic cues. Consequently, the place node dominated two separate nodes ARTICULATOR and APERTURE or TONGUE HEIGHT. The place features were therefore identical for vowels and consonants. The features and feature organization we will defend are based on universal principles of phonological alternations as well as perceptual mechanisms.

[1] The tier structures in the feature trees (1)–(4) are not relevant for this discussion.

[2] Hyman (1973) and Lahiri and Blumstein (1984) argued for a revival of the feature *grave*.

[3] The feature tree given in Halle et al. (2000, p. 389) does not indicate ± values. However, during the discussion of Irish assimilation, it is obvious that as before the features [HIGH, LOW, DISTRIBUTED, ROUND, ANTERIOR, BACK] are binary.

(4) Feature organization in FUL



These are all the features that are required to express segmental contrasts in the languages of the world. The feature [coronal] is always underspecified in the mental lexicon, and other features may also be underspecified depending on feature assignment which we discuss in the following sections.

There are two pairs of opposing binary features – consonantal or vocalic and sonorant or obstruent – which are the major class features available in all languages. The members of each pair are conflicting — i.e., consonantal implies *not* vocalic and vice versa. There are other features like high and low which are mutually exclusive, but these are not binary. This is because a vowel, for instance, cannot be both high and low, but it may be neither. The truly binary features do not have this possibility: a segment must be either consonantal or vocalic, and sonorant or obstruent. All other features are monovalent. The only dependencies we assume are universal and must be listed: [nasal] ⇒ [sonorant], [strident] ⇒ [obstruent] and [constriction] ⇒ [obstruent]. We assume that [high] or [low] can differentiate the various coronal consonants (dental, palatoalveolar, retroflex, etc.) instead of [± anterior]. A partial list of segment classification is given below.

(5) Features and segments

[labial]	labial consonants, rounded vowels
[coronal]	front vowels, dental, palatal, palatoalveolar, retroflex consonants
[dorsal]	back vowels, velar, uvular consonants
[radical]	pharyngealized vowels, glottal, pharyngeal consonants
[high]	high vowels, palatalized consonants, retroflex, velar, palatal, pharyngeal consonants
[low]	low vowels, dental, uvular consonants
[atr]	palatoalveolar consonants
[rtr]	retroflex consonants

It is worth noting that palatal and retroflex consonants are both [coronal] and [high] and therefore they are not distinguishable by these features alone. Three types of consonantal contrasts could be potential problems: palatal versus retroflex stops /c ʈ/ and nasals /ɲ ɳ/, and palatoalveolar versus retroflex sibilants /ʃ ʂ/. According to FUL, various different features keep these consonants apart as described below.

- First, palatals are definitely coronal (Keating & Lahiri, 1993; Lahiri & Blumstein, 1984); but we propose that a palatal versus retroflex underlying contrast in stops is only possible if the palatal stop is affricated or is an "alveopalatal" consonant, both of which would be [strident] (cf. Hall, 1997), or if one is derived from the other. In Malayalam, which has both retroflex and palatal stops, only retroflex stops occur in the underlying inventory. The palatal stops are derived in specific morphological environments from intervocalic velars when preceded by front vowels.[4]
- Second, if a language has both palatal and retroflex nasals the claim is that they cannot be truly contrastive. Either the palatal nasal /ɲ/ would consist of a nasal+glide sequence, or it would be an assimilated variant of an alveolar or dental /n/ in the context of a palatal or palatoalveolar stop, or the retroflex nasal is derived. Again, Malayalam is a good example since it has seven phonetic nasals derived from three underlying ones which are labial, dental and velar /m n ŋ/ (Mohanan & Mohanan, 1984, pp. 583–586, 596–598). Both palatal and retroflex nasals are derived by a homorganic nasal assimilation rule in the context of following palatal and retroflex stops, and the palatal stops are in turn derived from velars (see above).
- As for fricatives, the palatoalveolar [ʃ] in Malayalam (labelled as [s̄] in Mohanan & Mohanan, 1984) is also derived from palatalization and does not contrast with the underlying retroflex [ʂ]. However, although rare, the contrasts between retroflex and palatoalveolar sibilants are more frequent than such a contrast in stops. We would argue that to distinguish these consonants the features [atr]/[rtr] are used, where the palatoalveolar [ʃ] would be [atr].

1.1. Feature specification and underspecification

Other than the feature inventory itself, the related issue is lexical feature specification, and corresponding underspecification where some features might not be part of the lexical representation but can be generated by general or morpheme-specific production rules. Recent phonological research on features has predominantly stated that underspecification is not only unnecessary, but also misguided. Consequently, as Halle et al. (2000) emphasize, full specification is considered to be the norm, at least for features that are contrastive in any given language. Nevertheless asymmetries and markedness differences exist across features, feature distribution, and direction of phonological rules. These are dealt with by various methods. For instance, Calabrese (1995) proposes different types of feature representations, contrastive (determined by specific algorithms), marked and full, interspersed in the rule ordering. Mohanan (1993) advocates "fields of attraction" and dominance which provide a means of expressing different degrees of markedness. Although there is a strong objection to underspecification, most models do not specify non-contrastive features. Here we turn to Clements (2001) who presents a sophisticated model of specification and underspecification specifying non-contrastive features if they are active in phonology.

Clements proposes a model of feature representation that distinguishes between active features (which may refer to natural classes) and prominent features (which, for instance, play a role in spreading). Although he accepts coronal transparency, he points out

[4] Mohanan and Mohanan (1984) also suggest that given the complex conditioning of the palatalization rule, "Perhaps the right solution is to say that Palatalization is blocked when the segment has some ad hoc diacritical feature [-P]" (p. 589). Mohanan and Mohanan also make a distinction between *underlying* and *lexical* alphabet, the latter being derived by rules in the lexicon. Their claim is that the lexical alphabet "has significant consequences for human perception of speech sounds" (p. 598). It could be the case their lexical contrasts and our underlying contrasts would be the same. But this would require a complete rule-by-rule comparison, which is not possible here.

that "analyses accounting for coronal transparency in terms of coronal underspecification have tended to become discredited" (pp. 114–115), and presents an approach where coronal only becomes available when necessary. This works in the following way. In general, the feature [coronal] is absent from lexical specifications, "following the principle that unmarked features and feature values are lexically unspecified." However, [coronal] becomes specified in the phonology *if* it is available as a term in constraints and is projected to a separate tier if it is phonologically prominent. And thus, coronal transparency can be accounted for either by its absence where it is inactive and "by the nonprojection of [coronal] (in segment classes in which the feature is active but unprojected)" (p. 115). We discuss further details of his approach below. Our approach vitally does assume underspecification, which is built into phonological systems in an organized fashion. And where coronality is concerned, this feature is always underspecified. How, then, does one assign features in a model assuming underspecification? Following Ghini (2001b), we accept the notion "place first"; that is, a child acquiring a language will assign ARTICULATOR features first. Furthermore, based on Levelt (1995) and Fikkert and Levelt (2008), the assumption is that [LABIAL] is acquired as the specified feature in contrast to [CORONAL], a contrast which is assumed to exist in every language. Thus, the [CORONAL]/[LABIAL] contrast is established first, with [LABIAL] specified and [CORONAL] always remaining underspecified. It is listed in the feature geometry because the feature always exists and is filled in articulation. The contrasts [CONSONANTAL]/[VOCALIC] and [OBSTRUENT]/[SONORANT] are present in all languages. All other features depend on the phonological systems of individual languages. The assignment does not depend on whether any feature is active in a phonological rule, but only if it is necessary to establish a phonemic contrast. Furthermore, given the assumption of privative features, the absence of a feature maybe considered to be underspecification, but unlike coronal underspecification, such a feature is not realized in production. Consider the following examples:

- A feature like [VOICE] will be specified if the language contrasts a set of voiced vs. voiceless segments. Voicelessness does not exist as a feature and hence consonants which are not voiced will remain unspecified in production and this feature cannot be extracted by the perceptual system.
- Like [VOICE], features such as [NASAL], [STRIDENT], [LATERAL], etc. are specified if a contrast is established. The absence of these attributes plays no role in production or in perception.

To show how coronal transparency works in Clements (2001) as compared to our approach, we will draw on the example of Tahltan coronal harmony as analysed in Shaw (1991, pp. 144–152) and reconsidered in Clements. Shaw shows that Tahltan, an Athapaskan language with five series of coronal obstruents (cf. 6), has a coronal harmony process involving only three sets—apical, laminal and palatoalveolar coronals. Fricatives of these places of articulation assimilate to all coronal place features and stridency of any following coronal obstruent of one of these three sets.

(6) Tahltan coronal obstruents

simple	lateral	apical	laminal	palatoalveolar
d	dl	ʣ	dð	ʤ
t	tɬ	ts	tθ	ʧ
t'	tɬ'	ts'	tθ'	ʧ'
	ɬ	s	θ	ʃ
	l	z	ð	ʒ

The simple and lateral series are transparent to this harmony process and are not triggers or targets of the assimilation. Clements points out that these facts are impossible to account for in most other models without ad hoc rules, and argues that Shaw's original analysis assuming coronal transparency remains unrivalled. Shaw highlighted the fact that the coronal consonants which play a role in harmony must be specified for some marked value. The marked coronal features that were used by Shaw (1991) and Clements (2001) for the obstruents involved in the harmony are given below along with the features that would be utilized by FUL. Shaw and Clements defend their respective features and feature geometry in detail in their publications. What is relevant here is the way in which the harmony facts are accounted for in a language where there are five series of coronal obstruents but only three of the series are engaged in the harmonizing process. Crucially, Tahltan has more than one series of 'transparent' coronal obstruents.

(7) Relevant features for Tahltan coronal obstruents involved in the harmony process: lat = lateral, strid = strident, distr = distributed, ant = anterior, apic = apical, post = posterior

	Shaw ROOT				Clements ROOT				FUL ROOT			
	lat	coronal			lat	coronal			LAT	STRID	Tongue Height	
		strid	distr	ant		strid	apic	post			HIGH	LOW
d												
dl	+				+				√			
s ʣ		+				+				√		√
θ dð			+				–			√		
ʃ ʤ				–				+		√	√	

- Tongue height features [HIGH] and [LOW] will be both specified only if the language has a three-way height difference. Following Ghini (2001a), we will assume that [LOW] is specified first if a two-height system is established (cf. also Kabak, 2007 for Turkish). In a three-height system, mid-vowels are only a descriptive attribute and there is no feature *mid* which can be extracted from the signal.

In Shaw (1991) the [+strident], [+distributed] and [–anterior] are dominated by the coronal node and distinguish the three relevant series. The simple [d]-series, specified as [-continuant], and the lateral series as [+lateral] remain immune to the harmony process. Likewise, in Clements (2001), the features [strident], [apical] and [posterior] are dominated by the [coronal] node and only the marked feature values are specified, namely [+strident],

[−apical] and [+posterior]. Under Clements' analysis, the simple and the lateral series are unmarked for coronal, which is neither lexically specified nor active, and hence absent. Note that for both Clements and Shaw, [strident] is a property of coronal obstruents.

In FUL, all coronal consonants are unspecified for the ARTICULATOR node. Thus, the features separating the three crucial series cannot be distinguished by any dominating coronal node. That is, one cannot claim to have a set of consonants with specified coronal features and other sets which are not specified. Before we discuss the features distinguishing these consonants, it is important to note that Tahltan appears to have four series of affricates and fricatives. The only true stops belong to the simple series. Even in the lateral series, the consonants are not sonorants but obstruent affricates and fricatives. We view stridency as a feature which is not relevant for any particular place of articulation, but is a property which distinguishes between classes of affricates and fricatives. Consequently, the crucial distinction between the lateral, apical, laminal and palatoalveolar seems to be that the last three sets are STRIDENT and the lateral is not. Within FUL, following the principles of Jakobson et al. (1952), STRIDENT distinguishes higher degree of acoustic noise; consequently, the lateral fricatives and affricates would be released with noise, but less than the others.[5] For the harmonizing series, the TONGUE HEIGHT node is utilized to distinguish between them. The palatoalveolars are specified as [HIGH], the apical obstruents are specified as [LOW], and the laminal consonants would be unspecified for height. Recall that [STRIDENT] is independent of the ARTICULATOR node (see 4). In terms of acoustic characteristics, the TONGUE HEIGHT (HIGH, LOW) features are differentiated as follows:

(8) HIGH vs. LOW

LOW:	high F1 for vowels; concentration of more energy at higher frequencies
HIGH:	low F1 for vowels; concentration of more energy at lower frequencies

Now we turn to the harmony facts, which are summarized below. Note that during the harmony process fricatives and affricates retain their release characteristics and voicing remains unchanged.

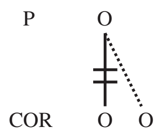
(9) Tahltan coronal harmony: target is within square brackets; and the trigger is underlined (data from Shaw, 1991)

/-s/ '1st sg. subject marker' /s/ > /θ/, /ʃ/				
(a)	mεθε[s]εθ	mεθε[θ]εθ	I'm wearing (on feet)	*fricative trigger*
(b)	na[s]tθ'εt	na[θ]tθ'εt	I fell off (horse)	*affricate trigger*
(c)	dε[s]kʷʊθ	dε[θ]kʷʊθ	I cough	*intervening syllable*
(d)	xaʔε[s]t'aθ	xaʔε[θ]t'aθ	I'm cutting the hair off	*intervening simple t'*
(e)	ε[s]ʤɪni	ε[ʃ]ʤɪni	I'm singing	*voiced affricate trigger*
(f)	ya[s]tɬεʧ	ya[ʃ]tɬεʧ	I'm singing	*intervening lateral tɬ*
(g)	ε[s]dan	ε[s]dan	I'm drinking	*no change*

/-θ/ '1st dual subject marker' /θ/ > /s/, /ʃ/				
(h)	u[θ]iʤε	u[ʃ]iʤε	we are called	*voiced affricate trigger*
(i)	dε[θ]it'as	dε[s]it'as	we are walking	*intervening simple t'*
(j)	xa[θ]iːdεts	xa[s]iːdεts	we plucked it	*intervening simple d*

Shaw's rule for the harmony process states that the rightmost specified coronal node spreads leftwards with the concomitant delinking of the previous coronal specification of the target. The target is an immediately adjacent specified coronal node. Since both trigger and target need to be specified coronal nodes, the lateral and simple series are unaffected by the spreading and cannot block harmony. The rule is given below indicating spreading of the rightmost coronal node leftward with concomitant deletion of the previous coronal specification of the target.

(10) Shaw's Tahltan harmony (p. 146)



Clements achieves the harmony effects very elegantly with a single AGREE constraint which states that "within the word, [coronal] nodes bearing marked feature values must be identical". The constraint ensures that the coronal features of the left marked consonant would match with the right consonant and consequently the unmarked simple and lateral series remain immune to harmony as shown in (9j). As Clements' argues, there can be no spreading, otherwise, a segment like /d/ would acquire the features of the following coronal consonants and will not be 'transparent'. Examples of how the constraint applies are given below. Note that agreeing with the rightmost consonant automatically means that the coronal feature of the preceding consonant is replaced. Under this analysis there is no need to delete or delink existing features.

(11) Harmony based on Clements (2001) (the intervening segments are in bold and the trigger is underlined)

θ	**d**	ʧ	→	ʃ	**d**	ʧ
[cont] [coronal] [−apical]		[coronal] [+posterior]		[cont] [coronal] [+posterior]		[coronal] [+posterior]
s	**tɬ'**	ʧ	→	ʃ	**tɬ'**	ʧ
[cont] [coronal] [+strident]		[cont] [coronal] [+posterior]		[cont] [coronal] [+posterior]		[coronal] [+posterior]

Although details differ, both Shaw and Clements achieve the harmony process by ensuring that the simple [d] and the lateral [dl] series are free of the coronal node, while the other series require features which are dominated by the coronal node. As we mentioned above, such a separation is not possible within FUL since all of these consonants are considered to be unspecified for coronal in the underlying representation. Their ARTICULATOR node remains unspecified. Although STRIDENT is necessary, it is also independent of coronality similar to the proposal in Chomsky and Halle (1968). Under our analysis, the harmony is restricted to those obstruents specified for STRIDENT.

[5] As always, the acoustic correlate of a feature may vary in degree from language to language. For instance, [high] could classify all vowels with F1 $<$ 400 Hz in one language but $<$ 450 in another depending on the inventory. Similarly, the degree of noise required for a fricative to be classified as [strident] may vary. In English, [θ] would not be [strident] but in Tahltan it is.

(12) Tahltan Harmony in FUL

Within a word, obstruents specified for STRIDENT must agree with the TONGUE HEIGHT features of any following strident.

The same result can also be achieved by spreading the Tongue Height node. Since the simple and lateral series are not specified for STRIDENT, these consonants remain immune to the harmony.[6] Some examples of harmony are given below.

(13) Harmony examples in FUL with and without intervening consonants (the intervening segments are in bold and the trigger is underlined)[7]

(i)	θ [–] [strid]	**d**	t͡ʃ [high] [strid]	→	ʃ [high] [strid]	**d**	t͡ʃ [high] [strid]
(ii)	s [low] [strid]	**tɬ’** [lateral]	ʃ [high] [strid]	→	ʃ [high] [strid]	**tɬ’** [lateral]	ʃ [high] [strid]
(iii)	s [low] [strid]	**d**	θ [–] [strid]	→	θ [–] [strid]	**d**	θ [–] [strid]
(iv)	θ [–] [strid]	**d**	ts [low] [strid]	→	s [low] [strid]	**d**	ts [low] [strid]
(v)	s [low] [strid]	θ [–] [strid]		→	θ [–] [strid]	θ [–] [strid]	

Neither series would be specified for [CORONAL]. Consequently, in (13iv), [LOW] can spread across /d/, which is unspecified for height. In (13ii), in the case of [s→ʃ], [HIGH] spreads across /tɬ’/, which is unspecified for [CORONAL], and the feature [LATERAL] is independent of the tongue height node and does not block any feature spreading. For [s→θ], the feature [LOW] is deleted in the context of an unspecified fricative.

Before ending this section, we will briefly touch on the differences between FUL and earlier approaches as radical underspecification (cf. Archangeli, 1988) and contrastive underspecification (cf. Steriade, 1995), two views which were heatedly discussed in the literature. Broadly speaking, these models differ crucially in their treatment of unmarked values. If [±high] differentiates two sounds, then only the marked value [−high] will be specified in radical underspecification, while both will be specified in contrastive underspecification. For the latter, only redundant features are not specified such as [+voice] for sonorants. Direct comparisons between FUL and these models can be rather fallacious, since most of the discussions concerning them were carried out using binary features. One fact where FUL would appear in agreement with radical underspecification is the notion that features rather than segments are the phonological primitives. Our assumption is that the child builds up its phonological system with features not with segments. The first cut for a child would be consonants vs. vowels, followed by [CORONAL] vs. non-coronal. If more vowels or consonants are discovered to be [CORONAL] and they need to be distinguished, further features will become necessary.

One major difference between radical underspecification and FUL is the treatment of features which are not specified in the underlying representation. In radical underspecification, the claim that only "unpredictable features are specified" (Archangeli, 1988) involves the addition of rules such as [+low]→[−high] if [−high] is not specified for low vowels (see 14a). Assumption of monovalent features automatically means that there is no feature [−high] and low vowels would be assigned no further features on the surface. Radical underspecification would also assume that if [−voice] is not specified, it will be subsequently added by feature filling rules. In FUL, however, voicelessness is not an attribute which will ever be available on the surface. Furthermore, unlike radical underspecification, FUL does not support multiple choices of feature inventories. To explicate this point, we use Archangeli's (1988, p. 193) example of five vowel systems [i e ɑ o u] and possible feature marking options.

(14) Options for specifying [ieɑou] in radical underspecification

	(a)	i	e	ɑ	o	u	(b)	i	e	ɑ	o	u	(c)	i	e	ɑ	o	u
high			–		–				–		–			+				+
low				+						+						+		
back					+	+		–	–								+	+

Since so many options would make a system unlearnable, the proposal in Archangeli (1988) is that universal principles guiding acquisition would take one of these options as the preferred one. In this paper she took (14a) to be the preferred option as a working hypothesis, although a footnote (p. 204, fn. 10) claims that the structure of the model does not depend on this to be the default set. It is, therefore, not obvious what the guiding principles may be. In any event, the feature specification of (14a) would never be the default option in FUL. As mentioned earlier, FUL has clear hypotheses of how features are specified during acquisition. First the ARTICULATOR node is considered, where CORONAL is assumed to always exist but remains underspecified, contrasting with another feature under the same node. Translating this into radical underspecification, [−back] would remain unspecified and presumably [+back] would be specified. This would be different from FUL's assumptions since the absence of CORONAL would not automatically mean the presence of its opposite category since there is no exact opposite. The feature that would contrast with [CORONAL] could be either [LABIAL] or [DORSAL]. Following Ghini, our preferred option would be to specify [LABIAL] next as in consonants. This step, however, needs further research (cf. also Fikkert & Levelt, 2008). A child will turn to TONGUE HEIGHT contrasts if the ARTICULATOR features are not sufficient to set apart the various words it is learning. The first TONGUE HEIGHT feature she acquires would be [LOW]. Thus, the assumptions in FUL are rather different from the ones that standard radical underspecification had made.

Now we turn to general arguments against the notion of underspecification and how one could deal with these under FUL.

1.2. *Referring to underspecified segments and differentiating coronal segments in FUL*

One of the most disturbing arguments against underspecification of coronals suggests that although harmony rules suggest coronal transparency, a number of early lexical phonotactic and morpheme-structure constraints need to refer to underspecified segments. Part of the problem has been the notion of *unmarked* vs. *marked* coronals (McCarthy & Taub, 1992, p. 365). English plain alveolars [t d n l r] are generally considered *unmarked* and

[6] Furthermore, STRIDENT is also not relevant for the velar, labialized velar or uvular series in Tahltan (Shaw, 1991). Under FUL, these would be distinguished by features dominated by the TONGUE HEIGHT and CONSTRICTION nodes and would not play a role in harmony.

[7] We use the notation [–] to indicate that a feature is not specified.

underspecified for [coronal], but interdental [θ ð] are [+anterior] and palatoalveolar fricatives and affricates [ʃ ʒ ʧ ʤ] are [−anterior, +distributed] and would be *marked* and specified for [coronal]. Thus, if one needs to make reference to features like [anterior] or [distributed], which are required to distinguish the segments [θ ð ʃ ʒ ʧ ʤ], one has to refer to the dominating feature [coronal], and hence coronal underspecification cannot work. Furthermore, it is difficult to refer to coronal consonants as a group when a language has more than one such consonant since some of the distinguishing features are dependent on [coronal]. However, in FUL, [CORONAL] does not have any dependent features, and consequently, this problem does not arise. For example, if necessary, the palatoalveolar consonants can be distinguished by [STRIDENT] and [HIGH], while the dental consonants would be [LOW]. Nevertheless, the question still remains as to how one can refer to coronal consonants as a class when there is no feature specified in the lexicon. How could one handle morpheme-structure constraints such as "the diphthong *aw* can be followed only by coronals, but *oy* is only followed by alveolars" without recourse to CORONAL and *anterior* or any other dependent feature (cf. McCarthy & Taub, 1992)?

Ghini (2001a, pp. 31–57)[8] argues extensively in favour of underspecification, pointing out flaws in the arguments presented in McCarthy and Taub (1992) and Steriade (1995). We will not repeat those in detail, but only mention a few points. He contends that during the onslaught against underspecification, there is the danger of forgetting that there are many instances where coronal underspecification is no doubt necessary. He points to Yip's (1991) observation of "cluster conditions" and the freedom of English coronals, which support that [coronal] is underspecified. Yip observes that in a stop-stop sequence C2 is represented by [t, d]; in a stop-fricative sequence, C2 is represented by [s, z]; in a fricative stop sequence, either C1 is [s] or C2 is [t, d], and so on. She comments that the English coronals enjoy the sort of freedom which is shared by glottal stops, which are usually treated as placeless.

As for referring to CORONAL consonants as a class without it being specified, one could handle these facts in a number of ways, either with an ordered set of negative constraints or with a reference to the ARTICULATOR node (see also Ghini, 2001a, chapter 1). Let us first consider the negative constraints. English *aw* and *oy* can only occur before coronal consonants and *oy* is further constrained to occur only before alveolars. The constraints limiting the distribution of English *aw* and *oy* could be described as follows.

(15) English *aw* occurs before all coronal consonants and *oy* only before alveolars.

(a)

* ROOT	+	ROOT
VOCALIC		CONSONANTAL
/aʊ, ɔɪ/		ART
		[LABIAL]

* ROOT	+	ROOT
VOCALIC		CONSONANTAL
/aʊ, ɔɪ/		ART
		[DORSAL]

Diphthongs aw and oy cannot be followed by labial and dorsal consonants

(b)

* ROOT	+	ROOT	
VOCALIC		CONSONANTAL	
/ɔɪ/		[STRIDENT]	TH
			[HIGH]

Diphthong oy cannot be followed by coronal palatoalveolar affricates or fricatives; i.e. it can only occur in front of alveolars.

[8] Ghini (2001a, b) still had the features [posterior] and [distributed] as dependents of coronal. We had been working on FUL with Mirco Ghini during this period and had been discussing the proper way of using the TONGUE HEIGHT features. Although he was completely in favour of the reanalysed FUL geometry without dependent features, the reanalysis was completed only after he passed away.

The lack of dependent features under [CORONAL] solves many of the problems raised in the literature. For instance, McCarthy and Taub citing Borowsky (1987) observe that plural, genitive and 3rd person present tense suffixes are constrained by the OCP, which requires the specification of both alveolars and palatoalveolars: *buses*, *churches*, *roses*, *judges*, etc. Here a schwa intervenes between the sibilant of the grammatical morpheme and the final sibilant of the noun or verb. However, one is not required to use coronal specification here either since the constraint refers to strident consonants. Under our feature specification the constraint can be as follows:

(16) Constraint for plural and genitive sequence of sibilants

* CONSONANTAL	+	CONSONANTAL
[STRIDENT]		[STRIDENT]

The strident consonants are the alveolar fricatives and the palatoalveolar obstruents, and hence any sequence of these will be blocked; e.g. *buses*, *catches*, *bridges*, *dodges*. The repair of the constraint would involve inserting a schwa or whatever analysis one prefers to have (recall that in English ð and θ are not strident, cf. fn. 5).

Our contention is that FUL's assumptions concerning lexical feature representation not only accounts for phonological systems, but also allows us to resolve many difficulties in language comprehension. Speech is variable and the model is built to account for asymmetries in phonological systems (e.g. markedness and transparency), but also to account for efficient language comprehension despite the variation. Underspecification, or rather sparse specification, is one way to solve the problem. The hypotheses are clear-cut and testable. Consequences of our approach for language comprehension and evidence in support of our claim are discussed in detail in Section 2, followed by a case study involving complex morphophonological alternations in Section 3, which we also assume are derivable from single stem morphemes.

2. Mapping from signal to representation—psycholinguistic and neurolinguistic evidence

FUL's assumptions about underspecification are related to contrast as well as variation. Building on earlier work, Lahiri and Reetz developed a more comprehensive model of feature representation, which plays an active role in language comprehension and production. Given that variation is so rampant in running speech, how does the system ever begin to recognize words? The idea is that the perceptual system does not hunt for detailed cues. Rather it tries to extract the most robust cues for features. The assumption is that the listener resolves the variation in speech in two steps:

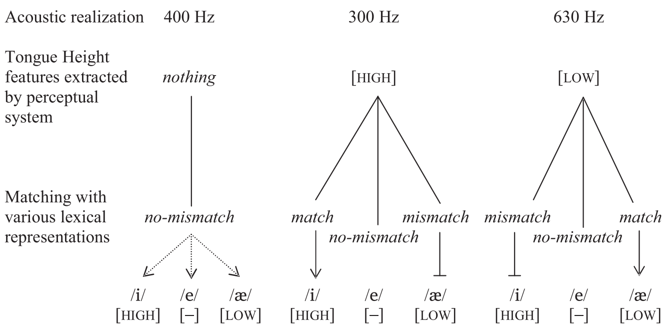
(A) The signal is parsed into features and not segments. The feature parsing is done using rough holistic acoustic parameters.
(B) A ternary mapping process matches the features as extracted from the acoustic signal to those stored in the mental lexicon.

Step-A allows multiple choices and delays decisions about individual segments. As more information comes in, the lexicon assists in making the final choice. Step-B allows a *nomismatch* possibility when the extracted feature does not find a complete match with underspecified segments. Thus, the mapping process can deliver three outcomes: (i) a *match* when features from the signal find an identical feature in the lexicon, (ii) a *mismatch* when both features exclude each other (e.g. [HIGH] mismatches with [LOW] and vice versa) and (iii) a *nomismatch* when a feature neither matches nor mismatches (e.g. [HIGH] does not mismatch with a vowel which is not specified for tongue height). Especially this third condition together with underspecification can lead to asymmetries mentioned earlier: CORONAL features extracted from the signal *mismatches* with DORSAL in the lexicon, but DORSAL extracted from the signal gives a *nomismatch* with an underspecified place (which usually receives its CORONAL place feature only in production).

A vital objective is not necessarily to find a perfect match, but not to access a conflicting form. The ternary algorithm tries to ensure that a certain amount of tolerance is possible when trying to find a match with the lexicon. For instance, in a language with a three-height system like Bengali, there are three front vowels /i/ [HIGH], /e/ [—], /æ/ [LOW]. A mid-vowel /e/ can be easily produced and perceived as high-mid or even high in fast speech. Such a mispronunciation of /e/ can turn the intended word into another word as in /beʃ/ 'nice' to /biʃ/ 'poison'. If the perceptual system extracts [HIGH] from the intended /e/ of /beʃ/, a better match will be the high vowel /i/ in /biʃ/; nevertheless /beʃ/ will still be activated since /e/ is underspecified for height. Consequently, misperception of /e/ as [i] will not throw out the intended word. Later, postperceptual or semantic contexts will lead to the correct reading but the initial access process will activate both words.

The figure below provides a graphic representation of how we envisage the perceptual system to cope with the variability in the signal. The figure explains how variable pronunciations of vowel /e/ is perceived during the matching process. On the basis of acoustic measures, let us assume that the F1 of a [i] would be less than 350 Hz and more than 600 Hz for [æ] (see Lahiri & Reetz, 2002).

(17) Variable pronunciation of [e]



No vowel is ever pronounced in the same way with identical formants. FUL's assumptions are that a mid-vowel like [e] may well be pronounced as a higher or a lower vowel depending on context. If it is pronounced as mid-high and suppose the extracted first formant triggers [HIGH], this variant of [e] would be a better match for /i/ than for /e/. Nevertheless, /e/ would still be activated. The low vowel /æ/ would not be activated since the extracted [HIGH] would conflict with the [LOW] of /æ/. Similarly, if [e] was pronounced as a lower vowel, and [LOW] was extracted, then it would be a better match for /æ/ but /e/ would still be activated. Furthermore, if the F1 value happens not to fall within the 350–600 Hz range, no height feature will be extracted and there is nomismatch with [HIGH], [LOW], or [–]. As a consequence, a [e] properly pronounced as a non-high, non-low vowel leads to a nomismatch with the feature specification of a /e/ in the lexicon; i.e., a [e] does not 'match' with a /e/ for TONGUE HEIGHT. This is one of the ways in which variation is handled in FUL.

Thus, FUL assumes that variation in speech may lead to inaccurate production and hence incorrect feature extraction. However, due to underspecification in the representation and ternary mapping logic, the inaccuracy will not exclude intended words from being recognized. Our assumption is that in normal pronunciation, the variation is always within a range. A low vowel /æ/ may become mid-low but will never be pronounced as high in any given context; that is, /gæni/ 'wise' will never become *[gini]. It may be pronounced mid-*[geni], but this is not a problem for the system (i.e. *[gini] will never occur as a variant of /gæni/ whereas *[geni] can). Similarly, a high vowel /i/ may be pronounced mid but it will never be so incorrectly pronounced in normal running speech to be perceived as [LOW], and again the system deals with it by a combination of underspecification and the ternary mapping logic. Consequently if a vowel /i/ is pronounced with a higher F1 and is closer to [e], then both /i/ and /e/ would be activated, but not /æ/. The lexicon would eventually sort out the best candidate. Crucially the acoustic signal is variable and the features that play a role in perception are extracted from this variable signal, but these alone do not constitute the decisive factor.

Our automatic speech recognition algorithm (Reetz, 2000) is also based on the feature extraction and ternary matching principles of FUL. As we reported in Lahiri and Reetz (2002, pp. 24–27), we ran the FUL recognizer and a HMM based recognition system on all 13 German vowels from the Kiel Corpus of spontaneous speech (IPDS, 1995–1997). For FUL, we used only 20 ms of the centre part of the vowels for this comparison. The vowels are classified by combinations of 7 features ([SONORANT], [LABIAL] [CORONAL], [DORSAL], [LOW], [HIGH], [RTR]) and the ternary logic described earlier. The hidden Markov model had three states for every phone and their left and right transitions, and eight mixtures to allow 8 variations of every phone to exist. The phones were modelled left-to-right without skipped states by 12 MFCCs (mel-frequency cepstral coefficients), the energy parameters, and the corresponding delta-values, giving a total of 26 parameters. The system was trained with a jack-knife procedure, where a subset of 80% of the recordings defined the pattern sequences for the phones and the remaining 20% (other speakers and other sentences) had to be 'recognized'. This procedure was repeated 5 times with different subsets from the database (i.e. each data set was used only once as test set) and the recognition results are averaged over these experiments.

The FUL system does *not* require any training and therefore there is no separation between training and test sets. For both systems, only the top-scoring vowels were counted as 'correct' recognition (lower scoring vowels are still contributing to the recognition, both in the HMM and in the FUL system, but were ignored in this comparison). Of the 4907 vowels in the database, the HMM reached 78% correct recognition and the FUL system achieved 81%. From these results it seems that the FUL system is able to hold its own in an evaluation format prescribed by stochastic models. Given that the vowels were from 36 different male and female speakers, which were equally well identified, we felt that we were on the right track.

Before we delve into further details, let us examine other models which differ in their assumptions.

2.1. Perception of assimilatory variants

When Lahiri and Marslen-Wilson (1991, 1992) introduced the notion of underspecification in language comprehension, the claim was that vowels, predictably nasalized in a following nasal context in both English and Bengali, need not be specified for the feature [nasal]. Consequently, although Bengali had underlying nasal and oral vowel phonemes, the nasal vowels were specified for nasality as in /kɑ̃dʰ/ 'shoulder' and the oral vowels were considered to be underspecified as in /kɑʤ/ 'work'. An oral vowel would become (predictably) nasalized when a nasal consonant followed, and a word like /kɑn/ 'ear' would be realized as [kɑ̃n], with a nasalized vowel. Bengali behaves here the same as English, where a word like /kæn/ *can* is usually pronounced as [kʰæ̃n]. In both Bengali and English, the nasality of the vowel in this context is due to the coarticulation of the upcoming nasal consonant and not a feature in the underlying representation. In Bengali, nasal/non-nasal vowel alternations emerge in morphophonological contexts as well. Consider the following vowel final verb roots.

(18) Nasal–oral vowel alternation in Bengali: FAM = FAMILIAR, HON = HONORIFIC, IMP = IMPERATIVE, PRES = PRESENT, FUT = FUTURE

	2P.FAM.IMP	3P.FAM.PRES	2P.HON.FUT	2P.HON.PRES
lie down	/ʃo/	/ʃo-e/>[ʃoe]	/ʃo - ben/>[ʃoben]	/ʃo-n/>[ʃõn]
touch	/ʧʰõ/	/ʧʰõ-e/>[ʧʰõʃ]	/ʧʰõ - ben/>[ʧʰõben]	/ʧʰõ-n/>[ʧʰõn]

If the underlying vowel is a nasal as in /ʧʰõ/ ‘touch’, it remains nasal throughout. If the underlying vowel is oral as in /ʃo/ ‘lie down’, then the vowel becomes nasalized in the context of the 2P.HON.PRES. morpheme /n/. Thus, the postlexical nasalization rule applies across morphemes as well as within monomorphemic words.

In the crucial gating experiment, Bengali listeners were presented with incrementing information from CVC, CṼC and CVN words. Bengali nasal vowel stimuli were cut from CṼC and CVN words (like /kɑ̃dʰ/ and /kɑn/, which are realized as [kɑ̃dʰ] and [kɑ̃n]) and oral vowels from CVC words (like /kɑdʰ/). Bengali listeners always assigned fragments like [kɑ̃], extracted from either [kɑ̃n] or [kɑ̃dʰ], to words with an underlying nasal, like /kɑ̃dʰ/, and not to words with nasalized vowels like [kɑ̃n]. Thus, the listeners perceived nasality equally from CVN and CṼC stimuli, but responded largely with CṼC stimuli. Only when they were presented with a longer fragment, which included part of the nasal consonant from the CVN stimuli, the listeners changed their answers from [kɑ̃dʰ] to words like [kɑ̃n]. Strikingly, Bengali listeners assigned fragments with oral vowels (extracted from CVC words) to words with underlying oral vowels like /kɑdʰ/ as well as CVN stimuli as in /kɑn/. The proportion of these CVC and CVN responses reflected the distribution of these words in the language. The authors argued that these results supported the claim that CVN words were underspecified for nasality and listeners did not associate nasalized vowels to these words. Furthermore, the claim was that the underlying vowels were accessed in the task and the acoustic surface match did not play a role since the CVN fragments (with surface nasalized vowels) were almost all responded to by underlying CṼC words. Crucially, the oral vowel fragments (from CVC words) were responded to by CVN words as well. Thus CVN words were considered to be a better match to oral vowels than to nasalized vowels, even though the CVN stimuli must have been nasalized or else listeners would not have responded to them with CṼC words.

Taking off from Lahiri and Marslen-Wilson's (1991, 1992) rather basic assumption that not all variants of assimilation need be represented, there has been an active research programme examining how listeners interpret phonetic variants due to assimilation. In the last decade, there have been three competing approaches to this issue. One view is that underspecification in representation is necessary to account for listeners ability to cope with assimilatory variants, which are very often asymmetric in all languages. With respect to place assimilation, Lahiri and colleagues have argued that underspecification in lexical form can give an explanation for asymmetry caused by assimilation in perception as well as production (Lahiri, 2007; Lahiri & Reetz, 2002; Scharinger, 2006; Scharinger, Reetz, & Lahiri, 2009; Zimmerer, Lahiri, & Reetz, 2008; Zimmerer, Reetz, & Lahiri, 2009). Here the asymmetry rests on the fact that CORONAL assimilates to other places of articulations but DORSAL and LABIAL do not. Asymmetries found in priming experiments, responses to deviant sounds in mismatch negativity (Eulitz & Lahiri, 2004) or lexical integration in EEG (Friedrich, Eulitz, & Lahiri, 2006; Friedrich, Lahiri, & Eulitz, 2008) can be accommodated by underspecification.

Marslen-Wilson, Nix, and Gaskell (1995) also found asymmetries when subjects were presented with minimal pairs like *late* and *lake*. Assuming underspecification and preferred direction of assimilations, one would predict that *lake* would be ambiguous since it could be an underlying dorsal or derived from assimilated [t] in phrases like *lake cruise*, which could be derived from *late cruise*. In a gating study, subjects responded to *lake* with both *late* and *lake*, but *late* (i.e. the coronal final word) was always responded to by a coronal. However, in following cross-modal repetition priming studies, both *fate* and *heap* primed their corresponding identical targets, but the opposite "word mismatches" did not prime; that is *fake* and *heat* did not prime *fate* and *heap*, respectively.

In further studies, Gaskell and Marslen-Wilson (1996) found "viability" effects where acceptability of assimilated variants only occur if the assimilation occurs in a viable context. Here the non-coronal nonword **leam* as well as the identity condition *lean* equally prime the target *lean*. However, when a viable or unviable context is presented, the acceptance rate goes down, such that **leam* activates *lean* in the context of **leam bacon*, but not in **leam gammon*. The authors claimed that in running speech listeners employed a system of inference which involved back-tracking, so that **leam* is acceptable only if the context is correct. Thus, if **leam* occurs in the context of *bacon*, it is acceptable, but not if the following word begins with a dorsal such as *gammon*. The viability hypothesis is also supported by other priming studies (Coenen, Zwitserlood, & Bölte, 2001; Gaskell & Marslen-Wilson, 2001).

Contrary to the viability hypothesis, however, Wheeldon and Waksler (2004) show that if the syntactic contexts are well chosen, then there is no evidence of inference or back-tracking. In their experiment, the assimilated nonword primes the word it is derived from, independent of the context. They claim that underspecification alone account for the data. Thus, both underspecification as well as inference seem to be necessary to understand all the results.

A third type of argument takes seriously the question of phonetic detail in assimilation and to what extent acoustic properties in the signal assist listeners to decode any coarticulation. Place assimilation like many other neutralization rules can be graded in that the change of one segment to another may not be complete (Nolan, 1992). An assimilated segment may share acoustic cues with both the altered articulation as well as the original segment in the assimilatory context. That is, [m] of the altered **leam* may share properties of labiality of [b] of *bacon*, but also keep some of the original properties of /n/ of *lean*. Gow (2002, 2003) argued that compensation for place assimilation involves a process of parsing acoustic properties, in which the two sets of place cues extracted from an assimilated segment must be associated with the appropriate underlying segments. He examined the acoustic information extracted by listeners from graded assimilated speech and how they interpreted it, taking seriously the notion of variance in assimilation (Gow, 2001, 2002, 2003). Based on natural assimilated tokens (rated by independent listeners), Gow argues that the coronal cues of an assimilated /n/ (as for example *lea[m/n]*) will be associated with the /n/ itself, whereas the bilabial cues will be used as evidence for an upcoming bilabial consonant.[9] In the absence of a following bilabial consonant (as in **leam gammon*), the bilabial cues cannot be accommodated, thus accounting for the viability effects of Gaskell and Marslen-Wilson (1998). This view of compensation for assimilation is attractive in that it makes use of the same process that is

[9] This is similar to the account of vowel nasalization as presented in Lahiri and Marslen-Wilson (1991, 1992) for allophonic assimilations in English. If the nasal information from a vowel could not find a match in the vowel itself, it anticipated an upcoming nasal consonant.

assumed to operate throughout the perception of connected speech. However, it functions only if there is partial assimilation and the listener is able to extract both coronal and labial information, for example, simultaneously from the same sound. If there is complete assimilation, and this does occur as we know from instances of language change, as well as on the basis of experimental evidence (Gaskell & Snoeren, 2008), the listener is unable to access the original segment. That is if **leam* has no hint of coronality left, the parser cannot access the original *lean*. Further, asymmetry remains unexplained in this model.

We turn now to the issue of the likelihood of complete assimilation. Three studies are relevant here. Dilley and Pitt (2007) observed that 9% of word final coronal stops and nasals assimilated to the following consonant. Acoustic measurements showed no difference between assimilated coronal consonants and unassimilated labials and dorsals in the preceding vocalic context. They concluded that assimilation was often (near) complete in running speech. A second study, which examined the variability in natural speech, showed that trained phoneticians labelled 6% of word final coronals as being assimilated in place when words with non-coronal consonant initials followed in the sentences; word final nasals in function words assimilated even in 18% in these cases (Zimmerer et al., 2009). These observations were corroborated by forced- and free-choice perception experiments as well as acoustic measures, showing consistently (a) that assimilations were often complete and (b) that assimilations were asymmetric, that is, /n/, /d/ and /t/ assimilated but /m/, /b/, /p/, /ŋ/, /g/ and /k/ did not or only rarely.

In another recent study, Gaskell and Snoeren (2008) also examined the possibility of strong versus weak assimilations, not in natural dialogues, but in controlled experimental situations where speakers were asked to produce sentences like *I think a quick run/rum picks you up* and *I think a quick run/rum does you good*, which may or may not lead to assimilations. The contexts were dubbed "viable" (i.e. assimilatory non-coronal) or "unviable" (i.e. coronal context). These sentences were used in a perception test where listeners had to choose between pairs of words as *run–rum, bride–bribe, right–ripe*, etc. that were presented on a screen at the end of the whole sentence. Listeners perceived significantly more labials and velars when an underlying coronal was followed by a labial (or velar) and not the other way around. Thus, intended *right* was perceived significantly more as *ripe* when a labial viable context followed than in a coronal context (13.7% shift). But surprisingly, the perception of underlying *ripe* shifted more to *right* in the viable (non-coronal) context (4.5% shift). Second, there was a gradient effect in perception depending on the manner of articulation where nasals and voiceless stops were perceived more often as assimilated compared to voiced stops. Third, the reaction times for labelling the assimilated nasals in a viable context were slower indicating that listeners were faced with two choices (e.g. *right* and *ripe*) since the assimilation caused lexical ambiguity.

Gaskell and Snoeren (2008) point out that the results of their first perception experiment show that "some assimilated coronal segments are effectively indistinguishable from their non-coronal counterparts" (p. 1643). This means that assimilation can be complete. From our perspective, if assimilation is complete and non-coronal is extracted from the signal, listeners would be able to access both—coronal and non-coronal segments. However, the other crucial point of this paper is the effect of "viability" on perception.

In a further perception test, Gaskell and Snoeren used a neutral as well as a biasing preceding context, and additionally cross-spliced the following context (i.e., there were four sets of stimuli). They used only a subset of the non-coronal stimuli (e.g. *bribe*) from the first experiment "avoiding cases where there was good agreement between participants on the underlying consonant of the target word in the viable condition" (p. 1641). This meant that the selected items were ambiguously perceived as sometimes coronal and sometimes non-coronal (32% coronal responses in a "viable," non-coronal context and 18% in "unviable," coronal context). In this second experiment, they presented these stimuli with a following coronal ("unviable") and non-coronal ("viable") context with a semantically neutral preceding phrase where both *bribe* and *bride* were possible: *A report about the bribe made the local paper/turned up in the local paper*. Although listeners always heard the same words with non-coronal (but ambiguous) consonants, viz. ambiguous *bribe*, responses with *bride* constituted 42% when the context was labial, but 28% when the context was coronal. When a further preceding semantic context was added biasing the percept *bride* (*The ceremony was held in June and the sunny weather added to the air of celebration*), these responses increased to 62% when a labial segment followed and 48% when a coronal followed. When the following context was cross-spliced and thus reversed (e.g. *turned up in the local newspaper* was swapped with *made the local newspaper*), the results changed. The cross-splicing led to the *bride* responses being reversed—55% for the new coronal context (*bribe*$_{made}$ *turned up*) and 62% for the new labial context (*bribe*$_{turned}$ *made*). We list in (19) the average of the semantically neutral and biased contexts for the four conditions on which Gaskell and Snoeren base their line of reasoning.

(19) Percentages of responses to non-coronal target words (which were ambiguously perceived) averaged over a preceding biasing and neutral context, and four different following contexts in Gaskell and Snoeren (2008) (percentages for the neutral and semantically biased conditions are given in brackets in column 5).[10]

Context	Targets (Examples)	CORONAL responses for these targets in Experiment 1	Viability of original context viable=non-coronal unviable=coronal	CORONAL responses	non-CORONAL responses
Original	*bribe*$_{made}$ *made*	32%	Viable	52% (42% 62%)	48%
	bribe$_{turned}$ *turned*	18%	Unviable	38% (28% 48%)	62%
Cross-spliced	*bribe*$_{made}$ *turned*	32%	Viable	45% (55% 35%)	55%
	bribe$_{turned}$ *made*	18%	Unviable	50% (62% 38%)	50%

These are intriguing results. Gaskell and Snoeren (2008) claim, quite rightly, that the following context is crucially responsible for the results. The cross-splicing indicates that the very same acoustic stimulus triggers different responses depending on the context. The results show that the surface form *bribe* is actually

[10] The third column indicates the way in which the chosen target items were perceived in Experiment 1 (Gaskell & Snoeren, 2008, p. 1641). The fifth and sixth are averaged over the preceding neutral and semantic biased context and are given in their text on p. 1642.

perceived as *bride* if the following context is labial ("viable" context). Their model, *probabilistic learned compensation model*, accommodates these results by invoking the lexicon and not just the coarticulated speech.

Gaskell and Snoeren's results indicate that (a) total assimilation does occur in spontaneous speech and (b) a following phonological context biases the listener's perception of assimilated stimuli. The first conclusion is important since some models as in Gow (2001, 2002, 2003) argue that assimilation is invariably partial and that underlying acoustic cues are always available to access the correct lexical item. The second conclusion argues against an analysis based entirely on underspecification. The results can be best understood in their view within the framework of learned compensation. Under such an approach, assimilations are learned by listeners, who also learn to compensate. Thus, in experiment 2, cross-splicing the context from, for example, labial /m/ to coronal /t/, the listeners' percept of a coronal [d] *decreased*, because listeners assume that a following coronal means a non-assimilated /b/.

These results are very interesting and one can also interpret them from the feature extraction perspective of FUL. For instance, if we really consider the context (and not just the cross-splicing alone) and compare the surface forms, e.g. [b][m] as in *bribe$_{made}$ made* and *bribe$_{turned}$ made*, the percentage [d] responses are 52% and 50%, respectively, then the difference is very small. Similarly, comparing [b][t] in *bribe$_{turned}$ turned* and *bribe$_{made}$ turned*, the [d] responses are overall less than above: 38% and 45%. That is, when the context is a labial then there are proportionately more coronal responses compared to a coronal context. Recall that the words with final labials like [b] were deliberately chosen such that they were ambiguous. Within FUL, in [b][m] context, an extracted LABIAL feature would not mismatch with an underspecified PLACE (although a *match* increases the score of a word, see Lahiri & Reetz, 2002) and an ambiguous target might not even lead to the extraction of any place information from the signal and therefore not mismatch with any place features in the lexicon ([LABIAL], [DORSAL] and [–]). Hence, when LABIAL is extracted from [m] and some of the preceding [b] it finds a match with the following [m], giving overall a 50% labial/coronal divide for the final segment in *bribe/bride*. For [b][t], if labiality is extracted from some of the [b] and one hears the following coronal, the LABIAL feature would be a nomismatch for the following coronal but the LABIAL would increase the score for the LABIAL responses for the last consonant of *bribe*. It is actually hard to see how the activation of both words can be explained without reference to underspecification.

The compensation model of Gaskell and colleagues, and all models where context is necessary to perceive a preceding sound correctly, are based on phenomena like assimilation, where one sound affects a contiguous sound. What is crucial for underspecification, however, is not assimilation alone but representation of contrasts, which is what we discussed in the first half of the paper. We will, therefore, describe an experiment where pure phonological context plays no role but underspecification does.

2.2. Variation without assimilation

Note, that beginning from Lahiri and Marslen-Wilson, the focus has been variation as a consequence of assimilation. Our hypothesis is stronger, namely that coronal underspecification is pervasive, with or without assimilation. That is, medial as well as initial coronal consonants remain underspecified unless the grammar resorts to special mechanisms. We summarize briefly some of the existing experimental evidence in support of our claim that FUL does not assume contextual underspecification and underspecification does not depend on syllable structure nor on possible allophonic processes. That is, if CORONAL is underspecified it would be so irrespective of any possible assimilation context, and consequently word initial or medial CORONAL consonants would be underspecified for place. Lahiri and Reetz (2002) report a semantic priming task testing medial as well as final place asymmetry. The task was lexical decision in a crossmodal design, which meant that subjects heard prime words and immediately saw a target word written on the screen to which they had to say whether it was a word or not. For the medial condition, where no assimilation is ever possible, words like *Ho[n]ig* 'honey' predictably primed *Biene* 'bee' and *Ha[m]er* 'hammer' primed *Nagel* 'nail'. Pseudoword variants of these primes, however, gave asymmetric results; **Ho[m]ig* successfully primed *Biene*, but **Ha[n]er* did not prime *Nagel*. That is, the LABIAL [m] of the pseudoword **Ho[m]ig* was tolerated as a variant of the underspecified /n/ in *Ho/n/ig*. But the CORONAL [n] of the pseudoword **Ha[n]er* was rejected by the underlying specified /m/ in *Ha/m/er*.

More recently, Friedrich et al. (2006) ran an electroencephalographic (EEG) study using words varying in medial CORONAL vs. non-CORONAL consonants to examine whether similar asymmetries would be found with a more direct technique measuring brain activity. As mentioned before, word medial coronals as in *Hor/d/e* 'horde' have no PLACE represented and the claim is that its corresponding non-coronal variant like **Hor[b]e* cannot mismatch this empty PLACE slot and therefore would activate *Hor/d/e*. A similar mapping would not occur with coronal pseudoword variants like **Pro[d]e* of the real word *Pro/b/e* 'test'. Although CORONAL is not represented in the lexicon, this feature can be perceived in the signal and can be used for lexical mapping and consequently, CORONAL extracted from **Pro[d]e* mismatches the specified LABIAL of *Pro/b/e* and therefore cannot activate this word. If the claim is correct, the prediction is that lexico-semantic memory search processes would be successful when a non-coronal variant is presented and activates the corresponding coronal target word, but not when a coronal variant is presented, which would lead to an immediate correct rejection as a non-existing lexical item. Thus, an asymmetry was predicted, at least for the initial N400 pseudoword effect, which most likely is related to lexico-semantic processing.

The task was speeded lexical decision to auditory stimuli. There were no significant differences in reaction time between coronal and non-coronal words nor for the different pseudowords. However, the error rates revealed significant differences. Non-coronal pseudowords like **Hor[b]e* (< *Hor[d]e*) had significantly more errors than coronal pseudowords like **Pro[d]e* (< *Pro[b]e*). This means that it was easier for subjects to recognize **Pro[d]e* as a nonword. In contrast, **Hor[b]e* was more difficult to reject as a nonword since it did activate the real word *Hor[d]e* producing significantly more errors. In the ERP-data, the early N400 results also showed a clear asymmetry in the earlier activation period of 100–250 ms. Mean amplitudes of the coronal pseudoword variants were significantly more negative than mean amplitudes for their non-coronal base words. By contrast, ERPs for non-coronal variants did not differ from their base words in this initial part of the N400 pseudoword effect. Furthermore, a significant difference between both types of pseudoword variants, but not between both types of words relates this early ERP deflection to mismatch detection in the case of coronal pseudowords. The asymmetric results do not support full specification or specification of phonetic detail in the lexicon.

Thus, medial coronal consonants, which are contrastive with dorsal and labial consonants, also show an asymmetric pattern. Non-coronal pseudowords with labial or dorsal consonants are accepted as a variant of the corresponding coronal word, but not vice versa as shown in the error data as well as in the early N400 effect. Medial consonants do not undergo any assimilation or any other alternation, which could have made them familiar to the listeners.

An even stronger hypothesis for a non-contextual, non-assimilation based abstract representation is placelessness of initial consonants. Friedrich et al. (2008) investigated this hypothesis using a lexical decision task with and without fragment priming. They compared the processing of German words and pseudowords that differed only in the place of articulation of the initial nasal or stop consonant. In the fragment priming experiment, two sets of words were used, one beginning with labials and dorsals (e.g. /grɛntsə/ *Grenze* 'border') and the other with coronals (e.g. /drɑxən/ *Drachen* 'dragon'). The words were primed with their first part such as [gɣɛn] and [dɣɑx] as well as with fragments where the place of the initial consonant was reversed *[dɣɛn] and *[bɣɑx] compatible with the nonwords **Drenze* and **Brachen*. Event-related brain potentials indicated that pseudoword fragments with initial non-coronal PLACE (e.g. *[bɣɑx]) activate words with initial coronal PLACE (e.g. *Drachen*) while coronal pseudowords (e.g. *[dɣɛn]) do not as effectively activate non-coronal words (e.g. *Grenze*). Thus, certain word onset variations do not hamper the speech recognition system, especially coronal and non-coronal onsets show an asymmetry. Friedrich et al. (2008) interpret this asymmetry as a consequence of underspecified coronal place of articulation in the mental lexicon. Neither exemplar models storing all information, nor contextual models could account for this asymmetry since both would assume that variants are accessed via experience. There could be no reason why German word initial stops and nasals would have been heard mispronounced in this particular way.

FUL makes the same predictions for vowels, viz. that given the underlying contrast between /o e ø/ in German, /o/ is specified for DORSAL and LABIAL, while the others are unspecified for CORONAL, and /ø/ is specified only for LABIAL. In terms of perception and mapping from the signal to the representation, FUL's predictions for German vowels are presented in (20).

(20) Mapping of features extracted from the signal to features in the lexicon of German vowels and their activation (extracted features or in normal print and features stored in the lexicon are in italics)

Sound	Extracted features *Stored features*		Matching	(Not) Activated vowels
[o]	[LABIAL]	[DORSAL]		
	[LABIAL]	*[DORSAL]*	MATCH	/o/
	[LABIAL]	*[–]*	NOMISMATCH	/ø/
	[–]	*[–]*	NOMISMATCH	/e/
[ø]	[LABIAL]	[CORONAL]		
	[LABIAL]	*[DORSAL]*	MISMATCH	o
	[LABIAL]	*[–]*	NOMISMATCH	/ø/
	[–]	*[–]*	NOMISMATCH	/e/
[e]		[CORONAL]		
	[LABIAL]	*[DORSAL]*	MISMATCH	o
	[LABIAL]	*[–]*	NOMISMATCH	/ø/
	[–]	*[–]*	NOMISMATCH	/e/

For example, in the perception of an [o] the features [LABIAL] and [DORSAL] are extracted from the signal, which are a nomismatch with the feature set [LABIAL] [–] of an /ø/ in the lexical representation, and /ø/ will be activated. The other way round, the features [LABIAL] and [CORONAL] will be extracted on hearing an [ø] and since [CORONAL] mismatches with [DORSAL], the /o/ will not be activated.

A magneto-encephalographic (MEG) study reported these topographic differences in the processing of mutually exclusive isolated CORONAL and DORSAL vowels in German (Obleser, Lahiri, & Eulitz, 2004). Eulitz and Lahiri (2004) used a component of the event-related brain activity, the *Mismatch Negativity* (MMN), to investigate the issue of asymmetry in mapping. MMN is assumed to be an automatic detection measure of the brain's ability to detect change in sounds, particularly to phonemes (Näätänen & Alho, 1997). If a sound is presented many times in a sequence (known as the *standard*), it is considered to tap the long-term sound representation, or, in other words, the underlying representation. If another sound is presented right after the sequence (i.e. a *deviant*), it would cause something of a jolt, and the brain would detect a change and respond accordingly. The classical MMN is high amplitude difference around 180 ms from the onset. Eulitz and Lahiri (2004) noted both an amplitude and a latency difference. As predicted by the matching algorithm, for the pair [o]–[ø], when [o] was the *standard* (i.e. underlyingly specified for DORSAL) and [ø] the *deviant* such that [CORONAL] is extracted, there was a higher and earlier MMN peak than the other way around. Similar predictable asymmetric pattern of results were obtained for the other pairs. Thus, just as for the consonants, the vowels showed asymmetric perceptual responses as predicted by FUL.

Underspecified representation, therefore, seems to be a reasonable candidate to explain the experimental data. We now turn to a detailed behavioural study to provide more experimental support for an underspecified representation for vowels.

3. A case study: German umlauted vowels

3.1. Background

German has underlying lexical rounded front vowels /y ʏ ø œ/ (often referred to as umlauted vowels), which contrast with rounded back vowels /u ʊ o ɔ/, and unrounded front vowels /i ɪ e ɛ/, e.g. *Tür*, *Tour*, *Tier* /y, u, i/ ('door', 'tour', 'animal'). These also play a role in morphophonemic alternations where a noun in the singular (SG) has rounded back vowels but the plural (PL) or the diminutive (DIM) has an umlauted vowel (e.g. $Sohn_{SG}$~$Söhne_{PL}$~$Söhnchen_{DIM}$ [zon~zønə~zønçən] 'son~sons~little son'). With few exceptions the umlauting occurs for all stressed back vowels in the diminutive, but not for all stems in the plural (e.g. $Boot_{SG}$~$Boote_{PL}$~$Bötchen_{DIM}$ [bot~botə~bøtçən] 'boat~boats ~little boat'). In phonological analyses, there are two ways of accounting for the umlauted vowels in such alternations in the surface output. Either particular suffixes carry the umlaut trigger, which is then attached to the stem vowel (cf. Kloeke, 1982; Wurzel, 1970), or specific alternating stems carry an umlaut floating feature, which is associated in appropriate contexts (Wiese, 2000). Our focus is on the perception and ultimately recognition of such words. The question we ask is how does an umlauted form like *Söhne* with an [ø] on the surface activate the base form *Sohn* which is pronounced with an [o]? Likewise, how do forms like *Söhnchen* and *Bötchen* activate their respective base forms? To answer these questions we need to make certain specific assumptions about the representations.

FUL makes two crucial assumptions regarding the morpho-phonology of the diminutive and the plural. Here we focus on the [LABIAL] front vowels, but the same holds for the unrounded front vowels. First, all diminutives have independent stem morphemes, separate from the noun (i.e., there are separate lexical entries for *Sohn* and *Söhnchen*). Since diminutives in German are not very productive, unlike in Swiss German or Dutch, one cannot easily form diminutives of all words. For instance, diminutives like *Balönchen* 'little balloon' or *Anoräkchen* 'little anorak' are highly suspect. Furthermore, the semantic relationship between the stem and the diminutive can easily be non-transparent: *Brot* 'loaf of bread', *Brötchen* 'roll'; *Hut* 'hat', *Hütchen* 'cone shaped devices'. The stressed vowel in the diminutive is not specified as DORSAL or CORONAL and surfaces as CORONAL. Second, the nouns, which do have umlauted plurals with rounded back vowels in the singular, also have unspecified stem vowels. The stem is lexically marked in the singular to surface as DORSAL (cf. Scharinger, 2006) while the plural takes the default specification of CORONAL. The nouns with back rounded vowels which do not change in the plural are specified for DORSAL. As a consequence, different representations can have the same surface forms. In particular, nouns with stem vowels specified as [LABIAL] or [LABIAL, DORSAL] in the underlying representation will both surface as [o] in the singular since the feature [LABIAL] without a specification for PLACE will generate the place feature [DORSAL] during the production due to the lexical marking of these words. This rule overrides the general default rule of filling in [CORONAL] for unspecified vowels. The detailed production rules are given below.[11]

(21) Production of words with specified and unspecified stems ($_x$ marks a morpheme-specific rule; plural morphemes of forms without vowel alternations are written in round brackets).

	Noun-stems			Diminutive-stems[a]	
	Boot(e)	*Sohn, Söhne*	*Löwe(n)*	*Bötchen*	*Söhnchen*
Underlying	*B/o/t* [LAB] [DOR]	*S/o/n$_x$* [LAB] [–]	*L/O/we* [LAB] [–]	*B/O/t* [LAB] [–]	*S/O/n* [LAB] [–]
Surface forms					
Singular-rule$_x$: [DOR] in SG if unspecified		*S*[o]*n* [LAB] [DOR]			
Unspecified default: [COR] if unspecified		*S*[ø]*ne* [LAB] [COR]	*L*[ø]*we(n)* [LAB] [COR]	*B*[ø]*TCHEN* [LAB] [COR]	S[ø]*NCHEN* [LAB] [COR]
No change:	*B*[o]*t(e)* [LAB] [DOR]				

[a] Words with a front rounded vowel in the base form have identical vowels in the diminutives and are not listed here.

The representational assumptions concerning umlauted stems are very similar to that of Wiese (2000). The difference lies in the features themselves and in the markedness rules. For instance, Wiese does not use the feature CORONAL. Rather his binary features [±front] ([+front] corresponding to CORONAL) is dominated by DORSAL, an assumption which is similar to the Halle–Sagey model where [−back] is dominated by DORSAL. Nevertheless, the basic principle that the stems (rather than the affixes) bear the property of being "umlautable" is identical.

As mentioned earlier, our focus here is on the extraction of the features from the signal and the perceptual consequences of the underspecified representations. Here our ternary matching algorithm plays a crucial role. The feature input of [LABIAL, CORONAL] from the signal will be a *nomismatch* with a [LABIAL][–] representation, but will *mismatch* with the [LABIAL, DORSAL] representation (cf. 20). Although the diminutive stems are morphologically separate, the stem of *Söhnchen* has inadvertently the same underlying features as the stem of *Sohn*, but *Bötchen* and *Boot* have different stems. Consequently, we predict a difference in activation depending on the underlying form although the surface forms of the stem vowel may be identical.

[11] We use the symbol /o/ for the underlying form that is specified as [LABIAL][DORSAL] and /O/ for the underlying form that is specified as [LABIAL][–] in the lexicon.

(22) Activation of noun-stems with /o/ and /O/ on perceiving [o] and [ø]

	Noun-stems	Diminu-tive-stem	Noun-stems		Diminu-tive-stem	Noun-stems
	Boot(e)	*Bötchen*	*Sohn*	*Söhne*	*Söhnchen*	*Löwe(n)*
(Prime) Input	*B*[o]*t(e)* [LAB] [DOR]	*B*[ø]*tchen* [LAB] [COR]	*S*[o]*hn* [LAB] [DOR]	*S*[ø]*ne* [LAB] [COR]	*S*[ø]*nchen* [LAB] [COR]	*L*[ø]we*(n)* [LAB] [COR]
Matching	*match*	*mismatch*		*nomismatch*		*nomismatch*
Underlying (Target)	[LAB] [DOR] *B*/o/*t*		[LAB] [–] *S*/O/*n$_x$*			[LAB] [–] *L*/O/*we*
Activation	YES	NO		YES		YES

Alternative hypotheses assuming full listing of all surface forms would not predict any such asymmetry. To test this hypothesis we used a lexical decision task with delayed priming. This experimental paradigm has shown priming effects for morphological relations, but not for semantic relations (Henderson, Wallis, & Knight, 1984; Marslen-Wilson, Frost, Deutsch, Gilboa, & Tannenbaum, 2000; Napps & Fowler, 1987; Scharinger, 2006). Within a nominal inflectional paradigm, the singular and plural are always semantically related. If transparent diminutives are chosen, then they too will be semantically related to the nouns. Since we wish to see if the similar and conflicting stem representations hinder or assist activation and recognition, a delayed priming task seemed more appropriate than an immediate repetition priming task. Unlike immediate repetition priming, where the unavoidable semantic relation of a morphological related form would always lead to priming, the delayed priming paradigm allows the separation of these two effects.

3.2. Material

In the auditory delayed priming task, several words and nonwords intervened between the prime and the target. Since diminutive formation is not completely regular, to ensure that only acceptable diminutives were presented, we first ran a separate off-line judgement task. A list of 150 diminutives were given in writing to 10 subjects who were asked to evaluate them on a scale from 1 (very acceptable) to 5 (not acceptable). From the set of words which received the highest scores, 26 words that form a plural with an umlaut (e.g. *Sohn*~*Söhne*) and 26 without an umlauted plural (e.g. *Boot*~*Boote*) were selected. Although the diminutives were in general lower in frequency than the non-diminutive nouns all words were controlled for frequency within each class.

For these 52 words, the nominative singulars formed the targets in the auditory lexical decision task. For each target there was an equal number of plural, diminutive, semantically related and unrelated control prime words (e.g., for the base form *Boot* 'boat', *Bötchen$_{DIM}$*, *Boote$_{PL}$*, *Schiff$_{SEM}$* 'ship' and *Wiese$_{CTRL}$* 'meadow' occurred as primes). The table in (23) provides the experimental design with the priming predictions. In the second column, we provide the critical features we assume are extracted from the acoustic signal when the prime is perceived. In the third column, the features of the stem vowel of the target are given. Since this is a delayed priming task, there should be no semantic priming. Thus, any priming we obtain must be due to the morphophonological activation.

(23) Experimental design with priming predictions

Conditions	PRIMES with extracted features of the stem vowel		TARGET with representation of the stem vowel			Priming effect
PLURAL	Boote	[LAB] [DOR]				yes
DIMINUTIVE	Bötchen	[LAB] [COR]	Boot	/o/	[LAB] [DOR]	**no**
SEMANTIC	Schiff 'ship'					no
CONTROL	Wiese 'meadow'					reference
PLURAL	Söhne	[LAB] [COR]				yes
DIMINUTIVE	Söhnchen	[LAB] [COR]	Sohn	/o/	[LAB] [–]	**yes**
SEMANTIC	Tochter 'daughter'					no
CONTROL	Tisch 'table'					reference

Targets and primes were split into four sets, so that each set contained 13 plurals, 13 diminutives, 13 semantically related words and 13 controls for the same 52 targets. Between each prime and target there were 6–8 filler items, and between the target and the next prime there were 3 filler items. 50% of all items were nonwords. 14 of the fillers ended in *-chen*, which, in word fillers, looked like a diminutive suffix, but was really part of a noun (e.g. *Kirche-n* 'church-PL'). Additionally, there were 20 fillers (half words and half nonwords) at the beginning of each set. The item lists of the four sets were identical with the exception of the (rotated) prime words. All words and nonwords were read by a male speaker in a sound treated room, recorded on digital tape (DAT), and transferred on a computer at 44.1 kHz. Each word was cut out with a signal editor program and care was taken to avoid clicks at the cuts.

In total, each subject heard 644 items, of which there were 322 nonwords, 218 filler words, 52 target words (singular form of nouns), and 52 primes (13 diminutive, plural, control and semantically related). The diminutives made up only 2.2% of all items that the subjects heard in the experiment. The occurrences of diminutive, plural, control and semantically related prime–target pairs were randomly distributed over the whole list and sequences of more than 4 word/nonwords in a row were removed by hand by swapping appropriate items.

Items were presented in the following sequence: after a warning tone of 300 ms and a pause of 200 ms, the items were presented followed by a period of 1500 ms silence during which the subjects had to make a word/nonword decision. Reaction times were measured from the offset of each item. The subjects were instructed orally and in writing to decide as fast as possible whether the items they heard were words of German or not. A push-button box was placed in front of them with buttons marked "yes" and "no", and they were required to press the relevant button with the index fingers of their preferred hand.

3.3. Subjects

56 subjects (students of the University of Konstanz and native speakers of German, with no hearing deficiencies) were paid for their participation in the experiment. They were divided into 4 groups of 14 subjects (equal number of male and female students). None of the subjects had taken part in the pre-evaluation of the diminutive words and each subject heard only one of the four experimental tapes and, hence, heard every word and nonword only once.

3.4. Results

Three subjects were excluded since they produced more than 15% wrong or missing responses to the word and nonword stimuli. Responses to the targets were excluded when their primes were not recognized correctly (i.e., when there was a 'nonword' or no response to the plural, diminutive, semantically related or control prime then the response to the respective target was disregarded). Incorrect responses to the targets and responses with reaction times more than two standard deviations above the mean were also removed from further evaluation. Eventually, 2031 responses went into the final analyses (921 of the *Boot*∼*Boote*∼*Bötchen* class and 1110 of the *Sohn*∼*Söhne*∼*Söhnchen* class). Of these targets, 503 had plural primes, 466 diminutive primes, 529 semantically related primes and 533 unrelated control primes.

An analysis of variance (ANOVA) was performed with the statistical suite JMP, Vers. 7.0.2 on a Macintosh computer. Reaction time was the dependent variable, UMLAUT (SPECIFIED *Boot*∼*Boote* vs. UNSPECIFIED *Sohn*∼*Söhnchen*), PRIME-TYPE (DIM, PL, SEM, CTRL), UMLAUT × PRIME-TYPE, and SUBJECT and TARGET (nested under UMLAUT), both as random factors, were the independent factors in this mixed model with a REML (restricted maximum likelihood) analysis.[12]

(24) Mean reaction times and amount of priming (Δ Mean) of the experiment

Stem vowel	Example	Condition	N	Mean (ms)	Δ Mean (ms)
/o/	Boote	Plural	230	327.68	15.45
	Bötchen	Diminutive	192	335.21	7.92
	Schiff	Semantic	248	344.40	−1.27
	Wiese	Control	251	343.13	reference
/o/	Söhne	Plural	273	343.89	18.08
	Söhnchen	Diminutive	274	349.31	12.66
	Tochter	Semantic	281	356.00	5.97
	Tisch	Control	282	361.97	reference

The analysis yielded significant effects for PRIME-TYPE ($F(3,3)=4.2$; $p=.006$) but not for UMLAUT ($F(1,48.98)=2.4$; $p=.124$). The post hoc analysis of the planned comparisons of the two UMLAUT groups for the PRIME-TYPES against the control (CTRL) showed significant effects for the plural (PL) primes (*Boote*∼*Boot*: $\Delta_{\text{CTRL-PL}}=18.43$ ms, $t=2.19$; $p=.028$; *Söhne*∼*Sohn*: $\Delta_{\text{CTRL-PL}}=18.42$ ms, $t=2.38$; $p=.017$) but not for the semantic (SEM) conditions (*Schiff*∼*Boot*: $\Delta_{\text{CTRL-SEM}}=-.26$ ms, $t<1$; ns; *Tochter*∼*Sohn*: $\Delta_{\text{CTRL-SEM}}=7.92$ ms, $t=1.03$; ns). Post hoc comparisons indicated that the difference between the priming in the SPECIFIED (*Boote*∼*Boot*) group was not significantly different from the priming in the UNSPECIFIED (*Söhne*∼*Sohn*) group (SPECIFIED-UNSPECIFIED: $\Delta_{\text{Boot:(CTRL-PL)-Sohn:(CTRL-PL)}}=.01$ ms, $t<1$; ns).

The crucial condition is the priming effect in the diminutive (DIM) condition. Here, the UNSPECIFIED (*Söhnchen*∼*Sohn*) umlaut group showed significant priming, but not the SPECIFIED (*Boote*∼*Boot*) group (*Bötchen*∼*Boot*: $\Delta_{\text{CTRL-DIM}}=9.27$ ms, $t=1.05$; ns; *Söhnchen*∼*Sohn*: $\Delta_{\text{CTRL-DIM}}=15.37$ ms, $t=1.99$; $p=.047$). Fig. 1 summarizes these results.

3.5. Discussion

The delayed priming experiment showed that semantically related words do not prime the target words; that is, the reaction times to the nominative singular target words after an unrelated or a semantically related prime word were essentially the same and are, hence, not significantly different. This is true for the SPECIFIED *Boote*∼*Boot* as well as the UNSPECIFIED *Söhne*∼*Sohn* groups.

[12] The REML estimation tests against the whole population and is always more conservative than the traditional EMS (Expected Mean Square) method.

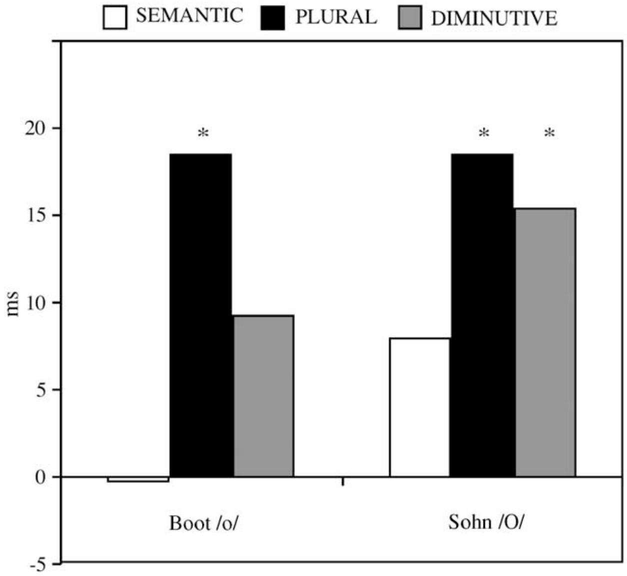


Fig. 1. Differences in reaction times (in ms) for the targets in relation to the control prime for the semantic, plural and diminutive conditions for the *Boot* and *Sohn* targets. Significant faster reactions of the targets than to control conditions are marked by an asterisk (*).

In contrast, the reaction times to the targets were significantly shorter than in the control condition when the plural form of the targets served as primes. Thus, there is no priming effect for semantically related words, but there is a priming effect when a morphologically related word is presented prior to the target although there are several filler words and nonwords interspaced. More important is the observation that this priming occurs although the surface form of the stem vowel has changed: the umlauted plural form (*Söhne*) does not hinder the priming of the non-umlauted singular nominative (*Sohn*). This suggests that the presentation of the plural form has activated the base (nominative) form and hence the reaction to this base form becomes faster. The reason for this priming could not be due to the semantic relationship alone since the presentation of a semantically related form like *Tochter* did not activate the base form *Sohn*. Another explanation could be that the umlauted vowel of the plural activates the non-umlauted form directly without any semantic contamination. This is in accordance with an underspecified representation of the rounded back vowel of the base form without specification for DORSAL (i.e., the vowel /o/ is represented as being [LABIAL]; perceiving [o] with the acoustic features [DORSAL, LABIAL] or [ø] with the features [CORONAL, LABIAL] would both activate the vowel /o/ with the lexical feature [LABIAL]).

The crucial case is the presentation of the diminutive forms which are always umlauted independent of whether the plural forms are umlauted or not. Here, the diminutives of the UNSPECIFIED *Söhnchen*∼*Sohn* group primed the nominative singular targets, but not the SPECIFIED *Bötchen*∼*Boot* group. Neither a semantic relation could explain this result (since as we have seen semantically related words did not prime) nor a purely morphological relation (since, in this case, both groups ought to behave in the same way). In light of the underspecified representation this result is consistent with the assumption that the /O/ in *Sohn* is specified only for [LABIAL], but that the /o/ in *Boot* is specified for [LABIAL] and [DORSAL]. The feature [CORONAL] from the signal mismatches with the features [DORSAL] in the lexicon for *Boot* (cf. 20 and 22) when the features [CORONAL, LABIAL] are perceived from [ø]. On the other hand, perceiving [CORONAL, LABIAL] from [ø] does not lead to a mismatch with the feature [CORONAL] of the lexical representation of /o/ in *Sohn*, which is only specified for [LABIAL] (22). In other words, there are two phonemes /o/ [DORSAL, LABIAL] and /O/ [LABIAL] in the lexicon which both surface as [o] [DORSAL, LABIAL] in the singular. On perceiving [CORONAL, LABIAL] from a signal, an /o/ ([DORSAL, LABIAL]) leads to a mismatch (and a base form like /bot/ *Boot* is not activated) whereas an /O/ ([LABIAL]) does not mismatch (and a base form like *Sohn* is activated).

4. Conclusions

The research programme sketched out in this paper offers a framework of feature representation which assumes underspecification, a finite set of universal distinctive features, and makes strong predictions concerning synchronic phonological systems as well as language processing. The research question is how the mental representation governs language comprehension. The feature geometry argued for by FUL assumes, on the one hand, that consonants and vowels share features, and on the other, that there are no feature dependencies (other than inherent ones). All features except those defining the major classes (SONORANT/OBSTRUENT and VOCALIC/CONSONANTAL) are monovalent. Asymmetries predicted by FUL are held to have consequences for phonological feature inventories and contrasts as well as for speech perception and language comprehension. We have provided psycholinguistic and EEG experimental data arguing for a processing system which tries to resolve variation in the signal on the basis of sparse underspecified phonological representations. We hope that the approach presented here will encourage a fresh look at constrained lexical representations and their consequences for phonological systems and language processing.

Acknowledgements

This research was partially supported by the *Deutsche Forschungsgesellschaft* (SFB 471 and *Leibniz Preis* to Aditi Lahiri) and the *Ministry of Science, Research, and the Arts of Baden-Württemberg*. We would like to thank Muna Pohl and Vincent Evers for assisting in preparing and running the experiments and two anonymous reviewers for giving valuable comments on a prior version of this article.

References

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology*, *5*, 183–207.
Blumstein, S., & Stevens, K. (1980). Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America*, *67*, 648–662.
Borowsky, T. (1987). Antigemination in English phonology. *Linguistic Inquiry*, *18*, 671–678.
Calabrese, A. (1995). A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry*, *26*, 373–463.
Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.
Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook*, *2*, 225–252.
Clements, G. N. (1989). *A unified set of features for consonants and vowels*. Unpublished manuscript, Cornell University, Ithaca, NY.
Clements, G. N. (2001). Representational economy in constraint-based phonology. In T. A. Hall (Ed.), *Distinctive feature theory* (pp. 71–146). Berlin: Mouton.
Clements, G. N., & Hume, E. V. (1995). The internal organization of speech sounds. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 245–306). Oxford: Blackwell.
Coenen, E., Zwitserlood, P., & Bölte, J. (2001). Variation and assimilation in German: Consequences of assimilation for word recognition and lexical representation. *Language and Cognitive Processes*, *16*, 535–564.
Dilley, L. C., & Pitt, M. A. (2007). A study of regressive place assimilation in spontaneous speech and its implications for spoken word recognition. *Journal of the Acoustical Society of America*, *122*, 2340–2353.
Eulitz, C., & Lahiri, A. (2004). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience*, *16*, 577–583.

Fikkert, P., & Levelt, C. C. (2008). How does Place fall into place? The lexicon and emergent constraints. In P. Avery, B. E. Dresher, & K. Rice (Eds.), *Contrast in phonology: Theory, perception, acquisition* (pp. 231–270). Berlin: Mouton.
Friedrich, C., Eulitz, C., & Lahiri, A. (2006). Not every pseudoword disrupts word recognition: An ERP study. *Behavioral and Brain Functions*, *2*, 1–36.
Friedrich, C., Lahiri, A., & Eulitz, C. (2008). Neurophysiological evidence for underspecified lexical representations: Asymmetries with word initial variations. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1545–1559.
Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 144–158.
Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 380–396.
Gaskell, M. G., & Marslen-Wilson, W. D. (2001). Lexical ambiguity and spoken word recognition: Bridging the gap. *Journal of Memory and Language*, *44*, 325–349.
Gaskell, M. G., & Snoeren, N. D. (2008). The impact of strong assimilation on the perception of connected speech. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1632–1647.
Ghini, M. (2001a). *Asymmetries in the phonology of Miogliola*. Berlin: Mouton.
Ghini, M. (2001b). Place of articulation first. In T. A. Hall (Ed.), *Distinctive feature theory* (pp. 71–146). Berlin: Mouton.
Gow, D. W. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*, *45*, 133–159.
Gow, D. W. (2002). Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology Human Perception and Performance*, *28*, 163–179.
Gow, D. W. (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics*, *63*, 575–590.
Hall, T. A. (1997). *The phonology of coronals*. Amsterdam: Benjamins.
Halle, M. (1995). Feature geometry and feature spreading. *Linguistic Inquiry*, *26*, 1–46.
Halle, M., Vaux, B., & Wolfe, A. (2000). On feature spreading and the representation of place of articulation. *Linguistic Inquiry*, *31*, 387–444.
Henderson, L., Wallis, J., & Knight, D. (1984). Morphemic structure and lexical access. In H. Bouma, & D. Bouhuis (Eds.), *Attention and performance* (pp. 211–225). London: Erlbaum.
Hyman, L. (1973). The feature Grave in phonological theory. *Journal of Phonetics*, *1*, 329–337.
IPDS. (1995–1997). *The Kiel corpus of spontaneous speech*, Vols. I–III. CD-ROM. Kiel: Institut für Phonetik und digitale Sprachverarbeitung.
Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to speech analysis. The distinctive features and their correlates*. Cambridge, MA: MIT, Acoustics Laboratory, Technical Report No. 13.
Jusczyk, P. (1986). Speech perception. In L. Kaufman, J. Thomas, & K. Boff (Eds.), *Handbook of perception and performance* (pp. 27.1–27.57). New York: Wiley.
Kabak, B. (2007). Hiatus resolution in Turkish: An underspecification account. *Lingua*, *117*, 1378–1411.
Keating, P., & Lahiri, A. (1993). Fronted velars, palatalized velars, and palatals. *Phonetica*, *50*, 73–101.
Kloeke, W. U. S. van Lessen (1982). *Deutsche Phonologie und Morphologie: Merkmale und Markiertheit*. Tübingen: Niemeyer.
Lahiri, A. (2000). Phonology: Structure, representation and process. In L. Wheeldon (Ed.), *Language production* (pp. 165–225). Hove: Psychology Press.
Lahiri, A. (2007). Non-equivalence between phonology and phonetics. In: *Proceedings of the 16th international conference of phonetic sciences*, Saarbrücken (pp. 31–34).
Lahiri, A., & Blumstein, S. E. (1984). A re-evaluation of the feature 'coronal'. *Journal of Phonetics*, *12*, 133–145.
Lahiri, A., & Evers, V. (1991). Palatalization and coronality. In C. Paradis, & J.-F. Prunet (Eds.), *The special status of coronals* (pp. 79–100). San Diego: Academic Press.
Lahiri, A., Gewirth, L., & Blumstein, S. E. (1984). A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study. *Journal of the Acoustical Society of America*, *76*, 391–404.
Lahiri, A., & Marslen-Wilson, W. D. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, *38*, 245–294.
Lahiri, A., & Marslen-Wilson, W. D. (1992). Lexical processing and phonological representation. In G. J. Docherty, & D. R. Ladd (Eds.), *Papers in laboratory phonology II* (pp. 229–254). Cambridge: Cambridge University Press.
Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven, & N. Warner (Eds.), *Labphon 7* (pp. 637–676). Berlin: Mouton.
Levelt, C. C. (1995). Segmental structure of early words: Articulatory frames or phonological constraints. In E. V. Clark (Ed.), *The Proceedings of the 27th annual child language research forum* (pp. 19–27). Stanford: CSLI.
Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.
Marslen-Wilson, W. D., Frost, R., Deutsch, A., Gilboa, O., & Tannenbaum, M. (2000). Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory and Cognition*, *28*, 1277–1288.
Marslen-Wilson, W. D., Nix, A., & Gaskell, M. G. (1995). Phonological variation in lexical access: Abstractness, inference and English place assimilation. *Language and Cognitive Processes*, *10*, 285–308.
McCarthy, J. J. (1988). Feature geometry and dependency: A review. *Phonetica*, *43*, 84–108.
McCarthy, J. J., & Taub, A. (Eds.). (1992). Review of Carole Paradis and Jean-François Prunet (eds.) (1991). The special status of coronals: internal and external evidence. *Phonology*, *9*, 363–370.
Mohanan, K. P. (1993). Fields of attraction in phonology. In J. Goldsmith (Ed.), *The last phonological rule—Reflections on constraints and derivations* (pp. 61–116). Chicago: University of Chicago Press.
Mohanan, K. P., & Mohanan, T. (1984). Lexical phonology of the consonant system in Malayalam. *Linguistic Inquiry*, *15*, 575–602.
Näätänen, R., & Alho, K. (1997). Mismatch negativity (MMN)—The measure for central sound representation accuracy. *Audiology and Neurotology*, *2*, 341–353.
Napps, S. E., & Fowler, C. A. (1987). Formal relationships among words and the organization of the mental lexicon. *Journal of Psycholinguistic Research*, *16*, 257–272.
Nolan, F. (1992). The descriptive role of segments. In G. J. Docherty, & D. R. Ladd (Eds.), *Papers in laboratory phonology II* (pp. 261–280). Cambridge: Cambridge University Press.
Obleser, J., Lahiri, A., & Eulitz, C. (2004). Magnetic brain response mirrors extraction of phonological features from spoken vowels. *Journal of Cognitive Neuroscience*, *16*, 31–39.
Reetz, H. (2000). Underspecified phonological features for lexical access. *Phonus (Institute für Phonetik, Universität des Saarlandes)*, *5*, 161–173.
Sagey, E. (1986). *The representation of features and relations in non-linear phonology. Doctoral Dissertation*. Cambridge, MA: MIT.
Scharinger, M. (2006). *The representations of vocalic features in vowel alternations: Phonological, morphological and computational aspects*. Doctoral Dissertation, University of Konstanz. URL 〈http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-24341〉.
Scharinger, M., Reetz, H., & Lahiri, A. (2009). Levels of regularity in inflected word form processing. *The Mental Lexicon*, *4*, 77–114.
Shaw, P. (1991). Consonant harmony systems: The special status of coronal harmony. In C. Paradis, & J.-F. Prunet (Eds.), *The special status of coronals* (pp. 125–157). San Diego: Academic Press.
Steriade, D. (1995). Underspecification and markedness. In J. Goldsmith (Ed.), *Handbook of phonological theory* (pp. 114–174). Oxford: Blackwell.
Stevens, K., & Blumstein, S. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, *64*, 1358–1368.
Wheeldon, L., & Waksler, R. (2004). Phonological underspecification and mapping mechanisms in the speech recognition lexicon. *Brain and Language*, *90*, 401–412.
Wiese, R. (2000). *The phonology of German*. Oxford: Clarendon Press.
Wurzel, W. U. (1970). *Studien zur Deutschen Lautstruktur (Studia Grammatica VIII)*. Berlin: Akademie-Verlag.
Yip, M. (1991). Coronals, consonant clusters, and the coda condition. In C. Paradis, & J.-F. Prunet (Eds.), *The special status of coronals* (pp. 61–78). San Diego: Academic Press.
Zimmerer, F., Lahiri, A., & Reetz, H. (2008). Harmful reduction?. *Laboratory Phonology (Wellington)*, *11*, 167–168.
Zimmerer, F., Reetz, H., & Lahiri, A. (2009). Place assimilation across words in running speech: Corpus analysis and perception. *Journal of the Acoustical Society of America*, *125*, 2307–2322.