

ZASPiL Nr. 37 – December 2004

**Papers in Phonetics and
Phonology**

Editors: Susanne Fuchs
and Silke Hamann

Table of Contents

<i>English syllabification as the interaction of markedness constraints</i>	1
T. A. Hall (Indiana University)	
<i>Opacity in Tiberian Hebrew: Morphology, not phonology</i>	37
Antony D. Green	
<i>Phonological Phrases in Xhosa (Southern Bantu)</i>	71
Sabine Zerbian (ZAS Berlin)	
<i>What African Languages Tell Us About Accent Typology</i>	101
Laura J. Downing (ZAS Berlin)	
<i>(Un)markedness of trills: the case of Slavic r-palatalisation</i>	137
Marzena Zygis (ZAS Berlin)	
<i>Prosody and Information Structure in Chichewa</i>	167
Laura J. Downing (ZAS Berlin)	
Al Mtenje (University of Malawi)	
Bernd Pompino-Marschall (Humboldt-Universität Berlin)	
<i>The phonetics of stop assibilation</i>	187
T. A. Hall (Indiana University)	
Silke Hamann (ZAS Berlin)	
Marzena Zygis (ZAS Berlin)	
<i>The Hungarian palatal stop:</i>	221
<i>phonological considerations and phonetic data</i>	
Christian Geng (ZAS Berlin)	
Christine Mooshammer (Universität Kiel)	

English syllabification as the interaction of markedness constraints*

T. A. Hall

Indiana University

The present study offers an Optimality-Theoretic analysis of the syllabification of intervocalic consonants and glides in Modern English. It will be argued that the proposed syllabifications fall out from universal markedness constraints – all of which derive motivation from other languages – and a language-specific ranking. The analysis offered below is therefore an alternative to the traditional rule-based analyses of English syllabification, e.g. Kahn (1976), Borowsky (1986), Giegerich (1992, 1999) and to the Optimality-Theoretic treatment proposed by Hammond (1999), whose analysis requires several language-specific constraints which apparently have no cross-linguistic motivation.

1 Introduction

Consider the syllabification of the following (broadly transcribed) English words:

- (1) lazy [leɪ.zi]
patron [peɪ.tɹ̩n̩]
bulky [bʌl.ki]

Although the parsings indicated in (1) are in line with universal principles to be made explicit below, there is a surprising lack of consensus in the literature on English phonology that words like the ones in (1) are syllabified as indicated. For example, while Giegerich (1992) syllabifies phonologically similar words of English as in (1), Kahn (1976) proposes an algorithm for the assignment of syllable structure which makes the /z/ in *lazy* and the /t/ in *patron* (but not the /l/

* Thanks are due to Susanne Fuchs, Antony Green, Silke Hamann, Christiane Kaden, Evan Mellander and Marzena Zygis for their comments on earlier versions of this article. All disclaimers apply.

in *bulky*) ambisyllabic. Hammond (1999) – whose parsings are similar to the ones in Selkirk (1982) – proposes an Optimality-Theoretic analysis of English in which *patron* and *bulky* are parsed [pɛt.r̩n] and [bʌlk.i] respectively.

Any study of syllabification in a single language needs to accomplish three general goals, namely to show that (i) the proposed parsings are supported by various phonological generalizations from the language in question, (ii) the arguments that have been proposed by alternate syllabifications can be reanalyzed in other ways, and (iii) the mechanism whereby segments are parsed into syllables is in line with universal principles. The present study offers an Optimality-Theoretic (henceforth OT; Prince and Smolensky 1993) analysis of syllabification in Modern English – in particular General American – in which the goals in (i)-(iii) are accomplished.

The present treatment is significant for several reasons. First, as indicated in the title, I argue that the proposed syllabifications (as in 1) fall out from universal markedness constraints – all of which derive motivation from other languages – and that syllabification is a process occurring in the surface which does not require steps in a derivation. My analysis can therefore be considered to be an alternative to the traditional rule-based analyses of English syllabification, e.g. Kahn (1976), Gussenhoven (1986), Borowsky (1986), Giegerich (1992), Rubach (1996), and most recently Giegerich (1999). My analysis can also be seen as an alternative to the OT treatment proposed by Hammond (1999), whose analysis requires several language-specific constraints which apparently have no cross-linguistic motivation.

This article is organized as follows. In §2 I present two syllable-based generalizations in English phonology which can be used as reliable tests for how intervocalic clusters should be parsed. In that section I posit an alignment constraint capturing one of these generalizations (ALIGN-3□) which is required in my analysis of English syllabification in §4 and §5. In §3 I posit well-known markedness constraints pertaining to syllable structure and onset well-formedness. §4 consists of a formal analysis of the syllabification of sequences of intervocalic consonants as in (1). In that section I demonstrate that syllabification can be accomplished given the seven constraints posited in §2 and §3 and a language-specific ranking. In §5 I discuss the syllabification of intervocalic sequences of consonants and glides and show that English requires reference to an additional markedness constraint penalizing sonorant plus palatal glide onsets (*□[SG]). In §6 I compare my OT analysis with the one proposed by Hammond (1999) and show that his alternate syllabifications are problematic and that his arguments for these parsings can be analyzed in other ways. §7 concludes.

2 Two syllable-based generalizations

In this section I introduce two syllable-based generalizations from English which I refer to below as evidence for and against various syllable parsings. These two generalizations are the distribution of the allophones of /p t k/, especially glottalization (in §2.1), and the distribution of superheavy syllables (in §2.2).

1.1 Allophones of /p t k/

The following data illustrate that the glottalized allophones of /p t k/ (= [p^ʔ t^ʔ k^ʔ]) occur word-finally (in 2a) and as the first C in word-medial clusters like the ones in (2b). I follow earlier authors who have argued that the two contexts in (2) can be collapsed into syllable-final (or, alternatively, rhymal) position (see, for example, Kahn 1976: 84ff., Gussenhoven 1986, Nespor and Vogel 1986: 77-78, Giegerich 1992: 220-221, Jensen 1993: 298ff., and Kenstowicz 1994: 69). See my analysis in §4.3.2 below, where I show how the clusters in (2b) are heterosyllabified.

(2)	a.	p	cap	[kæp ^ʔ]
		t	cat	[kæt ^ʔ]
		k	lack	[læk ^ʔ]
	b.	tk	catkin	[kæt ^ʔ .kɪn]
		tl	atlas	[æt ^ʔ .ləs]
		tm	utmost	[ʌt ^ʔ .moust]
		km	acme	[æk ^ʔ .mi]
		pn	shrapnel	[ʃɹæp ^ʔ .nɪ]
		tn	chutney	[tʃʌt ^ʔ .ni]
		kn	acne	[æk ^ʔ .ni]

The important point is that if /p t k/ are glottalized then it follows that they are in syllable-final position.¹

Some linguists argue that the rule required for the aspirated and flapped allophones also makes reference to syllable structure (e.g. Kahn 1976). I assume – following earlier treatments of Kiparsky (1979), Jensen (2000) and Davis and

¹ I account for the occurrence of [p^ʔ t^ʔ k^ʔ] in rhymal position as the consequence of an undominated constraint licensing the (nondistinctive) feature [+release] in syllable-initial position. If only unreleased segments like [p^ʰ t^ʰ k^ʰ] are glottalized then [p^ʰ t^ʰ k^ʰ] in the rhyme surface as [p^ʔ t^ʔ k^ʔ]. Since the constraints required in the analysis just described do not bear directly on the syllabification of intervocalic consonant clusters I ignore them here.

Cho (2003) – that /p t k/ are aspirated in foot-initial position, e.g. the /t/ in a word like *attain* (/ətɛɪn/) is aspirated because it is parsed [ə.(t^hɛɪn)_F]. I also assume an analysis for Flapping like the one proposed in Kahn (1976), according to which /t d/ are flapped when ambisyllabic. Although I do not present a formal analysis in the present article which predicts which intervocalic C's are made ambisyllabic (e.g. the /t/ in *later*) I provide some remarks in §4.2 on what a possible treatment might look like.

1.2 Superheavy syllables

A second syllable-based generalization for English can be gleaned from the examples in (3) (from Hall 2001), in which the distribution of superheavy (=VXC) syllables is summarized. The 'X' in VXC is a variable ranging over C (= consonant), G (=glide) and V (= vowel); hence VXC syllables subsume V:C (e.g. *seem*), VGC (e.g. *house*), and VCC (e.g. *farm*).

- (3) a. Three contexts in which superheavy (=VXC) syllables occur:
- | <i>context</i> | <i>VXC</i> |
|---------------------------------|--------------------|
| (i) before a word boundary | <u>arm</u> |
| (ii) before a compound boundary | <u>arm</u> -chair |
| (iii) before -CV(C) suffixes | doub <u>t</u> -ful |
- b. One context in which VXC cannot occur:
- | <i>context</i> | <i>VXC</i> |
|-----------------------|--------------------|
| (i) morpheme-internal | * <u>areel</u> .ba |

In (3a) we see that VXC syllables occur word-finally (in i), before a compound boundary (in ii) and before consonant-initial suffixes (in iii). By contrast, VXC generally does not occur within morphemes (see 3b). In morpheme-internal position heavy (i.e. VX) syllables occur regularly, e.g. *ego*, *garden*.²

In Hall (2001) it is argued that the stem in the three contexts in (3a) constitute separate phonological (=prosodic) words and that the evidence for the prosodic parsings required for words with the phonological and morphological

² The generalizations in (3) are also discussed by other authors, but their analyses differ from the one I adopt below. For example, Borowsky (1986, 1989) argues that the distribution of VXC falls out if the final C in a word is extrasyllabic and if the morphological rules of English are situated in two distinct lexical levels. Hammond (1999) captures the distribution of English VXC by adopting unconventional syllabifications, as I show in §6 below. Harris (1994) proposes a Government Phonology analysis of English which accounts for the distribution of VXC. See Hall (2001, 2002) for criticisms of Borowsky (1986, 1989) and §6 for criticisms of Hammond.

structure in (3a) correlates with independent evidence based on allomorphy in English morphology (see Raffelsiefen 1999). Given the prosodic constituent phonological word, the descriptive generalization can be made in (4a), which captures the facts in (3). This generalization is expressed formally with the constraint in (4b), which says that the right edge of a superheavy syllable (analyzed in Hall 2001 as trimoraic) aligns with the right edge of a phonological word.

- (4) a. A superheavy syllable must occur at the end of a phonological word.
b. ALIGN-3_σ: ([σσσ]_σ, right, phonological word, right)

Nonoccurring examples like **areelba*, **agelmda* violate ALIGN-3_σ because a trimoraic syllable would be internal to a pword.

There are two sets of systematic exceptions to (4) which I mention here because of their relevance to the following analysis: (i) V:C_a.C_bV (e.g. *chamber*, *shoulder*, *fealty*), and (ii) VC_aC_b.CV (e.g. *empty*, *antler*) occur iff C_aC_b is a sequence of nasal or liquid plus homorganic stop. Hall (2001) argues that the V:C_a and VC_aC_b sequences in (i) and (ii) are parsed exceptionally as bimoraic and that they therefore satisfy ALIGN-3_σ.

Although the bulk of the work in syllabifying intervocalic consonant clusters is accomplished with the markedness constraints I posit in §3, I show in §4 and §5 that ALIGN-3_σ is also necessary to rule out various incorrect forms. Since ALIGN-3_σ refers to moraic structure the output forms in the tableaux I posit need to contain moraic structure in addition to the segmental structure. For reasons of space I refrain from presenting nonlinear structures and simply refer the reader to Hall (2001, 2002) for a treatment of the moraic structure of Modern English. In the analysis below the constraint ALIGN-3_σ will be understood to be violated in a VCC.CV sequence, provided that the syllable-final CC cluster is not a stop plus homorganic nasal or lateral.

The upshot of the preceding discussion is that the generalization expressed in (4) pertaining to the distribution of VXC syllables and in (2) concerning the distribution of the glottalized allophones can be used as a test for the parsings of intervocalic consonant clusters.

3 Markedness constraints

In this section I posit several well-known markedness constraints pertaining to syllable structure which are required in order to syllabify sequences of intervocalic consonants and glides.

Four universal markedness constraints are posited in (5) (from Prince and Smolensky 1993, after Jakobson 1962, and Venneman 1988):

- (5) a. ONSET: Syllables have onsets.
 b. NOCODA: The syllable is open
 c. NOCOMPONSET: At most one segment occurs syllable-initially
 d. NOCOMPCODA: At most one segment occurs syllable-finally.

All four constraints are blatantly violated in the surface in many English words. Two sample words and the respective violations of the four constraints in (5) are presented in (6):

(6)	<i>example</i>	<i>constraints violated</i>
	ink [ɪŋk.]	ONSET, NOCODA, NOCOMPCODA
	blink [blɪŋk.]	NOCOMPONSET, NOCODA, NOCOMPCODA

Clearly English ranks the faithfulness DEP-V and MAX-C (see McCarthy and Prince 1995) in (7a) which prevent the insertion of vowels and the deletion of consonants respectively higher than the constraints in (5). This ranking is stated in (7b), where I refer to the two constraints in (7a) collectively as FAITH.

- (7) a. DEP-V: A vowel in the output corresponds to a vowel in the input.
 MAX-C : A consonant in the input corresponds to a consonant in the Output.
 b. FAITH » NOCOMPONSET, NOCODA, NOCOMPCODA, ONSET, NOCODA

The ranking in (7b) accounts for why English words with marked margins can surface as such. For example, a word like [ɔːɡ.] (from /dag/) is optimal in spite of the NOCODA violation because MAX-C and DEP-V are satisfied. By contrast, candidates like [dɑːgə.] or [dɑː.] – despite being in line with NOCODA – lose out because they violate DEP-V and MAX-C respectively. Since faithfulness constraints as in (7a) are not important in my analysis I ignore them in the remainder of this article.

In addition to the constraints in (5), any analysis of English requires a constraint which accounts for onset well-formedness. Occurring onsets like the ones in (8a) would satisfy this constraint, while the ones in (8b-c) would not:

(8) a.	□[tɹ	trip	b.	*□[ɹt	c.	*□[tɹ
	□[pɹ	play		*□[ɹp		*□[tɹ
	□[pɹ	puny		*□[ɹp		*□[nɹ

The well-formedness of the clusters in (8a) and the ill-formedness of the ones in (8b) is clearly a function of sonority, given the Sonority Hierarchy in (9) (see Clements 1990 for a similar version and Selkirk 1984, Vennemann 1972,

Hooper 1976, and Dogil and Luschützky 1990 for earlier ones) and the constraint in (10) (from Selkirk 1984: 116):

- (9) SONORITY HIERARCHY ('>' = 'more sonorous than'):
vowels > glides > rhotics > laterals > nasals > fricatives > stops
- (10) SONORITY SEQUENCING GENERALIZATION (SSG):
“In any syllable, there is a segment constituting the syllable peak that is preceded and/or followed by a sequence of segments with progressively decreasing sonority values”.

However, an examination of the ill-formed onset clusters in (8c) reveals that the constraint governing onset well-formedness cannot simply be the SSG alone because these sequences are well-formed with respect to sonority. For purposes of the present article I require a constraint referring to ONSET WELL-FORMEDNESS (OSW), which can be thought of as an abbreviation for the SSG in (10) or as the constraints ruling out sequences like the ones in (8c). The list in (8c) is not intended to be complete. Some additional restrictions that fit into this category will be discussed in the course of my analysis.

- (11) ONSET WELL-FORMEDNESS (OWF):
A cluster of two or more consonants or glides in syllable-initial position must conform to the SSG or to the constraints in (8c).

See Green (2003), who shows that the constraints governing onset well-formedness can occupy different niches in various language-specific constraint hierarchies. In the present treatment of English there is no reason for assuming that the SSG and the constraints in (8c) need to be split up in this way and therefore I refer to all of them collectively as OWF.³

4 Syllabification of intervocalic consonant clusters

4.1 Introduction

In this section I present a formal OT treatment of the syllabification of intervocalic clusters of consonants of English. In my analysis I divide the consonant sequences into three categories, as exemplified in (12). Here and

³ The ill-formedness of onset clusters like *tl* and *tn* in (8c) is sometimes attributed to the LAW OF INITIALS, which Vennemann (1988: 32) formalizes as follows: “Word-medial syllable heads are the more preferred, the less they differ from possible word-initial syllable heads of the language system”. (The term “syllable head” is synonymous with “syllable-initial cluster”). For OT treatments of English in which the LAW OF INITIALS is assumed see Raffelsiefen (1999) and Hall (2001).

below ‘C’ stands for a consonant, and ‘V’ represents a syllable peak (i.e. short vowel, syllabic sonorant, long vowel, diphthong). Unless otherwise noted the data are drawn from General American English.

- (12) a. V.CV sequences:
 V.CV (e.g. lazy)
 b. V.CCV sequences:
 tautosyllabic CC clusters:
 V.CCV (e.g. patron)
 heterosyllabic CC clusters:
 VC.CV (e.g. bulky)
 c. VCCC(C)V sequences:
 VC.CCV (e.g. empress)
 VC.CCCV (e.g. instrument)

In §4.2-§4.4 I discuss words like the ones in (12a-c) respectively. The most significant point I make below is that the variable parsings in (12a-c) fall out from the constraints in §2 and §3 and a language-specific ranking.

In the following analysis I only consider word-internal and not across-word syllabification. I also do not discuss the alignment constraints needed to account for the domain of syllabification (e.g. only within but not across each part of compounds). As I note below, many previous analyses of English have argued that certain intervocalic consonants are ambisyllabic (e.g. Kahn 1976). Since the generalizations I capture in my treatment can be done so without ambisyllabicity, I do not attempt to posit constraints and rankings which predict its occurrence. However, at several points in the following sections I mention why ambisyllabicity is often posited and suggest how my analysis could be modified to accommodate the additional data.

4.2 *The syllabification of VCV*

In this section I show that VCV is always parsed as V.CV and that this syllabification is a consequence of two universal markedness constraints ONSET and NOCODA.

The English examples in (13) contain a VCV sequence. In (13a) the first vowel is long and in (13b) it is short.

- (13) a. V:CV □ V:.CV
 g ego [i:.go]
 p open [ou:pən]
 z lazy [leɪ.zi]

- b. VCV □ V.CV
 g jagged [dʒæ.gəd]
 z wicked [wɪ.kəd]
 p happy [hæ.pi]

There are two reasons why I syllabify the words in (13) as V.CV (as opposed to VC.V). First, the C in this context is never glottalized if it is /p t k/. And second, the parsing VC.V would violate the generalization in (4) governing the distribution of superheavy syllables when the first V is long.⁴

The parsings in (13) fall out from the markedness constraints ONSET and NOCODA in (5a-b). This point is illustrated in the tableau in (14) for *lazy*.

(14)

	ONSET	NOCODA
a. □ [leɪ.zi]		
b. [leɪz.i]	*!	*

The ranking between the two constraints ONSET and NOCODA is not crucial, since the incorrect form in (14b) is penalized twice and the winner in (14a) not at all.⁵

It should be noted that the syllabification V.CV in (14) is predicted to be correct regardless of the stress pattern of the word; hence in the present analysis V.CV is also selected as optimal when the second vowel bears prominence (e.g. *about*). Some authors (e.g. Kahn 1976, Gussenhoven 1986) have argued that

⁴ Some linguists have argued that the intervocalic C in words like the ones in (13b) should be ambisyllabic because short vowels like [æ] otherwise only occur in closed syllables (see, for example, Giegerich 1992: 172, Hammond 1999). If this is a true syllable-based generalization then one could either analyze the post-short vowel C in words like *happy* as ambisyllabic (Giegerich 1992) or as a geminate consonant (Hammond 1999). I leave these two possibilities open for further study. The nonglottalization of an ambisyllabic/geminate consonant follows from the proposals made through the years on ‘geminate inalterability’, e.g. Hayes (1986).

⁵ Note that words which end in /t̚/ are also predicted to syllabify as V.CV, e.g. *eaten* as [i:.t̚]. Words like *eaten* are apparently problematic for my analysis because the (syllable-initial) /t/ surfaces as [t̚] (recall from §2.2 that the glottalized allophones only surface syllable-finally). I maintain that the predicted parsing [i:.t̚] is correct and that the /t/ is glottalized because – as the first part of a nasally released stop – it must by definition be orally unreleased (see Ladefoged 1993: 53 for discussion of nasally released stops in English – in his terminology ‘nasal plosion’). From a formal point of view the reason /t/ in /t̚/ sequences is unreleased is a consequence of an (undominated) constraint saying that /t̚/ sequences are nasally released stops. A precise formalization of this constraint exceeds the goals of the present paper.

English has stress-dependent rules of ambisyllabification referred to in §2.2 above (which are not to be confused with the non-stress related argument for ambisyllabicity described in note 4). In the present analysis I concentrate solely on the constraints responsible for tauto- or heterosyllabifying intervocalic sequences and not on the ones which might be needed to create ambisyllabic structures. However, for the sake of thoroughness I conclude this section with some brief remarks on the motivation for stress-dependent ambisyllabicity and what an analysis thereof might look like.

Following Kahn (1976) (and Gussenhoven 1986) the generalization is that within English words the C in V(:)CV is ambisyllabic if the second V is unstressed. This type of ambisyllabicity is said to be confirmed by native speaker intuitions and is simultaneously argued to be the environment for Flapping, i.e. /t d/ in American English are flapped only when ambisyllabic. (However, see Steriade 2000: 325ff. for an OT constraint for Flapping which does not make reference to ambisyllabicity.) For example, in the two words *later* and *attain* only the /t/ in the former item is analyzed as ambisyllabic (given its stress pattern) and hence flapped. One could treat the C in 'V(:)CV as ambisyllabic given an OT analysis like the one proposed by Green (1997) for Irish, who posits a constraint called CLOSECONTACT – inspired by Trubetzkoy's *Silbenschnittkorrelation* – according to which a C within a foot (i.e. in 'V(:)CV) is tautosyllabic with the first vowel. If CLOSECONTACT outranks NOCODA (see tableau 16), then the ambisyllabic candidate would be selected as the optimal one.⁶

4.3 The syllabification of VCCV

In this section I argue that two adjacent intervocalic consonants are parsed either tauto-syllabically (i.e. syllable-initially), e.g. [tɹ] in *patron*, or heterosyllabically, e.g. [lk] in *falcon*. The alternate parsings fall out from the seven markedness constraints above and a language-specific ranking.

⁶ Word-final /t d/ are also flapped in American English if the following word begins with a vowel, e.g. the /t/ in *meet Ann* (vs. the aspirated /t/ in *see Ted*). See McCarthy and Prince (1993: 130-131), who argue that a word-final consonant of English is made ambisyllabic by ranking ONSET ahead of an alignment constraint which says that the left edge of a stem coincides with the left edge of a phonological word and a constraint requiring that a phonological word ends in a consonant. Given this ranking the /t/ in *meet Ann* (but not the /t/ in *see Ted*) would be made ambisyllabic and would therefore be flapped.

4.3.1 Tautosyllabic CC clusters

VCCV is parsed as V.CCV in two instances: (i) the CC exhibits a sonority rise (going from left to right) and the CC is a permissible word-initial cluster of English, or (ii) the CC consists of [s] + voiceless stop.

The English examples in (15) contain an intervocalic CC sequence for category (i) (in 15a) and category (ii) (in 15b). The phonetic transcriptions in (15) and below are broad, unless otherwise noted.

(15) a. obstruent + liquid

pɹ	apron	[eɪ.pɹɪ]
tɹ	patron	[peɪ.tɹɪ]
kɹ	sacred	[seɪ.kɹəd]
bɹ	cobra	[kou.bɹə]
dɹ	hydrant	[haɪ.dɹənt]
gɹ	flagrant	[fleɪ.gɹɪnt]
kl	cyclone	[saɪ.kloun]

b. s + voiceless stop:

st	feisty	[faɪ.sti]
sp	aspect	[æ.spɛkt]
sk	fiscal	[fɪ.skɪ]

Recall from note 4 that the consonant following a short vowel (as in the final two examples in 15b) might in fact close the first syllable if this consonant is analyzed as ambisyllabic.

Given the ranking NOCODA » NOCOMPONSET the syllabification in (15a) is correctly predicted to be optimal. This point is illustrated in the tableau in (16) for the word *cobra*:

(16)		NOCODA	NOCOMPONSET
a.	□ [kou.bɹə]		*
b.	[kou.bɹ.ə]	*!	
c.	[kou.b..ɹə]	*!	

An examination of this tableau reveals that (16a) is optimal because the closest competitors in (16b-c) violate NOCODA. Note that the same results could be achieved by substituting NOCODA for ALIGN-3□. Although this alternative would work in words like *cobra*, in which the CC cluster is preceded by a long vowel, it fails to tautosyllabify CC when preceded by a short vowel (see below).

The constraints in (16) also predict that intervocalic /st/ as in (15b) is tautosyllabic. (There is general agreement in English phonology that [sp st sk]

clusters cannot be heterosyllabified because the [s] consistently blocks the aspiration of the following stop; see, for example, Kahn 1976, Selkirk 1982, Hammond 1999). This syllabification V.stV is illustrated in the tableau in (17) for the word *feisty*:

(17)

	NoCODA	NoCOMPONSET
a. □ [faɪ.sti]		*
b. [faɪst.i]	*!	
c. [faɪs.ti]	*!	

In this tableau we can observe that (17b-c) lose out to (17a) because they are not in line with NoCODA.

The tableau in (18) for the word *aspect* is important because it shows that the intervocalic sC cluster is correctly tautosyllabified even if a short vowel precedes:

(18)

	NoCODA	NoCOMPONSET
a. □ [æ.spɛkt]		*
b. [æs.pɛkt]	*!	
c. [æsp.ɛkt]	*!	

At first glance, one might assume that the ranking NoCODA » NoCOMPONSET incorrectly predicts that *all* fricative-stop sequences are parsed V.CCV, i.e. not only the /st/ in *feisty* but also, for example, the /ft/ in *after*. I show in §4.3.2 below that fricative+stop sequences other than s+stop are heterosyllabified (i.e. VC.CV) and that this parsing follows from the ranking of two constraints over NoCODA.⁷

⁷ Antony Green (personal communication) points out that one curious gap suggests that [sp st sk] are actually heterosyllabic: only [st] but not [sp sk] can be preceded by long vowels (e.g. *Easter*). Given the parsing Vs.pV and Vs.kV the gaps *V:spV and *V:skV would follow from the generalization in (4), and V:s.tV would be allowed if the homorganicity restriction discussed in §2.2 included not only nasal/liquid + homorganic stop but also fricative+homorganic stop. That V:sp and V:sk are also illicit word-finally might then follow from the treatment in Hall (2001), which is independently required to ban other illegal word-final sequences, e.g. *V:lk, *V:lp.

While I have at present no analysis for [sp st sk] which accounts for both the phonotactic facts described in the preceding paragraph as well as the lack of aspiration of the stop portion, I suggest tentatively that the [s] in intervocalic [sp st sk] is ambisyllabic, although I leave open how a formal analysis would predict such parsings.

1.1.2 Heterosyllabic CC clusters

There are three types of heterosyllabic CC sequences:

- (19) a. CC clusters in which the second C is less sonorous than the first (e.g. /lk/ in *bulky*)
b. CC clusters in which the second C is more sonorous than the first and which are not well-formed onsets of English e.g. /tl/ in *atlas*)
c. CC clusters in which the two C's are equally sonorous (e.g. /pk/ in *napkin*)

Note that none of the CC clusters in (19a-c) satisfies OWF.

The data set in (20) contains words of the structure (19a). I have placed these words into two groups: (i) sonorant + obstruent sequences (in 20a), and (ii) fricative + stop sequences (in 20b), where the fricative is anything other than [s]. (Examples with [s] plus stop between vowels were discussed in 17-18). As indicated in the phonetic transcriptions, the clusters listed here are analyzed as heterosyllabic.

- (20) a. sonorant + obstruent
- | | | |
|----|--------|-----------|
| lk | bulky | [bʌl.ki] |
| lt | guilty | [gɪl.ti] |
| lp | pulpit | [pʌl.pɪt] |
| nv | envy | [ɛn.vi] |
| ŋk | ankle | [æŋ.kl] |
- b. fricative (other than [s]) + stop
- | | | |
|----|-------------|--------------|
| ft | after | [æf.tɪ] |
| zb | husband | [hʌz.bʌnd] |
| zd | mazdar | [mæz.dɑː] |
| ʃk | shish kebab | [ʃɪʃ.kə.bab] |

The syllabification VC.CV (as opposed to VCC.V) in (20) derives support from two sets of facts. First, the /p t k/ in the examples in (20) are not glottalized, and second the CC sequences in (20) can only be preceded by short vowels (provided that CC is nonhomorganic, recall §2.2). Were long vowels permissible in this position, then the otherwise valid generalization in (4) involving the distribution of superheavy syllables would be violated.

The constraints posited above conspire to predict the correct syllabification for words like the ones in (20a-b), as shown in the tableaux in (21-22) for the words *bulky* and *after* respectively:

(21)

	ALIGN-3□	OWF	NoCODA	NoCOMPONSET
a. □ [bʌl.ki]			*	
b. [bʌ.lki]		*!		*
c. [bʌlk.i]	*!		*	

(22)

	ALIGN-3□	OWF	NoCODA	NoCOMPONSET
a. □ [æf.tɹ]			*	
b. [æ.ftɹ]		*!		*
c. [æft.ɹ]	*!		*	

An examination of these tableaux reveals that the (b) and (c) candidates lose out to the respective (a) forms because they violate either OWF or ALIGN-3□. Note that the (c) forms also violate the constraint ONSET (recall 5a). The reason ALIGN-3□ cannot be substituted for ONSET is that for certain words ungrammatical candidates cannot be ruled out without reference to ALIGN-3□, e.g. candidate (29b) below violates only ALIGN-3□ but not ONSET.

The additional data in (23) – like the ones in (15a) – contain an intervocalic CC, where the second C is more sonorous than the first (=19b). The phonetic transcriptions below are narrow, since they indicate the glottalized allophones of voiceless stops.

(23) VCCV □ VC.CV

tl	atlas	[æt ^ʔ .ləs]
tm	utmost	[ʌt ^ʔ .moust]
km	acme	[æk ^ʔ .mi]
pn	shrapnel	[ʃɹæp ^ʔ .nɪ]
tn	chutney	[tʃʌt ^ʔ .ni]
kn	acne	[æk ^ʔ .ni]

In contrast to the examples in (15a), the words in (23) are parsed VC.CV and not V.CCV. One piece of evidence that the adjacent consonants in (23) are heterosyllabic is that /p t k/ are glottalized, as indicated in the phonetic transcriptions. A second argument for the heterosyllabification of the clusters in (23) is that the first C in CC is consistently preceded by a short vowel (see Hall 2001 for additional examples showing this and for a formal analysis of the V:C.CV gap). Were long vowels permissible in such words then superheavy

syllables would occur within a phonological word, contrary to the generalization in (4).

A comparison of the obstruent + sonorant consonant sequences in (15a) and (23) reveals a difference between the two: The consonant sequences in the latter words are in line with the SSG in (9), but not with the filters in (8c). These examples therefore show that onset well-formedness cannot be reduced to the SSG alone. Consider the tableau for the word *chutney*, which is a representative example of the words in (23):

(24)

	ALIGN-3σ	OWF	NoCODA	NoCOMPONSET
a. σ [tʃʌt.ni]			*	
b. [tʃʌ.tni]		*!		*
c. [tʃʌtn.i]	*!		*	

In this tableau we see the same results as in the ones in (21-22): Candidate (24a) wins out because the two closest competitors are not in line with OWF and ALIGN-3σ.

The final set of data in this section involves a sequence of two intervocalic C's which are equally sonorous (=19c), as in (25). Here we see combinations of two stops (in 25a), two fricatives (in 25b), and two nasals (in 25c).

(25) VCCV σ VC.CV

a. stop + stop

pt	optical	[ɒp.tɪ.kəl]
kt	actor	[æk.tɪ.ə]
pk	napkin	[næp.kɪn]

b. fricative + fricative

sf	asphalt	[æs.fɒlt]
----	---------	-----------

c. nasal + nasal

mn	amnesty	[æm.nə.sti]
nm	enmity	[ɛn.mɪ.ri]

I parse words like the ones in (25) as VC.CV for two reasons. First, the first C can be glottalized if it is a voiceless stop⁸ and second, the CC clusters in (25) can only be preceded by short vowels. Again, gaps such as [V:k.tV] and [V:m.nV] follow from (4).

⁸ As pointed out by Kahn (1976), the glottalization of American English /t/ occurs more readily than the glottalization of /p/ and /k/. For this reason the glottalization of /p/ and /k/ in the examples in (24a) is not always obvious to many native speakers.

The analysis presented up to this point correctly selects the parsing VC.CV in (25) as optimal, as illustrated in the tableau in (26) for *napkin*:

(26)

	ALIGN-3□	OWF	NOCODA	NOCOMPONSET
a. □ [næp.km]			*	
b. [næ.pkm]		*!		*
c. [næpk.m]	*!		*	

Again, the first candidate in this tableau wins out over the final two because it satisfies the highest two constraints.

To summarize, the constraints ALIGN-3□ and OWF, when ranked over NOCODA, correctly syllabify all of the English data presented up to this point. Note too that these constraints account for the principle referred to in the literature as Onset Maximization, i.e. a sequence VCCV is parsed V.CCV only if CC is a possible onset of the language in question.

1.4 The syllabification of VCCCV and VCCCCV

In this section I examine the syllabification of (monomorphemic) English words containing three or four intervocalic consonants. (There are apparently no examples with five or more adjacent consonants). The data are divided into four categories, which have been summarized in (27):

- (27) a. $C_a C_b C_c$ clusters in which $C_b C_c$ are well-formed onsets of English and in which $C_b C_c$ rise in sonority (from left to right)
 b. $C_a C_b C_c$ clusters in which $C_b C_c$ are not well-formed onsets of English and in which $C_b C_c$ rise in sonority (from left to right)
 c. $C_a C_b C_c$ clusters in which $C_b C_c$ form a sonority plateau
 d. $C_a C_b C_c C_d$ clusters

All occurring three member clusters of English will be shown to belong to one of (27a-c).⁹ I show below that (27a-b, d) can be syllabified given the ranking ALIGN-3□, OWF » NOCODA shown above. Context (27c) will be shown to require the constraint ONSET.

⁹ Note that there are a number of gaps, e.g. there are no $C_a C_b C_c$ clusters in which C_c is less sonorous than C_b and C_b is less sonorous than C_a , but where $C_b C_c$ is not a well-formed onset (e.g. the hypothetical [ɪlmta]). Such gaps are not important for the present analysis and will therefore be ignored. See Hall (2001), who shows that these gaps often follow from the generalization in (4).

1.1.1 Three member clusters (=27a)

The following words illustrate (27a). In the first column we see the relevant sequence and on the corresponding line of the second column a representative example. These words can have either a nasal (in 28a) or a liquid (in 28b) as the first C.

- (28) VCCCV □ VC.CCV
- a. mpɪ empress
 - mpl accomplish
 - mbɪ ambrosia
 - mbl emblem
 - mfl pamphlet
 - nθɪ anthrax
 - ŋkɪ pancreas
 - ŋgɪ angry
 - b. ldɪ caldron
 - lgɪ pilgrim
 - lpɪ culprit
 - ɹtɪ cartridge
 - ɹθɪ arthritis

I analyze the syllabification of all of the words in (28) as VC.CCV. Language specific support for this parsing (as opposed to the syllabification VCC.CV) comes from the distribution of the glottalized allophones of /p t k/: since the second C in the clusters in (28) never surfaces as glottalized (if it is a voiceless stop) we can conclude that it must be syllable-initial and not syllable-final. The generalization in (4) can also be used as a diagnostic for the parsing as indicated in (28) (and against the syllabification VCC.CV), since the first V in VC.CCV can only be short, provided that the adjacent C's are nonhomorganic (recall §2.2).

Given the constraints posited above my analysis correctly syllabifies the words in (28) as VC.CCV. This is illustrated in the tableau in (29) for the word *pilgrim*:

(29)

	ALIGN-3□	OWF	NoCODA	NoCOMPONSET
a. □ [pɪl.gɪm]			**	*
b. [pɪlg.ɪm]	*!		**	
c. [pɪ.lgɪm]		*!	*	*
d. [pɪlgɪ.ɪm]	*!		**	

Candidates (29b, d) lose out because they violate ALIGN-3 and (29c) because it is not in line with OWF.

The following examples contain intervocalic CCC sequences in which all three segments are obstruents. Like the clusters in (27) the ones in (30) illustrate (26a).

- (30) VCCCV VC.CCV
 ksp expand
 kst extend
 pst capstan

Given the constraints above the words in (30) are syllabified as indicated. This is shown in tableau (31) for the word *extend*:

(31)

	ALIGN-3	OWF	NoCODA	NoCOMPONSET
a. [k.stend]			**	*
b. [I.kstend]		*!	*	*
c. [iks.tend]	*!		**	
d. [Ikst.end]	*!		**	

Forms (31c, d) are clearly less harmonic than (31a) because they violate ALIGN-3 and (31b) violates OWF.

1.1.2 Three member clusters (=27b)

The following words illustrate intervocalic $C_aC_bC_c$ clusters in which C_bC_c are not well-formed onsets of English and in which C_bC_c rise in sonority (from left to right) (=27b):

- (32) VCCCV VCC.CV
 ntl antler
 ntn vintner
 ndl chandler

Since the /t/ in the first three examples is (optionally) glottalized before a nasal or a lateral (see Kahn 1976), we can conclude that the syllabification indicated in (32) is correct. Recall from §2.2 that the VCC sequences in (32) do not violate (4) because the first two C's consist of a nasal or liquid plus a homorganic stop (cf. *empty*).

The syllabification VCC.CV in (32) obtains given the constraints above. This is illustrated in the tableau in (33) for the word *antler*, which is representative of (32):

(33)

	ALIGN-3□	OWF	NOCODA	NOCOMPONSET
a. □ [ænt.lɹ]			*	
b. [æn.tlɹ]		*!	*	*
c. [æ.nɫɹ]		*!		*
d. [æntl.ɹ]	*!		*	

In this tableau we see that (33b-d) are less harmonic than (33a) because they violate either OWF or ALIGN-3□.¹⁰

The tableau in (33) is important because it illustrates why ALIGN-3□ cannot be substituted for NOCOMPCODA. Were NOCOMPCODA the highest ranked constraint in (33) then the winner in (33a) would incur a violation due to the [nt] cluster in the coda of the first syllable and candidate (33b) would incorrectly be selected as optimal. See below for additional examples of English words which can only be successfully syllabified with ALIGN-3□ and not NOCOMPCODA. This being said, it will be demonstrated in §5.2.1 below that my analysis must refer to NOCOMPCODA to account for the syllabification of VCCjV.

1.1.3 Three member clusters (=27c)

The words in (34) consist of three intervocalic consonants C_aC_bC_c in which C_bC_c exhibit a sonority plateau and therefore illustrate (27c). Note that C_bC_c is also not in line with OWF.

(34) VCCCV □ VCC.CV

mpt	empty
mpk	pumpkin
mpt̆	sumptuous
ŋkt̆	juncture
ŋkt	plankton

¹⁰ The one three member cluster like the ones in (32) which I have omitted is [ɹtn], e.g. *partner*. Since the [t] is glottalized in such examples the parsing is VCC.CV. From a purely descriptive point of view this example does not conform to the generalization in (4a). Since words like *partner* are idiosyncratic exceptions to (4a) my analysis is not intended to explain them.

I analyze the syllabification of the words in (34) as VCC.CV – a parsing which derives support from the fact that the second C can surface as glottalized if it is /p t k/. Recall from §2.2 that words like *empty* are systematic exceptions to ALIGN-3 because the first two C’s in these five sequences are homorganic.

The syllabification VCC.CV in (34) is correctly selected as optimal given the constraints above, as illustrated in the tableau in (35) for *empty*. Examples like this one are important because they illustrate the necessity of the constraint ONSET, which must be ranked below OWF but is unranked with respect to NoCODA:

(35)

	ALIGN-3	OWF	ONSET	NoCODA	NoCOMPONSET
a. ∅ [emp.ti]			*	*	
b. [ε.mpti]		*!	*		*
c. [εm.pti]		*!	*	*	*
d. [empt.i]			**!	*	

In this tableau we see that (35b-c) are not optimal because they violate OWF. A comparison of the winner in (35a) with the closest competitor in (35d) reveals that the latter one loses out to the former because it incurs two ONSET violations. (Note that the addition of ONSET to the constraint hierarchy does not effect the outcome in the tableaux presented above.)

1.1.4 Four member clusters (=27d)

Hammond (1999: 83) notes that there are few examples of English words with four intervocalic consonant clusters in which these consonants are tautomorphemic. Some examples are provided in (36). Two types of examples which in the present analysis consist of four separate consonants are listed in (36):

- (36) VCCCCV ∅ VC.CCCV
- a. nskɪ conscript
 - nstɪ instrument
 - lstɪ maelstrom
 - b. kstɪ extra
 - kspl explain
 - bstɪ abstract
 - kskl exclaim

The tableau in (37) for *instrument* shows how the correct syllabification in (36) obtains:

(37)		ALIGN-3 _l	OWF	ONSET	NOCODA	NOCOMPONSET
a.	∅ [ɪn.stɹə.mɪnt]			*	**	*
b.	[ɪn.s.tɹə.mɪnt]	*!		*	**	*
c.	[ɪnst.ɹə.mɪnt]	*!		*	**	
d.	[ɪnstɹ.ə.mɪnt]	*!		**	**	
e.	[ɪ.nstɹə.mɪnt]		*!	*	*	*

In this tableau it is illustrated that the two top ranked constraints ALIGN-3_l and OWF suffice to rule out the final four forms.

5 Syllabification of VC(CC)GV

I treat the syllabification of VC(CC)GV separately because it does not follow from the constraints in the preceding sections alone. It will be shown below that intervocalic C(CC)G clusters can be correctly syllabified give the ranking ALIGN-3_l, OWF » ONSET, NOCODA » NOCOMPONSET established above, with the addition of a constraint barring sonorant consonant plus glide onsets (i.e. *_l[SG) and NOCOMPCODA.

1.1 The syllabification of VCGV

VCGV words have been divided into three categories, as shown in (38):

- (38) a. VCGV sequences in which C is a stop, where CG is a well-formed onset (e.g. *opulent* with [pj])
 b. VCGV sequences in which C is anything but a stop and in which CG is not a well-formed onset (e.g. *onion* with [nj])
 c. VCGV sequences in which C is anything but a stop and in which CG is a well-formed onset (e.g. *amulet* with [mj]).

I discuss first (38b), followed by (38a) and then (38c).¹¹ Context (38c) will be shown to require the constraint *_l[SG referred to above.

¹¹ Note that there is a fourth logical possibility in (38), namely VCGV sequences in which C is a stop, where CG is not a well-formed onset. No clear examples of such CG sequences seem to exist in English. One possible example is [gj], which is attested in the rare word *gules* and as an optional pronunciation for *gubernatorial*. In the following analysis I assume that [gj] is a wellformed onset cluster of English.

1.1.1 VCGV sequences in which C is not a stop (=38b)

The following words contain intervocalic CG sequences in which the C is not a stop and in which CG is not a well-formed onset, i.e. [nj lj ɹj] and [lw] in American English.¹² As indicated in (39) I analyze intervocalic [nj lj] as heterosyllabic – following earlier studies on English phonology (e.g. Borowsky 1986, Jensen 1993) – even though the V.CGCV parsing is not obviously bad from the point of view of syllable markedness.¹³

(39)	VCGV	□	VC.GV
	nj		onion
	lj		million
	ɹj		erudite
	lw		always

In most of the examples discussed below the glide in VC(CC)GV is [j], since there are often few or no examples with [w].

¹² Note that there are no Cj sequences in American English in which the C is a coronal, i.e. [tj dj θj ðj sj zj ʃj ʒj nj lj ɹj].

¹³ Borowsky (1986) and Jensen (1993) posit the Vn.jV and Vl.jV parsing (when the first vowel is stressed) to account for the fact that the [j] in VnjV and Vl.jV deletes in certain dialects when the *second* vowel is stressed, e.g. *volume* (with [j]) vs. *voluminous* (without the [j]) and *continue* (with [j]) vs. *continuity* (without the [j]). Thus, if these dialects ban syllable-initial [lj nj], the deletion of [j] in *voluminous* and *continuity* is a consequence of the parsing V'.ljV and V'.njV. (It should be noted here that many speakers of American English pronounce *voluminous* and *continuity* with the [j]; see also Kenyon and Knott 1953).

In the present analysis it is unclear how to account for the stress-conditioned ‘resyllabification’ of [A.B] to [.AB], as described in the preceding paragraph. Note that it only occurs when A □ [l n] and B is [j], whereas other syllable contacts remain unaffected. For example, heterosyllabic [t.l] is not converted to a [tl] onset when the following vowel is stressed because the /t/ is consistently glottalized, regardless of the stress pattern, cf. the /t/ in *atlas*, *Atlanta*, which is realized in both words as glottalized. I therefore assume for speakers of the dialect in which no [j] is realized in words like *voluminous* and *continuity* the existence of a constraint guaranteeing that V.njV and V.ljV are optimal given the existence of stress on the second vowel. At present I have no explanation for why there should be a constraint nor can I say whether or not it is attested in other languages.

That the sequences [nj] and [lj] in (39) can be preceded by long vowels, e.g. *union*, *amonia*, *begonia*, *failure*, *peculiar*, suggests that the parsing V.CGCV is correct. See Hall (2001), who attributes the occurrence of long vowels before njV and ljV to the fact that [n l j] are all coronal.

The following tableau for the word *onion* shows that the correctly syllabified output form is selected as optimal given the constraints presented in the preceding sections:

(40)

	ALIGN-3	OWF	ONSET	NoCODA	NoCOMPONSET
a. [ʌn.jŋ]			*	*	
b. [ʌ.njŋ]		*!	*		*

The winner in (40a) is more harmonic than (40b) because the latter one violates OWF.¹⁴

1.1.2 VCGV sequences in which C is a stop (=38a)

The examples in (41) contain VCGV sequences in which the C is a stop, and in which CG is a well-formed onset cluster of (American) English. (These clusters subsume [pj bj kj gj kw gw]). In contrast to the examples in (39), I analyze CG as tautosyllabic in such words. There are no sequences of CG in American English in which the C is [t] or [d] (recall note 12).

(41) VCGV V.CG V

pj	opulent
bj	tabular
kj	accurate
gj	regular
kw	equity
gw	jaguar

The reason the words in (41) are parsed V.CG V and not as VC.GV is that the stop is never glottalized if it is voiceless. Further evidence for the tautosyllabic parsing in (41) is that long vowels can occasionally be found before CG sequences when the C is a voiceless or voiced stop. Examples have been provided in (42):¹⁵

¹⁴ In varieties of English in which [nj lj] are well-formed (e.g. RP *news* [njuwz]), the treatment presented up to this point predicts that they should be tautosyllabic in intervocalic position (e.g. [ʌ.njŋ] for *onion*). See §5.1.3. In that section I show that the constraint *_{CG}[SG correctly heterosyllabifies examples like *onion* in British English.

¹⁵ The number of examples like the ones in (42) is small, however. Another peculiarity regarding these words is that the only long vowels permitted before CG are [i:] and [u:]. I leave open whether or not this is systematic or accidental.

The commonly occurring mispronunciation of the word *nuclear* as [nu:kjəlɪ] lends further support to my claim that long (high) vowels can precede CG.

(42) Long vowels before stop-glide sequences:

pj	scrupulous
bj	tubular, zebu
kj	nucule
gj	rugulous
kw	equal, sequence, sequel

It should also be noted here that RP allows long vowels before [tj] and [dj] sequences which are absent in American English, e.g. *mutual* [mju:tjwəl].

In the tableau in (43) for the word *opulent* I have shown that the correctly parsed forms in (41) and (42) are selected by the constraint rankings presented up to this point:

(43)		ALIGN-3	OWF	ONSET	NoCODA	NoCOMPONSET
a.	□ [ɑ.pju.lɪnt]			*	*	*
b.	[ɑp.ju.lɪnt]			*	**!	

In contrast to the previous examples in this section, the winner in (43) can only be selected by considering the low ranked constraint NoCODA.

1.1.3 VCGV sequences in which C is not a stop (=38c)

The following words contain VCGV sequences in which C is anything but a stop and in which CG is not a well-formed onset. These sequences are restricted to [fj vj mj]. As indicated in (44) I analyze CG as heterosyllabic, even though the V.CG.V parsing is not obviously bad from the point of view of syllable markedness.

(44) VCGV □ VC.GV

fj	nephew
vj	purview
mj	amulet

The reason I analyze the CG sequences in (44) as heterosyllabic is that the vowel preceding this cluster is (typically) short, as illustrated in (45) below:¹⁶

¹⁶ To my knowledge no study to date has considered the cooccurrence restrictions regarding the vowels which can and cannot occur before CG in English. Impressionistically it seems that there are few English examples of [vj] clusters which are tautomorphemic with the preceding vowel. There is one example to my knowledge of a word with a [vj] cluster which is preceded by a long vowel, namely *uvular*. Given the word *uvular*, in which the long vowel [u:] surfaces before [vj], one could either conclude that the generalization in

- (45) Short vowels before CG, where C = [f v m]:
 mj stimulous, emulate, amulet, amuse
 fj effusive, refuge, refuse, diffuse
 vj review

Given the generalization concerning the distribution of superheavy syllables in (4) the lack of long vowels before CGV makes sense if CG is heterosyllabic.

The constraints posited up to this point are unable to select the correctly parsed form in (44) and (45). This point is illustrated in the tableau in (46) for the word *amulet*. (‘□’ represents here and below the incorrect winner)

(46)		ALIGN-3□	OWF	ONSET	NoCODA	NoCOMPONSET
a.	[æm.ju.lət]			*	**!	
b.	□ [æ.mju.lət]			*	*	*

In this tableau we see that the intended winner in (46a) and the intended loser in (46b) incur three constraint violations each. (Candidate 46b satisfies OWF because [mj] is a well-formed onset of English, e.g. *music*).

The tableau in (46) shows that an additional constraint is needed which penalizes certain onset clusters which obey OWF, e.g. [mj]. This constraint has been presented in (47). See Casali (1998: 74-88) and Kiparsky (1998) who posit a general constraint *□[CG on the basis of data from the Benue-Congo language Emai (Nigeria) and Gothic respectively. The constraint in (47) only penalizes □[CG onsets in which the C is a sonorant consonant.¹⁷

- (47) *□[SG: Sonorant plus glide in syllable-initial position is prohibited.

Since English allows certain Sj sequences word-initially (e.g. [mj] in *music*), *□[SG is obviously not surface true and must be outranked by the faithfulness constraints (7a).

The reason for the avoidance of SG onsets can be attributed to the observation made by Vennemann (1988) that languages prefer to have onset clusters which are separated by a relatively great sonority value, e.g. [pj] is

(45) is correct and that this is a marked exception, or that long vowels can occur before Cj if the C is an obstruent. These are questions I leave open for further research.

¹⁷ The constraint *□[CG is independently required in languages which allows sequences of obstruent+liquid in the onset but not obstruent+glide.

Jensen (1993: 67) posits an identical constraint for American English, although his reasons are very different than my own.

preferred to [pn] because the sonority distance between [p] and [j] is greater than the distance between [p] and [n] on the Sonority Hierarchy in (9).¹⁸ Vennemann (1988: 14ff.) shows that various synchronic and diachronic patterns are triggered by this tendency, e.g. the historical deletion of /g k/ in initial /kn gn/ clusters in English (e.g. *knee*, *gnome*). I see the avoidance of SG onsets in English in a similar way.

The addition of *_□[SG to the constraints posited above correctly predicts that the optimal form for the words in (44) is selected. This point is illustrated in the tableau in (48) for the word *amulet*.

(48)

	ALIGN-3 _□	OWF	ONSET	* _□ [SG	NoCODA	NoCOMP ONSET
a. □ [æm.ju.lət]			*		**	
b. [æ.mju.lət]			*	*!	*	*

A comparison of tableau (48) with the earlier one in (46) reveals that the addition of the constraint *_□[SG to the hierarchy enables us to select the winner in (48a) over its closest competitor in (48b). The crucial ranking here is that *_□[SG must dominate NoCODA; in (35) below it will be shown that *_□[SG must be ranked below ALIGN-3_□. (Please note that the addition of *_□[SG to the constraint hierarchy for English does not affect the outcome of the tableaux presented earlier).

Note that candidate (48b) also satisfies OWF in RP, because this dialect has many words beginning with [nj], e.g. *news* [njuwz] (recall note 14). Candidate (48a) represents the correct syllabification for this dialect to account for the deletion of /l n/ (recall note 12). This candidate is selected because candidate (48b) violates *_□[SG.

1.2 The syllabification of VCC(C)GV

In this section I divide VCC(C)GV sequences into the following two categories:

- (49) a. VC_aC_bjV in which C_a is anything other than [s]
 b. VC_aC_b(C_c)jV in which C_a is [s]

¹⁸ Vennemann employs ‘strength’, which is the converse of ‘sonority’. According to his Head Law (1988: 13): „A syllable head [=onset, T. A. H.] is more preferred: ...the more sharply the Consonantal Strength drops from the onset toward the Consonantal Strength of the following syllable nucleus.“

I discuss first (49a) and then (49b). The former context will be shown to require reference to the constraint NOCOMP CODA.

1.1.1 VC_aC_bjV in which C_a is anything other than [s] (=49a)

The following words consist of VCCjV sequences. Note that the words are parsed so that only the second C forms an onset with the glide, thus the constraint *_l[SG is consistently violated in (50b) but not in (50a). I have divided the words in (50) into two groups. In (50a) we see VC_aC_bjV sequences in which C_b is a stop (50b) in which C_b is anything other than a stop.

- (50) VCCGV □ VC.CGV
- a. .ɟj argue
 - lkj calculus
 - mkw cumquat
 - b. .mj formula
 - bfj obfuscate

The parsings in (50) derive support from the two generalizations in §2.2. First, the second C in VCCGV is never glottalized if it is a voiceless stop (e.g. the /k/ in *calculus*) and second, the VCC.GV parse would violate (4), since C_aC_b are nonhomorganic.

I have provided a tableau below in (51) for the word *argue* and *formula*, which are representative of (51a) and (51b) respectively. The example in (52) is important because it shows that ALIGN-3□ must dominate *_l[SG.

(51)

	ALIGN-3□	OWF	ONSET	* _l [SG	NoCODA	NoCOMPONSET
a. □ [ɑ.ɟju]			*		*	*
b. [ɑ.ɟ.ju]	*!		*		*	

(52)

	ALIGN-3□	OWF	ONSET	* _l [SG	NoCODA	NoCOMPONSET
a. □ [fo.ɹ.mju.lə]				*	*	*
b. [fo.ɹm.ju.lə]	*!				*	

In these tableaux the (a) candidates are optimal because their closest competitors are not in line with ALIGN-3□.

A final set of data to consider is presented in (53). These words consist of VCCGV, where the adjacent C's are a nasal plus a homorganic stop:

- (53) VCCGV \square VC.CGV
 mpj computer
 mbj ambulance
 ŋgj angular
 ŋgw linguist
 ŋkw banquet

Recall from §2.2 above that tautosyllabic VNC sequences are analyzed as bimoraic and not trimoraic, so that the parsing VC.CjV does not violate ALIGN-3 \square .

The following tableau provides an evaluation of the word *computer*. The winner in (54a) can only be selected by adding the constraint NOCOMP CODA to the constraint ranking presented up to this point and that this constraint must outrank NOCOMP ONSET. For reasons of space I do not list ALIGN-3 \square , OWF, ONSET or * \square [SG, which are satisfied by both candidates.

(54)		NoCODA	NoCOMP CODA	NoCOMP ONSET
a.	\square [kəm.pju.rɪ]	*		*
b.	[kəmp.ju.rɪ]	*	*!	

An inspection of this tableau reveals that the winner in (54a) and its closest competitor in (54b) are in line with ALIGN-3 \square , OWF, ONSET, * \square [SG and they incur one violation each of NOCODA.

It is worth noting that certain examples in (50) and (53) would violate the VCC.GV parse for independent (phonotactic) reasons. For example, in *cumquat* and *obfuscate* in (50) the parsing VCC.GV is bad because the CC strings are otherwise impermissible codas, i.e. *[mk] and *[bf]. The same point holds for a word like *angular* in (53), since there are no English syllables ending in [ŋg].

1.1.2 VC_aC_b(C_b)jV in which C_a is [s] (=49b)

Two final data sets have been presented in (55) and (56). In (55) we see VCCjV and in (56) VCCCjV; in both data sets the first C is an [s]. (There are no examples to my knowledge of words belonging to the latter category in which CCCG are obviously tautomorphic.) As indicated below the CC(C)G sequence is situated in onset position.

- (55) VCCGV \square V.CCGV
 skj rescue
 spj dispute
 skw sesqui-

(56) VCCCGV □ VC.CCGV

kskw exquisite

kskj excuse

Evidence for the parsing in (55) and (56) is that the stop portion is never aspirated (recall §2.2 in which it was noted that only foot-initial /p t k/ aspirate). Since no aspiration occurs even when the vowel following the glide is primarily stressed (e.g. *askew*) we can safely assume that the [s] is in onset position.

In (57) I have provided a tableau for the word *rescue*, which is representative of (55). In (57) and (58) below I only include NOCOMPONSET and not NOCOMP CODA because the winning candidate violates the former constraint and not the latter one.

(57)

	ALIGN-3□	OWF	ONSET	*□[SG	NoCODA	NoCOMP ONSET
a. □ [ɹɛ.skju:]						*
b. [ɹɛs.kju:]					*!	*
c. [ɹɛsk.ju:]	*!				*	

In this tableau (57c) is clearly the least harmonic candidate because it is not in line with ALIGN-3□. The closest competitor to the winner in (57a) is candidate (57b), which is not optimal because it violates NOCODA.

In (58) I have provided a tableau for the word *exquisite*, which is representative of (56):

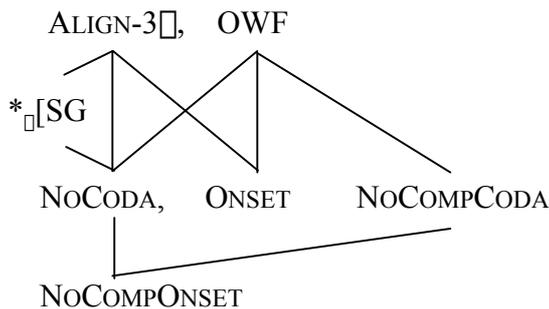
(58)

	ALIGN-3□	OWF	ONSET	*□[SG	NoCODA	NoCOMP ONSET
a. □ [ɛk.skwi.zɪt]			*		**	*
b. [ɛ.kskwi.zɪt]		*!	*		*	*
c. [ɛks.kwi.zɪt]	*!		*		**	*
d. [ɛksk.wi.zɪt]	*!		*		**	

In this tableau (58b-d) are less harmonic than the winner in (58a) because they violate either ALIGN-3□ or OWF.

In (59) I summarize the constraint ranking posited in the preceding analysis:

(59) Summary of rankings:



In §4.2 I suggested that certain consonants in English are ambisyllabic (following the claims made by earlier researchers, e.g. Kahn 1976). If there are indeed ambisyllabic consonants then additional constraints (e.g. CLOSE CONTACT) need to be added to the ranking in (59).

2 A criticism of Hammond (1999)

Hammond (1999) has an OT approach to English which differs radically from the one proposed above because he assumes – contrary to tradition in phonology – that codas and not onsets are maximized.¹⁹ Thus, Hammond proposes that under circumstances to be made explicit below VCV can be parsed VC.V and VCCV as VCC.V. I show here that the generalizations Hammond tries to capture with coda maximization can (and should) be expressed in other ways and therefore reject parsing such as VC.V and VCC.V on principled grounds.

Hammond (1999: 49ff.) argues that the syllabification V.CV is correct unless the C is [ʒ]. In the case of V₃V the consonant is argued to close the first syllable, e.g. [li:ʒ.ɹ] for *leisure* and not [li:ʒ.ɹ], which is the syllabification predicted by the analysis in §2 to be correct. Hammond's argument that V₃.V (and not V.ʒV) is the correct parsing is based on the distribution of [ʒ], which is summarized in (60):

- (60) a. word-final: beige, rouge
 b. intervocalic: leisure, seizure, pleasure
 c. word-initial: * [ʒa]
 d. postconsonantal: * [ɹl.ʒou]

In (60c-d) we see that [ʒ] is barred from onset position and in (60a) that the same sound can occur as a (word-final) coda. If the [ʒ] in the intervocalic environment

¹⁹ Hammond's coda maximization approach is assumed by some earlier authors for English, e.g. Hoard (1971) and Selkirk (1982).

in (60b) were syllable-final – so the argument goes – then the generalization is that English [ʒ] cannot occur in the syllable onset. For this reason, Hammond proposes a constraint he calls *ONSET/_ (p. 49), which bars [ʒ] from onset position. Since this constraint outranks ONSET and NOCODA, the optimal syllabification for words like *leisure* is predicted to be Vʒ.V.

I reject the parsing Vʒ.V because it would violate the generalization in (4) if the first V is long (e.g. in *leisure*). I hold that the phonotactic facts in (60) can be explained without recourse to the syllabification Vʒ.V (or the constraint *ONSET/_). In order to account for the gaps in (60c-d) we simply need a (positive) constraint stating that [ʒ] occurs only in postvocalic position. Given a purely linear constraint, formulated positively and not negatively, there is no reason for the proposed syllabification Vʒ.V.²⁰

According to Hammond the words of the structure in (15) (e.g. *patron*) and (20) (e.g. *bulky*) are syllabified differently than in the present analysis. Some examples of specific clusters which are argued to be parsed VCC.V are listed in (61) (Hammond 1999: 133ff.):

(61) VCC.V parsings according to Hammond (1999):

- Vlt.V: filter
- Vlp.V: pulpit
- Vlk.V: falcon

For reasons of space I do not consider the arguments Hammond adduces for all of the two member clusters which he claims have the VCC.V parsing, nor do I consider his arguments for the VC.CV parsing (e.g. for *cobra*, which I argue is parsed V.CCV). Instead I restrict my discussion to the examples presented in (61) and show how these specific sequences need not have a parsing which is diametrically opposed to the principle of Onset Maximization. I leave open for further study how Hammond's arguments for the VCC.V parsings for other CC clusters or for VC.CV in words like *cobra* can be reanalyzed so that they are in line with the analysis presented in §4.

²⁰ See Green (2001: 37ff.) for an OT analysis of the distribution of English [ʒ] which does not assume Hammond's syllable parsings.

The same reasoning can be applied to Hammond's (1999: 51) analysis of the distribution of the velar nasal. Hammond formulates a constraint (p. 51) stating that [ŋ] does not occur in the onset, concluding that intervocalic [ŋ] is in the coda (e.g. *dinghy* is parsed [dɪŋ.i]). An alternative is that the constraint expressing the distribution of [ŋ] be stated positively: [ŋ] only occurs after a short vowel. Given a constraint along these lines there is no reason to assume the parsing VC.V in words like *dinghy*. See Green (2001: 37ff.) for an OT constraint capturing the distribution of English [ŋ] which refers to moraic structure.

Hammond argues that VCC.V is correct for the sequences in (61) because the cooccurrence restrictions governing vowels that can and cannot occur before the relevant consonant cluster is the same (or nearly the same) when the cluster is word-medial and word-final. Consider the following argument for the syllabification Vlt.V (p. 132): Both word-final [lt] and word-medial [lt] cannot be preceded by the two diphthongs [aʊ oɪ], meaning that there are gaps in English like *[ɔɪlt] and *[ɔɪlti]. Hammond reasons that gaps like these follow if two assumptions are made:

- (62) a. the constraint ruling out [ɔɪlt] and [awlt] is syllable-based
 b. medial sequences of [lt] are affiliated to the left: Vlt.V

I reject both assumptions in (62) below.

The final two clusters in (61) also show a cooccurrence restriction with the previous vowel which Hammond captures with the VCC.V parsing: The long vowels and diphthongs [i: eɪ u: oʊ aɪ aw oʊ] cannot occur before [lp lk], regardless of whether or not [lp lk] are word-final (in 63a) or word-medial (in 63b).

- (63) a. *[i: eɪ u: oʊ aɪ aʊ oɪ] + [lp lk] # (z.B. *[aʊlk])
 b. *[i: eɪ u: oʊ aɪ aʊ oɪ] + [lp lk] V (z.B. *[aʊlki])

Again, Hammond argues that the gaps in (63) can be ruled out if the constraint ruling out [aʊlk] etc. is syllable-based (62a), and if medial sequences of more than one consonant are affiliated to the left, i.e. Vlk.V and Vlp.V (=62b).

Hammond argues that the VCC.C parsings for the words in (61) are selected as optimal due to the interaction between various syllable-markedness constraints and a high ranked constraint he calls MAX-CODA, which say that as many consonants should be affiliated to the left as possible when there is more than one (p. 219). If MAX-CODA is ranked ahead of ONSET, we predict the VCC.V parsing in (61), as illustrated in the tableau in (64) for *fealty*:

- (64) Evaluation of *fealty* after Hammond (1999: 231):

	MAX-CODA	ONSET
[fi:l.ti]	*!	
□ [fi:lt.i]		*

According to Hammond's treatment the parsing V:C.CV loses out over V:CC.V in (64) because the former syllabification violates MAX-CODA and because MAX-CODA is higher ranked than ONSET.²¹

The alternative to Hammond's treatment is to reject both of the assumptions he makes above in (64). Thus, in the case of the [ɔɪlt] and [aʊlt] gaps, one could simply write a highly specific *purely linear* constraint (i.e. one not referring to the syllable) ruling out the sequences [ɔɪlt] and [aʊlt], e.g. *[ɔɪlt]/[aʊlt]. Given the absence of the syllable in this constraint we can therefore parse VltV sequences as Vl.tV and not as Vlt.V.²² A linear constraint like the one just described might derive further support from the fact that the diphthongs [ɔɪ] and [aʊ] have an additional restriction that has nothing to do with syllable structure: neither of them can be followed by a noncoronal consonant, regardless of whether or not that consonant would be tautosyllabic with the diphthong or in the onset of the following syllable (i.e. there are no words like [maʊp], [maʊk], [moɪ.fə] or [moɪ.kə]).²³ Thus, if *[ɔɪ]/[aʊ]+noncoronal is linear, the proposed constraint *[oɪ]/[aʊ]+[lt] could be linear as well.

Hammond tries to account for the gaps involving vowels before the final two clusters in (61) by maximizing codas, but they can be captured by independent constraints which are in line with onset maximization. Thus, consider (63b). Given the (correct) syllabifications Vl.kV and Vl.pV, the lack of long vowels and diphthongs before the lateral can be seen as a consequence of the generalization stated in (4) above, since the superheavy V:l sequence in V:l.kV and V:l.pV is not final in the phonological word. The gaps in (63a) are indicative of a larger gap in English phonotactics, namely V:CC# is only possible if the final C is a coronal obstruent (e.g. *colt*, *scold*, *seems*, *seemed*). Many writers have tried to account for this gap by positing a maximal syllable template for English which can only be exceeded by an (extrasyllabic) coronal

²¹ Hammond analyzes the VC.CV parsing in examples like *cobra* as the result of the ranking of MAX-CODA over NOCODA, i.e. the winner [kɔʊb.rɪə] is better than [kɔʊ.b.rɪə] because the latter incurs a fatal violation of the former constraint. The reason [kɔʊb.rɪə] wins out over [kɔʊb.rɪə] is that the latter form violates a high ranked cover constraint Hammond dubs GOODCODA, which prohibits sequences like [b.rɪ] in syllable-final position.

²² Technically speaking, the constraint *[ɔɪlt]/[aʊlt] is not purely linear, since these sequences can occur if they are separated by a word or compound boundary, e.g. ...owl today. In this case the segments in the constraint *[ɔɪlt]/[aʊlt] must belong to the same phonological word. Thus, a string like [aʊlt] is permissible only when these segments belong to separate phonological words.

²³ The linear constraints described above would also have the advantage that they would not require one to analyze the diphthongs [aʊ] and [oɪ] as trimoraic, as suggested by Hammond (1999).

obstruent. For a recent treatment of these gaps in which no extra-syllabificity is assumed see Hall (2001, 2002). While Hammond's treatment of the gaps involving [lp lk] has the apparent advantage of being able to capture the *[V:lp] and the *[V:lpV] gaps in the same way the cost is that universal Onset Maximization is sacrificed for a language specific constraint maximizing codas. What is equally problematic from a language specific point of view is that Hammond's parsing Vlp.V and Vlk.V cannot account for the fact that /p k/ are never realized as glottalized in this environment.

3 Conclusion

This article has presented a formal OT analysis of syllabification in Modern English in which the surface patterns follow from seven markedness constraints (including an alignment constraint), all of which are independently motivated in other languages, and a language-specific ranking. The syllabification of intervocalic consonant clusters requires the ranking listed in (59).

I have argued above that the proposed analysis is superior to earlier ones (in particular the one proposed by Hammond 1999), in which parsings are assumed which are diametrically opposed to the established principle of Onset Maximization. The reasons for the alternate syllabifications proposed by Hammond (1999) were shown to be compatible with the present analysis.

References

- Borowsky, T. 1986. *Topics in the Lexical Phonology of English*. Ph.D. dissertation. University of Massachusetts at Amherst [Published 1990: New York: Garland].
- Borowsky, T. 1989. Structure preservation and the syllable coda in English. *Natural Language and Linguistic Theory* 7: 145-166.
- Casali, R. F. 1998. *Resolving Hiatus*. New York: Garland.
- Clements, G. N. 1990. The role of the sonority cycle in core syllabification. In: J. Kingston & M. E. Beckman (eds.) *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press. 283-333.
- Davis, S. & M.-H. Cho 2003. The distribution of aspirated stops and /h/ in American English and Korean: an alignment approach with typological implications. *Linguistics* 41.4: 607-652.
- Dogil, G. & H. C. Luschützky 1990. Notes on sonority and segmental strength. *Rivista di Linguistica* 2.2: 3-54
- Giegerich, H. 1992. *English Phonology. An Introduction*. Cambridge: Cambridge University Press.
- Giegerich, H. 1999. *Lexical Strata in English. Morphological Causes, Phonological Effects*. Cambridge: Cambridge University Press.

- Green, A. 1997. *The Prosodic Structure of Irish, Scots Gaelic, and Manx*. Ph.D. dissertation. Cornell University.
- Green, A. 2001. The tense-lax distinction in English vowels and the role of parochial and analogical constraints. *Linguistics in Potsdam* 16: 32-57.
- Green, A. D. 2003. Extrasyllabic consonants and onset well-formedness. In: C. Féry and R. van de Vijver (eds.) *The Syllable in Optimality Theory*. Cambridge: Cambridge University Press.
- Gussenhoven, C. 1986. English plosive allophones and ambisyllabicity. *Gramma* 10: 119-141.
- Hall, T. A. 2001. The distribution of superheavy syllables in Modern English. *Folia Linguistica* XXXV-3/4: 339-442.
- Hall, T. A. 2002. Against extrasyllabic consonants in German and English. *Phonology* 19.1: 33-75.
- Hammond, M. 1999. *The Phonology of English. A Prosodic Optimality-Theoretic Approach*. Oxford: Oxford University Press.
- Harris, J. 1994. *English Sound Structure*. Oxford: Blackwell.
- Hayes, B. 1986. Inalterability in CV phonology. *Language* 62: 321-351.
- Hoard, J. 1971. Aspiration, tenseness, and syllabication in English. *Language* 47: 133-140.
- Hooper, J. 1976. *An Introduction to Natural Generative Phonology*. New York: Academic.
- Jakobson, R. 1962. *Selected Writings I: Phonological Studies*. Den Haag: Mouton.
- Jensen, J. 1993. *English Phonology*. Amsterdam: Benjamins.
- Jensen, J. 2000. Against ambisyllabicity. *Phonology* 17: 187-235.
- Kahn, D. 1976. *Syllable-Based Generalizations in English Phonology*. Ph.D. dissertation. MIT.
- Kenstowicz, M. 1994. *Phonology in Generative Grammar*. Cambridge: Blackwell.
- Kenyon, J. S. & T. A. Knott 1953. *A Pronouncing Dictionary of American English*. Springfield: Merriam.
- Kiparsky, P. 1979. Metrical structure assignment is cyclic. *Linguistic Inquiry* 10: 421-441.
- Kiparsky, P. 1998. Sievers Law as prosodic organization. In: J. Jasanoff, H. C. Melchert & L. Oliver (eds.) *Mir Curad: Studies in Honor of Calvert Watkins*. Innsbruck: Institut für Sprach-wissenschaft der Universität Innsbruck. 345-360.
- Ladefoged, P. 1993. *A Course in Phonetics*. Fort Worth: Harcourt and Brace.
- McCarthy, J. & A. Prince 1993. Generalized alignment. In G. Booij & J. van Marle (eds.) *Yearbook of Morphology 1993*. Dordrecht: Kluwer. 79-153.
- McCarthy, J. & A. Prince 1995. Faithfulness and reduplicative identity. In: J. N. Beckman, L. Walsh Dickey and S. Urbanczyk (eds.) *Papers in Optimality Theory*. Amherst, Mass.: GLSA. 249-384.
- Nespor, M. & I. Vogel 1986. *Prosodic Phonology*. Dordrecht: Foris.
- Prince, A. & P. Smolensky 1993. *Optimality Theory*. Ms. Rutgers University and University of Colorado.
- Raffelsiefen, R. 1999. Phonological constraints on English word formation. In: G. Booij and J. van Marle (eds.) *Yearbook of Morphology 1998*. Kluwer. 225-287.
- Rubach, J. 1996. Shortening and ambisyllabicity in English. *Phonology* 13.2: 197-237.
- Selkirk, E. O. 1982. The syllable. In: H. van der Hulst & N. Smith (eds.) *The Structure of Phonological Representations. Part I*. Dordrecht: Foris. 337-382.

- Sekirk, E. O. 1984. On the major class features and syllable theory. In: M. Aronoff & R. Oehrle (eds.) *Language Sound Structure*. Cambridge, Mass.: MIT Press. 107-137.
- Steriade, D. 2000. Paradigm uniformity and the phonetics-phonology boundary. In: M. Broe & J. Pierrehumbert (eds). *Papers in Laboratory Phonology 5*. Cambridge: Cambridge University Press. 313-334.
- Vennemann, T. 1972. On the theory of syllabic phonology. *Linguistische Berichte* 18: 1-18.
- Vennemann, T. 1988. *Preference Laws for Syllable Structure and the Explanation of Sound Change*. Berlin: Mouton.

Opacity in Tiberian Hebrew: Morphology, not phonology*

Antony D. Green

The phenomenon of phonological opacity has been the subject of much debate in recent years, with scholars opposed to the Optimality Theory (OT) research program arguing that opacity proves OT must be false, while the solutions proposed within OT, such as sympathy theory and stratal OT, have proved to be unsatisfying to many OT proponents, who have found these proposals to be inconsistent with the parallelist approach to phonological processes otherwise characteristic of OT. In this paper I reexamine one of the best known cases of opacity, that found in three processes of Tiberian Hebrew (TH), and argue that these processes only appear to be opaque, because previous analyses have treated them as pure phonology, rather than as an interaction between phonology and morphology. Once it is recognized that certain words of TH are lexically marked to end with a syllabic trochee, and that the goal of paradigm uniformity exerts grammatical pressure on phonology, the three processes no longer present a problem to parallelist OT. The results suggest the possibility that *all* crosslinguistic instances of apparent opacity can be explained in terms of the phonology-morphology interface and that purely phonological opacity does not exist. If this claim is true, then parallelist OT can be defended against its detractors without the need for additional mechanisms like sympathy theory and stratal OT.

1 Introduction

Phonological opacity is the phenomenon of a process applying even though its environment is not present on the surface (overapplication), or of a process failing to apply even though its environment is present on the surface (underapplication). In rule-based frameworks, opacity is frequently accounted for by allowing rules to apply in counterbleeding order (for overapplication) or counterfeeding order (for underapplication) (Kiparsky 1968, Iverson 1995), although not all

* This paper is an early version of a chapter to appear in Green (in prep.). Thanks to Laura Downing, Caroline Féry, and Tracy Hall for helpful comments.

instances of counterbleeding and counterfeeding result in opacity.¹

Spencer (1996: 169) gives an example of counterbleeding with resultant opacity from Bulgarian. In this language, the rules of velar palatalization (1) and yer deletion (2) apply in counterbleeding order (i.e. the second destroys the environment of the first after the first has already applied), resulting in apparent overapplication of velar palatalization: the velar /k/ has palatalized to [č] without the motivation for the change present on the surface, as shown in the derivation of [mračna] ‘dark (fem.)’ in (3).

(1) Velar palatalization

/k/ → [č] / __ [-cons, -back]

/k/ goes to [č] before a front vowel.

(2) Yer deletion

/ĩ/ → ∅ / __ C₀ V_{full}

/ĩ/ (a so-called “yer vowel”) is deleted before a syllable containing a full vowel.

(3) Derivation of [mračna] ‘dark (fem.)’

Underlying representation /mrakĩna/

Velar palatalization č

Yer deletion ∅

Surface representation [mračna]

An example of opacity due to counterfeeding is taken from German (Hall 2000: 142). Here, the rules of dorsal assimilation (4) and r-vocalization (5) apply in counterfeeding order, with apparent underapplication of dorsal assimilation: the front dorsal [ç] appears on the surface after the back vocoid [ɐ], as shown in the derivation of *durch* [dʊɐç] ‘through’ in (6).

(4) Dorsal assimilation

/ç/ → [+back] / [-cons, +back] __

/ç/ goes to [x] after a back vocoid.

(5) R-vocalization

/R/ → [ɐ] / [-cons] __ C₀]_σ

Coda /R/ is vocalized to [ɐ].

¹ Two rules are said to apply in counterbleeding order if the second rule destroys the environment of the first rule (after the first has already applied), and in counterfeeding order if the second rule creates the environment of the first rule (after the first has already failed to apply).

(6)	Derivation of <i>durch</i> [dʊɾç]	‘through’
	UR	/dʊɾç/
	Dorsal assimilation	—
	R-vocalization	ɤ
	SR	[dʊɾç]

The normal application of dorsal assimilation is seen in forms like *Buch* /bu:ç/ → [bu:x] ‘book’ and *Bach* /baç/ → [bax] ‘creek’.

Although derivational phonology thus allowed opaque interactions to be expressed, it was acknowledged that they are less natural than transparent interactions, and it was often argued that rule orderings tend to switch from opaque to transparent over time (Kenstowicz and Kisseberth 1971, 1977, Kiparsky 1971, 1973).

But in a surface-based theory like OT, it is very difficult to explain why faithfulness should be violated excessively when it does not result in improved markedness, which is usually the problem in cases of opacity. Both supporters and opponents of optimality theory point to opacity as a serious flaw in the theory. René Kager, a prominent OT phonologist, says in his textbook on OT (1999: 377):

Opacity appears to be a direct empirical refutation of the *surface-based* evaluation of well-formedness constraints in OT. Since opacity is OT’s Achilles heel, researchers have attempted to find solutions for it which maximally preserve the theory’s advantages.

But in his subsequent discussion he finds that each attempted solution has certain advantages and disadvantages and none of them seems to truly solve the opacity problem.

Perhaps the best known proposal for solving the opacity problem in OT is sympathy theory (McCarthy 1999, 2003), which extends the domain of faithfulness constraints beyond input/output and output/output relations to include relations between competing candidates. As we shall see below, sympathy theory enables opaque candidates to be selected as optimal by proposing constraints enforcing some degree of faithfulness between one candidate and another, so-called “sympathetic,” candidate, which would have been optimal if a specific constraint had been ranked high rather than low.

Another approach to opacity is provided by stratal OT (McCarthy and Prince 1993 (appendix); Kenstowicz 1995; Booij 1996, 1997; Noyer 1997; Paradis 1997; Rubach 1997; Kiparsky 2003), which allows evaluation to proceed in more than one stage, with the possibility of constraint reranking between stages; the input for each stage after the first is the output of the previous stage,

rather than the “underlying” input. This approach also enables the selection of opaque candidates, because faithfulness constraints consider only the relationship of the surface form to an intermediate input rather than the underlying input. I will argue against both sympathy theory and stratal OT below, and show that both approaches are not only too powerful but also unnecessary for an analysis of apparent opacity.

Idsardi (1997, 1998, 2000), an opponent of OT, argues that traditional OT cannot handle the opacity found in the alternations between stops and fricatives in Tiberian and Modern Hebrew at all, and that even sympathy theory, which may be able to handle opacity, predicts that historical language change should decrease opacity (a hypothesis made also by derivational phonologists, as mentioned above), contrary to the facts, since Modern Hebrew has more opaque interactions than Tiberian Hebrew had. He considers opacity to be “*the* single most important issue in current phonological theory” (Idsardi 2000: 337). Other opponents of OT (e.g. Chomsky 1995 and McMahon 2000) concur, arguing that opacity proves that OT is false and that phonology *must* be derivational.

In this paper I will argue that the apparent opaque relationships of Tiberian Hebrew (henceforth TH) can be analyzed in a fully parallel, monostratal version of OT without recourse to mechanisms like sympathy theory and stratal OT. In particular, I contend that there is no phonologically productive opacity in TH; rather, opaque interactions are almost always limited to certain morphological classes or environments.² In TH at least, and perhaps more generally, problems of opacity are never purely phonological in character but are always dependent on morphological/lexical information in some way. If this assertion turns out to be true crosslinguistically, then opacity is a red herring in OT phonology, because truly phonological processes are always transparent. Opacity becomes a question of the interaction of morphology and phonology, rather than a question of the interaction between markedness and faithfulness, which previous accounts have made it out to be.

The paper is organized as follows: in §2 we are introduced to the phonology of TH, and in the following sections we examine in turn three cases of opacity found in TH: the interaction between epenthesis and stress placement in §3, that between epenthesis and ?-deletion in §4, and that between spirantization and syncope in §5. In §6 we see how opacity has been handled up to now in OT,

² See Sanders (2003) for a similar claim about opacity in Polish. I make the qualification “almost” to the statement because some cases of apparent opacity, such as that of German *durch* ‘through’ mentioned above, do seem to be purely phonological. However, even such cases may turn out to have other explanations: in the case of *durch*, my own impression is that the nonsyllabic vocoid corresponding to underlying /R/ is actually rather more front before [ç] than the transcription [ɤ] would indicate. Perhaps a spectrographic analysis will reveal that a transcription like [dʊɛç] or [dʊɪç] would be more accurate.

namely by discarding parallelism either covertly (as in sympathy theory) or overtly (as in stratal OT). In §7 we reexamine the TH data and find that an analysis taking morphological constraints into consideration allows for a fully parallel interpretation of OT: no additional machinery is necessary to account for opacity. §8 sums up with conclusions.

2 Tiberian Hebrew

Tiberian Hebrew (Brown et al. 1906, Gesenius 1910, Hetzron 1987, Malone 1993, Khan 1997, Steiner 1997, Churchyard 1999) is the language in which almost the entire Old Testament is written.³ The term “Tiberian” refers not the region in which the language was spoken, but the region where the scholars (called Masoretes) lived who devised the pointing system that eventually became standard. It is important to be aware that all information about vowels, stress, and spirantization is indicated by this Tiberian pointing system and that Hebrew had died out as a language of everyday communication several centuries before the pointing system was invented. Thus virtually everything that modern linguists assume about the structure of TH depends on information provided by people who were native speakers of Aramaic, not Hebrew.

The surface phone inventory of TH is as shown in (7). The sounds enclosed by boxes are allophones of a single phoneme, traditionally considered to be the stop, but by richness of the base the constraint hierarchy should be organized in such a way that the correct distribution of allophones is predicted regardless of which allophone is assumed in the input. The so-called “emphatic” coronals [t̤ s̤] were probably originally ejective and later uvularized or pharyngealized (Churchyard 1999: 126); the uvular stop [q] also belongs to the class of “emphatic” obstruents. The sound symbolized [š] descends from Proto-Semitic [ʃ] but had already merged with [s] five hundred years before the Masoretic period (Churchyard 1999: 126). If [š] was ever a nonlateral sibilant distinct from [s], it is unknown what the distinction was: McCarthy (1979/85, 13) suggests [š] may have been palatalized; Malone (1993: 28) assumes [š] is [–distributed] (i.e. apical) while [s] is [+distributed] (i.e. laminal). Note that the pharyngeals [ħ ʕ] and laryngeals [h ʔ] are considered glides.

³ A few passages, namely Ezra 4:8–6:18, 7:12–26, Jeremiah 10:11, and Daniel 2:4–7:28, are written in Aramaic.

(7) Surface phones of TH
Consonants⁴

	Labial	Dental	Alveolar	Palatal	Velar	Uvular	Pharyngeal	Laryngeal
Voiceless stops	p	t			k			
Voiceless fricatives	f	θ	s, ś	š	x			
“Emphatic” obstruents		ṭ	ṣ			q		
Voiced stops	b	d			g			
Voiced fricatives	v	ð	z		ɣ			
Nasals	m	n						
Liquids			l, r					
Voiceless glides							ħ	h, ʔ
Voiced glides	w			y			ʕ	

Vowels⁵

	Long	Short	Reduced
i:	u:	i	u
e:	o:	e	o
ɛ:	ɔ:	ɛ	ɔ
		ǣ	ǝ
		a	ǻ

TH has at least three interesting phenomena conventionally described as opaque interactions, which we will discuss in turn in the following sections: there is an epenthesis process that interacts opaquely both with stress placement and with a process of ʔ-deletion, and a spirantization process that interacts opaquely with a syncope process.

⁴ The labial fricatives may have been bilabial [ɸ β] rather than labiodental [f v].

⁵ I follow Hetzron (1987), Malone (1993), Khan (1997), Rendsburg (1997), Steiner (1997), and Bye (2003) in assuming TH had seven different full vowel qualities [i, e, ɛ, a, ɔ, o, u]. This contrasts with the transcription system used by McCarthy (1979, 1999), Idsardi (1997, 1998, 2000), Benua (1998), Churchyard (1999), and Coetzee (1999), which allows only five different vowel qualities [i, e, a, o, u].

3 Epenthesis and stress placement

Among the declension classes of TH is one comprising the forms known as *segolate* nouns.⁶ Segolate nouns, which are all masculine, are characterized by their penultimate stress and by the vowel [ɛ] (but [a] adjacent to pharyngeals and [i] after [y]) in the final syllable. In Proto-Semitic, the ancestors of the TH segolate nouns had the root form CVCC, and most phonologists working on TH assume this shape in the underlying forms of the segolate nouns for synchronic analyses as well. Some examples of segolate nouns are shown in (8).

(8) Segolate nouns (Gesenius 1910: 264–65)

	UR	Surface form	Gloss	Citation
a.	/malk/	mélɛx	‘king’	Gen. 14:7
b.	/sepr/	séfɛr	‘book’	2 Sam. 11:14
c.	/qodš/	qóðɛš	‘sacredness’	Exod. 3:5
d.	/mawt/	mówɛθ	‘death’	Deut. 19:6
e.	/naʕr/	náʕar	‘a youth’	Gen. 37:2
f.	/nešḥ/	néšah	‘perpetuity’	1 Sam. 15:29
g.	/poʕl/	póʕal	‘deed’	Hab. 1:1
h.	/zayt/	záyiθ	‘olive’	Gen. 8:11

Suffixed forms like those in (9) have no epenthetic vowel, indicating that the locus of epenthesis is a syllable- or word-final consonant cluster.

(9) Suffixed forms of segolate nouns

a.	/malk-i:/	malkí:	‘my king’	2 Sam. 19:44
b.	/sepr-i:/	sifrí:	‘my book’	Exod. 32:33
c.	/qodš-i:/	qóðší:	‘my sacredness’	Lev. 20:3
d.	/nešḥ-i:/	nišḥí:	‘my perpetuity’	Lam. 3:18

Malone (1993: 93–94) proposes the rule of segolate epenthesis shown in (10), which inserts the vowel [ɛ] into a word-final consonant cluster.⁷

(10) Segolate epenthesis (Malone 1993: 93–94)

$\emptyset \rightarrow \varepsilon / C _ C \#$

[ɛ] is inserted into a word-final consonant cluster

⁶ From *segol* [səyó:l], the Hebrew name of the vowel point representing [ɛ].

⁷ Malone’s formulation is more complex, taking into account the fact that the vowel surfaces as [a] adjacent to pharyngeals and as [i] after [y]. Malone states his rules using an *SPE*-style formalism, though most phonologists working in 1993 would probably have stated the rule of segolate epenthesis in terms of syllable structure.

In contrast to the segolate nouns, most polysyllabic nouns of TH have final stress (with concomitant vowel lengthening in both the stressed and the pretonic syllable, processes that will not concern us here), as shown in (11) (list adapted from Bye 2003, example (7)).

(11) TH nouns with final stress

	UR	Surface form	Gloss	Citation
a.	/dabar/	dɔ:vó:r	‘word’	Gen. 18:14
b.	/šaba/	ʃɔ:vó:	‘army’	Num. 1:3
c.	/ḥašer/	ḥɔ:šé:r	‘court’	1 Kings 7:8
d.	/zaqen/	zɔ:qé:n	‘old man’	Gen. 43:27
e.	/lebab/	le:vó:v	‘heart’	Deut. 28:28
f.	/šešar/	še:šó:r	‘hair’	Gen. 25:25
g.	/katep/	kɔ:θé:f	‘shoulder’	Zech. 7:11
h.	/rašabo:n/	rəšɔ:vó:n	‘hunger’	Ps. 37:19

Verbs also have a strong tendency to stress the final syllable, with the exception of some unstressed endings or suffixes, as shown in (12).

(12) Stress patterns in verbs

a. with final stress

lɔ:máð	‘he has learned’	Isa. 26:10
yilmáð	‘he learns’	Deut. 17:19
?elməðó:	‘I learn’	Ps. 119:73
yilməðú:n	‘they (m.) learn’	Deut. 4:10
limðú:	‘learn! (m.pl.)’	Isa. 1:17
limmáð	‘he taught’	Eccles. 12:9
yəlammé:ð	‘he teaches’	Ps. 25:9
?əlamməðó:	‘I teach’	Ps. 51:15
yəlamme:ðú:n	‘they teach’	Deut. 4:10
lummó:ð	‘he was taught’	Jer. 31:18

b. with unstressed endings or suffixes

lɔ:máð-ti:	‘I have learned’	Prov. 30:3
limmaðtá-ni:	‘you (m.sg.) have taught me’	Ps. 71:17
təlamməðén-nu:	‘you (m.sg.) teach him’	Ps. 94:12
lamməðé:-ni:	‘teach (m.sg.) me!’	Ps. 25:4

On the basis of forms like those in (11) and (12)a, we could propose a rule building an iamb at the right edge of a prosodic word (cf. Malone 1993: 53–54), shown in (13).

(13) Stress placement

Build an iamb at the right edge of the prosodic word.

This rule interacts opaquely with epenthesis, since epenthetic vowels in final syllables are unstressed, as we saw in the segolate nouns in (8). But segolate nouns are not the only words with epenthesis. Two other classes of words that typically show penultimate stress also have an epenthetic vowel in the final syllable: feminine nouns ending in [-εθ] and “truncated” forms (including jussives and imperfect consecutives) of *lamed-he* verbs (see Benua 1998: ch. 4, and Churchyard 1999: ch. 1 for discussion).⁸ Some examples are shown in (14); related suffixed forms with no epenthesis are shown beneath the forms with epenthesis.

(14) Other forms with epenthesis

a. Feminine nouns in [-εθ]

gulgóleθ	2 Kings 9:35	‘skull’
gulgolt-ó:	Judg. 9:53	
gəvéréεθ	Isa. 47:5	‘mistress’
gəvirt-í:	Gen. 16:8	
šéveθ	1 Kings 10:19	‘dwelling’
šivt-ó:	Obad. 3	
məmšéleθ	Ps. 136:8	‘rule’
məmšalt-ó:	1 Kings 9:19	

b. “Truncated” forms of *lamed-he* verbs⁹

way-yívez	Gen. 25:34	‘despise’
way-yivz-é:hu:	1 Sam. 17:42	
way-yíven	Gen. 2:22	‘build’
yivn-é:hu:	Job 20:19	
yíyel	Job 20:28	‘remove’
yiyl-ú:	Amos 6:7	
way-yéyel	2 Kings 17:6	‘exile’
way-yayl-é:hə:	2 Kings 16:9	

⁸ The vast majority of Hebrew verb roots are considered to consist of three consonants. The *lamed-he* verbs are those whose third consonant (etymologically [y] or [w]) is orthographically *h* and phonologically never present on the surface.

⁹ Forms with [waC-] ‘and’ (where C = copy of the following consonant) are imperfect consecutive; forms without it are jussive.

yíxɛl	Job 33:21	‘be consumed’
yixl-ú:	Isa. 1:28	
wat-témɛr	Ezek. 5:6	‘rebel’
tamr-ú:	1 Sam. 12:14	
way-yáʕal	Gen. 8:20	‘bring up’
yaʕl-é:m	Deut. 28:61	
way-yífen	Exod. 2:12	‘turn’
yifn-é:	1 Sam. 13:17	
way-yéfer	Ps. 105:24	‘make fruitful’
yafr-əxó:	Gen. 28:3	
way-yíʕaʕ	Gen. 4:4	‘look’
yíʕ-é:	Isa. 17:7	
way-yéθaʕ	2 Chron. 33:9	‘cause to stray’
way-yaθʕ-é:m	Job 12:24	
yírev	Gen. 1:22	‘multiply’
yirb-əyú:n	Deut. 8:13	

Stress placement affects the rightmost vowel in a word; segolate epenthesis supplies a vowel that is only one segment removed from the right edge of a word. Thus the two rules stand in a potential feeding relationship: if epenthesis precedes stress placement, the former feeds the latter, but if epenthesis follows stress placement, the former counterfeeds the latter. In fact, since the epenthetic vowel is not stressed, the two rules must stand in counterfeeding order. Also relevant to the present discussion are the rules of midding assimilation in (15) and spirantization in (16).

(15) Midding assimilation (Malone 1993: 64)

$a \rightarrow \varepsilon / _ _ C \varepsilon$

[a] is raised to [ɛ] in an open syllable before an [ɛ] in the following syllable

(16) Spirantization (Idsardi 1998)¹⁰

$[-\text{son}] \rightarrow [+cont] / X \quad _$
 $\quad \quad \quad |$
 $\quad \quad \quad \text{Nucleus}$

A stop becomes a fricative after a vowel.

The derivation of [mélex] ‘king’ is then as shown in (17).

(17) Derivation of [mélex]

Underlying representation	/malk/
	(×)
Stress placement	malk
Segolate epenthesis	lɛk
Midding assimilation	ɛ
Spirantization	x
Surface representation	[mélex]

The fact that stress falls on a nonfinal syllable shows that the ordering of stress placement before segolate epenthesis is opaque. The opacity would remain in a surface-oriented constraint-based analysis, because the facts of (11) and (12)a require a constraint ranking that results in final stress, which would be inconsistent with this form as well as with the other unsuffixed forms in (8) and (14). We will return to the interaction of stress placement and epenthesis in §7.1, where I argue that unsuffixed forms like those in (8) and (14) have penultimate stress for morphological rather than phonological reasons.

4 Epenthesis and ?-deletion

The second case of opacity found in TH involves ?/∅ alternations of the kind shown in (18).

(18) ?/∅ alternations

a.	/śane?/	śo:né:	Deut. 12:31	‘hate’
		śane:ʔ-ś:h	2 Sam. 13:15	

¹⁰ The “emphatic” stops [t] and [q] do not undergo spirantization; Idsardi suggests that spirantization adds not only [+cont] but also [+spread glottis] and that TH prohibits segments from being marked for all three of an oral place feature (coronal or dorsal), a laryngeal feature ([+spread glottis]), and a tongue root feature (on the assumption that the emphatics are marked [+RTR]). Geminates are also excluded from spirantization because of geminate inalterability (Guerssel 1977, Kenstowicz 1982, Hayes 1986, Schein and Steriade 1986, Inkelas and Cho 1993, Elmedlaoui 1993, Kirchner 2000).

b.	/ħetʔ/	ħet ħetʔ-ó:	Lev. 19:17 Lev. 24:15	‘sin’
c.	/gayʔ/	gay ge:ʔ-ó:θ	Num. 21:20 Ezek. 31:12	‘valley’
d.	/yarʔ/	way-yár yirʔ-é:	Gen. 18:2 Gen. 22:8	‘see’

Alternations like those in (18) lead to the conclusion that [ʔ] is deleted in syllable-final position. Idsardi’s (1998) rule of laryngeal deletion is stated in (19).

(19) Laryngeal deletion

$$\begin{array}{l} X]_{\sigma} \\ \dagger \\ \text{ʔ} \end{array}$$

But there is some inconsistency regarding word-final /Cʔ/ clusters. Sometimes the /ʔ/ is simply deleted and the preceding consonant surfaces as the word-final sound, as in (18)b–d. But in other cases, an epenthetic [ɛ]¹¹ appears word-finally after the consonant, as in (20).

(20) Word-final [ɛ] where input ends in /Cʔ/

a.	/gabʔ/	géve gəvɔ:ʔ-ɔ:w	Isa. 30:14 Ezek. 47:11	‘pool’
b.	/dašʔ/	déše dašʔ-ó:	Gen. 1:11 (unattested)	‘grass’
c.	/ʔanʔ/	ʔéne ʔanʔ-ǎxó:	Deut. 26:4 Deut. 28:5	‘basket’
d.	/kelʔ/	kéle kilʔ-ó:	1 Kings 22:27 2 Kings 25:29	‘imprisonment’
e.	/pelʔ/	péle pilʔ-ǎxó:	Exod. 15:11 Ps. 89:6	‘wonder’
f.	/perʔ/	pére pəɔ:ʔ-í:m	Hosea 8:9 Jer. 14:6	‘wild ass’

¹¹ In this and similar forms I follow McCarthy (1999) in transcribing the final vowel as short. In fact, it is not known for certain whether this final unstressed vowel was long or short; Malone (1993) assumes such vowels were long.

g.	/yedʔ/	way-yéðe yiðʔ-é:	Ps. 18:11 Deut. 28:49	‘fly swiftly’
h.	/telʔ/	wat-téle tilʔ-é:	Job 4:5 Job 4:2	‘be weary’
i.	/yerʔ/	yére yirʔ-é:	Exod. 5:21 Gen. 22:8	‘see’

In these forms, there is an opaque interaction of ʔ-deletion with segolate epenthesis. ʔ-deletion counterbleeds segolate epenthesis in these forms by destroying part of the latter’s structural description after it has already applied. On the surface, segolate epenthesis appears to have overapplied, since the epenthetic vowel is not followed by a consonant. The derivation of [péle] is shown in (21).

(21) Derivation of [péle]

UR	/peʔ/
	(×)
Stress placement	peʔ
Segolate epenthesis	leʔ
ʔ-deletion	∅
SR	[péle]

Ordering epenthesis before ʔ-deletion provides a derivational explanation for the opacity in this form, as does the sympathy theory analysis of McCarthy (1999), which brought TH opacity to the attention of phonologists. Bye (2003) proposes an analysis in Declarative Phonology, according to which the epenthetic [e] appears before the position of a ʔ/∅ alternation, but is not regulated by syllable structure.

But none of these analyses is satisfying, since none of them accounts for the fact that not all instances of word-final /Cʔ/ surface opaquely as [Ce]: recall (18)b–d in which /Cʔ/ surfaces transparently as [C]. Malone (1993: 60) simply calls the rule of ʔ-deletion “uneven”, while Coetzee (1999: 76, 178) denies that the forms without epenthesis genuinely have /ʔ/ in their underlying representation. McCarthy (1999) and Bye (2003) fail to mention the transparent forms at all in their analyses. But the issue is addressed by Bruening (1999), who proposes that the transparent and opaque forms belong to different morphological classes with different prosodic requirements, a suggestion I pick up on and elaborate in §7.2.

5 Spirantization and syncope

The third opaque rule interaction found in TH is between spirantization (a singleton nonemphatic stop becomes a fricative after a vowel; see (16)) and a rule I call pretonic syncope, stated in (22).¹²

(22) Pretonic syncope

$$V \rightarrow \emptyset / VC _ CV$$

The weak vowel of an iamb in an open syllable after an open syllable is deleted.

By deleting a vowel, syncope can remove the environment for spirantization. But spirantization applies before syncope does, so syncope counterbleeds spirantization, resulting in an opaque interaction where spirantization appears to have overapplied. In the segolate nouns, this phenomenon is encountered in the construct plural, as shown by the derivation of [malxè:] ‘kings (construct)’ in (23).¹³

(23) Derivation of [malxè:]

UR	/malake:/
	(. ×)
Stress placement	malake:
Spirantization	x
Pretonic syncope	∅
SR	[malxè:]

Other forms also show the opaque interaction between spirantization and syncope, such as [kiθvú:] ‘write (imperative plural)’ < /kotobu:/. In §7.3 I will argue that overapplication of spirantization in TH is attributable to paradigm uniform-

¹² The rule I give as pretonic syncope is simplified from the rules given by Malone (1993) and Idsardi (1998). Malone (1993: 87) proposes a rule called Reduction, which is written in such a way as either (i) to reduce a pretonic vowel in an open syllable to [ə], or (ii) to delete a vowel in that environment altogether if the preceding syllable is light and the intervening consonant is nonguttural. Idsardi (1998) breaks this up into two rules, one of vowel reduction that deletes a foot-initial grid mark before another grid mark, and one of schwa deletion that deletes “a reduced vowel (i.e. one without a grid mark) in an open syllable when the onset is not a guttural and the preceding syllable is also open.” Here I have conflated the two rules and omitted the nonguttural restriction on the first syllable, since sometimes deletion does happen after gutturals, and also because I am not convinced that “nonguttural” is really a natural class.

¹³ The construct is part of the same stress unit (Bruening 1999 suggests the prosodic word) as the noun that follows it, which is why the stress indicated on construct forms is secondary.

ity considerations, but first we turn our attention to how the problem of opacity has been approached by previous OT researchers.

6 Opacity in OT

In derivational phonology rule ordering is sufficient to account for the presence of opaque relationships like the ones described above. In OT phonology, however, over- and underapplication of phonological processes are very difficult to explain. Consider the interaction of epenthesis and ?-deletion discussed above in §4. In an OT analysis, both of these processes involve violation of faithfulness constraints: epenthesis violates DEP-IO(V) and ?-deletion violates MAX-IO(C). But the markedness constraints against complex codas and syllable-final [ʔ] outrank the faithfulness constraints, as is seen in (24) and (25). (Other processes, such as the raising of [a] to [ε] and the spirantization of [k] to [x] in (24) and the lengthening of [a] to [ɔ:] in (25), are not discussed as they are tangential to the issue at hand.)

(24) $*CC]_{\sigma} \gg \text{DEP-IO(V)}$, from [mélɛx] ‘king’

/malk/	$*CC]_{\sigma}$	DEP-IO(V)
malk	* !	
☞ mélɛx		*

(25) $*ʔ]_{\sigma} \gg \text{MAX-IO(C)}$, from [śɔ:né:] ‘he hated’

/śaneʔ/	$*ʔ]_{\sigma}$	MAX-IO(C)
śɔ:néʔ	* !	
☞ śɔ:né:		*

In the correspondence relationship /dašʔ/ \Re [děšɛ], both markedness constraints are obeyed and *both* faithfulness constraints are violated; but the violation of DEP-IO(V) appears gratuitous. The competing candidate *[daš] also obeys both markedness constraints while violating only one of the faithfulness constraints, and is therefore predicted to win the evaluation.¹⁴ (The symbol ☹ indicates the selection of an ungrammatical candidate.)

¹⁴ Note that closed syllables are freely allowed in Hebrew, e.g. [su:s] ‘horse’; thus it is not the fact that *[daš] is a closed syllable that renders it ungrammatical.

(26) Constraint ranking fails on [děšɛ] ‘grass’

/dašʔ/	*CC] _σ	*ʔ] _σ	DEP-IO(V)	MAX-IO(C)
dašʔ	* !	*		
děšɛʔ		* !	*	
děšɛ			* !	*
☹ daš				*

McCarthy’s (1999) solution to this paradox is sympathy theory. According to sympathy theory, the optimal candidate in an evaluation must not only maximize unmarkedness and faithfulness to the input in the conventional manner, but also maximize faithfulness to the so-called “sympathetic candidate” (marked with ☼ in tableaux) chosen by a special constraint called the selector constraint (marked with ☆ in tableaux). The selector constraint may be ranked anywhere in the hierarchy but behaves as if it were top-ranked for purposes of selecting a sympathetic candidate. In the case of /dašʔ/ \mathfrak{R} [děšɛ], the selector constraint is MAX-IO(C) and the sympathetic candidate is [děšɛʔ]. The optimal candidate must obey a faithfulness constraint to the sympathetic candidate, in this case MAX-☼O(V). This constraint compares each candidate not against the input /dašʔ/ but against the sympathetic candidate [děšɛʔ] and gives a violation mark to any candidate lacking a vowel present in the sympathetic candidate. The optimal candidate [děšɛ] obeys MAX-☼O(V) while the transparent candidate [daš] fatally violates it. The tableau for the entire sympathy interaction is shown in (27).

(27) Sympathy analysis of [děšɛ] ‘grass’

/dašʔ/	*CC] _σ	*ʔ] _σ	MAX-☼O(V)	DEP-IO(V)	☆MAX-IO(C)
dašʔ	* !	*	*		
☼ děšɛʔ		* !		*	
☞ děšɛ				*	*
daš			* !		*

The selection of the sympathetic candidate is achieved by imagining that the selector constraint, in this case MAX-IO(C), is top-ranked. If it were, [děšɛʔ] would win, because [děšɛ] and [daš] would fatally violate MAX-IO(C), and [dašʔ] would fatally violate *CC]_σ.

A very similar result is obtained in a different OT approach to opacity, namely stratal OT (McCarthy and Prince 1993 (appendix); Kenstowicz 1995; Booij 1996, 1997; Noyer 1997; Paradis 1997; Rubach 1997; Kiparsky 2003), which follows lexical phonology in assuming different levels of phonological activity. Under stratal OT, the output of one level becomes the input to the next

level. In the TH example, a stratal OT analysis would argue that level 1 has the crucial ranking $\text{MAX-IO(C)} \gg *?]_{\sigma}$, picking [dɛʃɛʔ] as the optimal output to the input /daʃʔ/. At the next level, the input is /dɛʃɛʔ/ and MAX-IO(C) is demoted below $*?]_{\sigma}$; the crucial ranking is now $*?]_{\sigma} \gg \text{MAX-IO(C)}, \text{MAX-IO(V)}$. The output of the level 2 constraint interaction is [dɛʃɛ].

(28) /daʃʔ/ \Re [dɛʃɛ] in stratal OT

a. Level 1

/daʃʔ/	MAX-IO(C)	*CC] _σ	*?] _σ	MAX-IO(V)	DEP-IO(V)
daʃʔ		* !	*		
☞ dɛʃɛʔ			*		*
dɛʃɛ	* !				*
daʃ	* !				

b. Level 2

/dɛʃɛʔ/	*CC] _σ	*?] _σ	MAX-IO(C)	MAX-IO(V)	DEP-IO(V)
dɛʃʔ	* !	*		*	
dɛʃɛʔ		* !			
☞ dɛʃɛ			*		
dɛʃ			*	* !	

The problem with stratal OT is that it eliminates the parallelism of traditional OT. One of the main characteristics that distinguish OT from derivational phonology is that OT establishes a correspondence between an input and an output that does not assume any serial derivation or change over time from the input to the output. Parallel OT does not assume any intermediate stages “between” the input and the output; the input does not come “before” the output in any way. Stratal OT, on the other hand, is derivational: first there is an evaluation at one level, then there is a second evaluation with a new input and a new constraint ranking. Moreover, in a case like this there is no independent evidence, such as the addition of a morpheme, for two separate levels. Traditionally in lexical phonology each level corresponded to some level of morphological affixation. That is not always the case in opacity cases, rendering stratal OT analyses of opacity somewhat ad-hoc.

Sympathy theory is an attempt to sidestep this problem by ostensibly allowing a fully parallel selection of both the sympathetic candidate and the optimal candidate, but it is unclear to what extent this is conceptually really possible. If the selection of the optimal candidate depends on faithfulness to the sympathetic candidate, then the selection of sympathetic candidate must happen in some sense “before” the selection of the optimal candidate. If this is the case, then there is no substantial difference between sympathy theory and stratal OT,

and the tableau in (27) is just a shorthand for the two tableaux in (28).

A further conceptual problem with both sympathy theory and stratal OT is their reliance on faithfulness to a nonexistent form. The output [dɛʃɛ] clearly exists, as this is the form the speaker articulates and the listener perceives; the input exists as well, as this is the form that the speaker (and the listener as well) has listed in the lexicon. But the form [dɛʃɛʔ] cannot be said to exist in the same way: it is a *hypothetical* form that is neither the lexical representation nor the surface form, and yet the grammar is somehow supposed to compare the candidates for faithfulness against both the lexical input /daʃʔ/ and the hypothetical form [dɛʃɛʔ]. Serious questions about learnability are raised here: the learner acquires the output [dɛʃɛ] by hearing it, and the lexical input /daʃʔ/ by induction, comparing it to related forms like [daʃʔó:] ‘his/its grass’, but how does the learner acquire the “sympathetic candidate” or “intermediate input” [dɛʃɛʔ]? For that matter, how does the learner learn to use the relatively low ranked selector constraint (in sympathy theory) or to rerank constraints between levels (in stratal OT)? The most successful models of learning OT grammars, such as the gradual learning algorithm (Boersma 1998, 2000) make no provision for treating a low ranking constraint as if it were high ranking or for reranking constraints in the course of a single instance of harmonic evaluation.

The [dɛʃɛ] paradox relies crucially for its existence on the assumption that the /daʃʔ/ \Re [dɛʃɛ] correspondence is purely phonological in character. If morphological constraints play a role, then the output [dɛʃɛ] could be superior to its competitor *[daʃ] for nonphonological reasons. In the next section I will argue that this is indeed the case.

7 A reexamination of the TH facts

In order to come up with a successful analysis of the TH facts, it will be necessary in this section to reexamine the data in the context of TH morphology.

7.1 Coda clusters in TH

The first fact to be considered is that it is not the case that coda clusters are completely forbidden in TH, as both the derivation epenthesis rule (10) and the constraint ranking $*CC]_{\sigma} \gg \text{DEP-IO(V)}$ would imply. There are an (admittedly very small) number of nouns which end in consonant clusters, such as [qoʃt] ‘truth’ (Prov. 22:21) and [nerd] ‘spikenard’ (Song of Sol. 4:14), which Coetzee (1999: 183) considers to be lexically marked as exceptional. In verbs, however, final clusters are regularly found in jussive/imperfect consecutive forms of *lamed-he* verbs (cf. (14)b) when the cluster is of falling sonority.

- (29) Final clusters in jussive/imperfect consecutive forms of *lamed-he* verbs
- | | | |
|----------|------------|-----------------|
| way-yišb | Num. 21:1 | ‘take captive’ |
| way-yift | Job 31:27 | ‘be simple’ |
| way-yešt | Gen. 9:21 | ‘drink’ |
| way-yevk | Gen. 27:38 | ‘weep’ |
| yešt | Prov. 7:25 | ‘turn aside’ |
| way-yašq | Gen. 29:10 | ‘give water to’ |

Final clusters are the norm in the second person singular feminine perfective form of verbs, where the ending [-t] is attached directly to the final consonant of the root.

- (30) Second person singular feminine perfective verb forms
- | | | |
|---------|-------------|------------|
| hɔ:láxt | Ezek. 16:47 | ‘walk’ |
| yɔ:láðt | Judg. 31:3 | ‘bear’ |
| liqqátt | Ruth 2:19 | ‘glean’ |
| yɔ:náqt | Isa. 60:16 | ‘suck’ |
| niθpást | Jer. 50:24 | ‘be taken’ |

Benua (1998: ch. 4) and Churchyard (1999: ch. 1) attribute the presence of final clusters in (29) and (30) to output-output faithfulness between a base and its truncated form. The forms in (29) are held to be truncated from full forms of the imperfective like those shown in (31).

- (31) Full forms corresponding to (29)
- | | | |
|--------|--------------|-----------------|
| yišbé: | (unattested) | ‘take captive’ |
| yifté: | Deut. 11:16 | ‘be simple’ |
| yišté: | Gen. 44:5 | ‘drink’ |
| tivké: | 1 Sam. 1:7 | ‘weep’ |
| tišté: | Num. 5:29 | ‘turn aside’ |
| yašqé: | Num. 5:26 | ‘give water to’ |

The presence of word-final clusters in (29) can be explained by ranking output-output faithfulness between full forms and truncated forms above *CC]_σ.

As for the forms in (30), these are considered to be truncated from /-ti:/ (on the basis of suffixed forms like [yəliðti:-ni:] ‘you (f. sg.) have borne me’, Jer. 15:10). When not truncated, this ending is homophonous with that of the first person singular, shown in (32).

- (32) First person singular equivalents of the forms in (30)
- | | | |
|-----------|-----------|--------|
| hɔ:láxti: | Ruth 1:21 | ‘walk’ |
| yɔ:láðti: | Gen. 21:7 | ‘bear’ |

liqqátti:	(unattested) ‘glean’
yɔ:náqti:	(unattested) ‘suck’
niθpásti:	(unattested) ‘be taken’

However, the argument that the forms in (29)–(30) stand in a base-truncation correspondence relationship with the forms in (31)–(32) derives solely from the desire to explain the presence of consonant clusters in these forms. There is otherwise no independent evidence that these forms need to be in especially close correspondence with each other, nor that truncation is employed in TH inflectional morphology. Truncation usually has the nature of a hypocoristic (as the nickname *Lar* [lær] truncated from *Larry*, discussed by Benua in chapter 2); it is not clear that it is an advantage to morphological theory to propose that one verb form can be derived from another by truncation. A much simpler explanation would be that the ending of the second person singular feminine perfective is [-t], and its status is entirely parallel to that of other endings like [-ti:] (first person singular perfective) and [-tɔ:] (second person singular masculine perfective). Similarly, the input for a “truncated” jussive/imperfect consecutive form like [yišb] would be simply /yišb/, not /ya-šbɛ:-TRUNC/ as Benua proposes.

If, then, it is the case that the inputs of, say, [liqqátt] and [yišb] are /liqqatt/ and /yišb/, then it is clear that DEP-IO(V) outranks *CC]_σ in TH, i.e. that coda clusters are phonologically tolerated and surface faithfully. Nouns like [qošt] and [nerd] are no longer a problem either. If this is the case, then correspondence relationships like /malk/ \mathfrak{R} [mélɛx] and /mɛmšalt/ \mathfrak{R} [mɛmšéleθ] need to be rethought and a new motivation for the vowel epenthesis found.

Bruening (1999) has argued that morphological classes in TH make reference to prosodic templates, but that these templates are not inputs (as was generally held in pre-OT templatic morphology theory) but are rather output conditions (consistent with the output-based orientation of OT generally). Under this analysis, the difference between segolate nouns and other nouns is that segolate nouns have to meet a trochaic template, while other nouns have to meet an iambic template. Bruening assumes that the trochee of the template is a moraic trochee and thus equally well achieved by (L L) or by (H). Both iambic and trochaic templates are to be aligned at the left edge of the prosodic word; this has the effect of allowing the suffixed forms of segolate nouns to conform to the template, e.g. [(mal)(kí:)] ‘my king’ with a moraic trochee (mal) at the left edge. But Bruening does not address feminine nouns in [-ɛθ] like [gəvéréθ] ‘mistress’ where there is no trochee at the left edge of the word.

If we modify Bruening’s suggestion and assume that the trochaic template of TH is syllabic, not moraic, we can account for the difference between [mélɛx] with epenthesis and [qošt] and [nerd] without epenthesis. [qošt] and [nerd] belong to the first declension, whose template is an iamb (H) or (L H). Iambic are

always moraic, as has been shown by Prince (1990), Kager (1993), Hayes (1995), and Eisner (1997). [mélex], on the other hand, belongs to the second declension, whose template is a syllabic trochee, i.e. a foot of the shape (ó σ).

In fact, it is not necessary to assume a lexically marked iambic template. Rather, since the iambic pattern is the more widespread in TH, we may assume that it is the pattern called for by the purely phonological constraints of the language. Therefore we can say that segolate nouns and feminines in [-εθ] are marked in the lexicon with a diacritic requiring them to end in a trochee, while other nouns have no lexical marking and take the iambic pattern by default. I will indicate the trochaic diacritic as “Tr”.

(33) Nouns with and without the trochaic diacritic

Input	Output	Gloss
/malk/ _{Tr}	mélex	‘king’
/gabert/ _{Tr}	gəvéréεθ	‘mistress’
/qošt/	qošt	‘truth’
/dabar/	də:vó:r	‘word’
/raʕabo:n/	rəʕə:vó:n	‘hunger’

A constraint corresponding to the stress placement rule of (13) is STRESSRIGHT (34).

(34) STRESSRIGHT

Align-R(PWord, ó)

The right edge of a prosodic word is aligned with the right edge of a stressed syllable.

This constraint achieves the final stress seen in the forms in (11) and (12)a. But for words marked with the trochaic diacritic, the constraint TROCHEE takes precedence.

(35) TROCHEE

Align-R(PWord_{Tr}, Trochee)

The right edge of a prosodic word associated with the diacritic Tr is aligned with the right edge of a trochee.

TROCHEE outranks STRESSRIGHT (a case of Pāṇinian ordering since TROCHEE is more specific than STRESSRIGHT), as shown in the tableaux in (36). This results in epenthesis in [mélex] and [gəvéréεθ], both with the trochaic template, but not in [qošt] with no lexically marked template.¹⁵

¹⁵ Bruening himself actually assumes phonological epenthesis (i.e. *CC]_σ >> DEP-IO(V)), but then he does not address forms like [qošt].

(36) Nouns with and without lexically specified templates

a. Trochaic diacritic: [mélex] ‘king’

/malk/ _{Tr}	TROCHEE	STRESSRIGHT	DEP-IO(V)	*CC] _σ
(malk)	* !			*
☞ (mélex)		*	*	

b. Trochaic diacritic: [gəvéreθ] ‘mistress’

/gabert/ _{Tr}	TROCHEE	STRESSRIGHT	DEP-IO(V)	*CC] _σ
ga(vért)	* !			*
☞ gə(véreθ)		*	*	

c. No diacritic: [qošt] ‘truth’

/qošt/	TROCHEE	STRESSRIGHT	DEP-IO(V)	*CC] _σ
☞ (qošt)				*
(qóšet)		* !		
(qəšé:t)			* !	

Epenthesis in segolate and [-eθ] nouns is thus driven not by the desire to avoid final clusters, but rather by the lexically imposed requirement to have a (σ σ) trochee at the end of the word.

Verbs almost always follow the iambic pattern (abstracting away from unstressed suffixes and endings like those seen in (12)b), so that [liqqátt] and [yišb] surface without epenthesis.

(37) Verb forms not marked with Tr diacritic

a. [liqqátt] ‘glean’ (2 sg. fem. perf.)

/liqqatt/	STRESSRIGHT	DEP-IO(V)	*CC] _σ
☞ (liq)(qátt)			*
(liq)(qátεθ)	* !		
(liq)(qətḗ:θ)		* !	

b. [yišb] ‘take captive’ (jussive)

/yišb/	STRESSRIGHT	DEP-IO(V)	*CC] _σ
☞ (yišb)			*
(yíšev)	* !		
(yəšḗ:v)		* !	

Under this analysis, there is no opacity between stress placement and epenthesis in TH after all. It is not the case that epenthetic vowels are unstressed because they are added after stress has already been assigned. Rather, epenthetic vowels are unstressed because they are added only into words that are required to end in

a syllabic trochee and thus in an unstressed syllable.

Compliance with the lexical diacritic may perhaps be overridden by phonotactic considerations. Consider the jussive/imperfect consecutive forms of various *lamed-he* verbs: the forms in (38)a–f are repeated from (29) and the forms in (38)g–r from (14)b.

(38) Jussive/imperfect consecutive forms of *lamed-he* verbs

a.	way-yíšb	Num. 21:1	‘take captive’
b.	way-yíft	Job 31:27	‘be simple’
c.	way-yéšt	Gen. 9:21	‘drink’
d.	way-yév̄k	Gen. 27:38	‘weep’
e.	yešt̄	Prov. 7:25	‘turn aside’
f.	way-yášq	Gen. 29:10	‘give water to’
g.	way-yívez	Gen. 25:34	‘despise’
h.	way-yíven	Gen. 2:22	‘build’
i.	yíxel	Job 20:28	‘remove’
j.	way-yéxel	2 Kings 17:6	‘exile’
k.	yíxel	Job 33:21	‘be consumed’
l.	wat-témer	Ezek. 5:6	‘rebel’
m.	way-yáfal	Gen. 8:20	‘bring up’
n.	way-yífen	Exod. 2:12	‘turn’
o.	way-yéfer	Ps. 105:24	‘make fruitful’
p.	way-yíšaʕ	Gen. 4:4	‘look’
q.	way-yéθaʕ	2 Chron. 33:9	‘cause to stray’
r.	yírev	Gen. 1:22	‘multiply’

The forms in (38)a–f all end in a cluster with falling sonority; furthermore they all have final stress and thus conform to basic TH stress pattern. The forms in (38)g–q would all end in a cluster with level or rising sonority if the epenthetic vowel were not present; they all have penultimate stress and thus conform to the trochaic template, but because level and rising sonority clusters are prohibited in word-final position in TH it is unclear whether these words are trochaic for phonotactic reasons or because they are conforming to a lexically specified trochaic diacritic. But the form [yírev] in (38)r is clearly lexically marked for the trochaic diacritic since the alternative without the epenthetic vowel, *[yirb], is phonotactically permitted as it ends in a cluster with falling sonority. Thus we may conclude that some jussive/imperfect consecutive forms of *lamed-he* verbs are marked for the trochaic diacritic, while others are not so marked, but for many it is ambiguous to which they belong.

7.2 Epenthesis and ?-deletion

Turning to the next supposed example of opacity in TH, that between epenthesis and ?-deletion, we find that the derivational analysis, under which epenthesis precedes (and is counterbled by) ?-deletion, predicts that all cases of underlying word-final /C?/ surface as [Cε], as we saw above in (20). However, this is not the case. There are nouns like [hət̚] ‘sin’ ((18)b) and [gay] ‘valley’ ((18)c) and verb forms like [way-yár] ‘see (imperfect consecutive)’ ((18)d) where [?] is deleted without epenthesis before it. Recall from (26) that /C?/ \mathfrak{R} [C] is the predicted, transparent relation. In fact, the only place where the correspondence relationship /C?/ \mathfrak{R} [Cε] holds is in segolate nouns (e.g. [dɛʃε] ‘grass’) and in the “truncated” forms (jussives and imperfect consecutives) of *lamed-he* verbs whose second consonant is [?], e.g. [way-yéd̚ε] ‘fly swiftly (imperfect consecutive)’ (root *d-ʔ-h*). As we saw in the previous section, segolate nouns and “truncated” forms of some *lamed-he* verbs are lexically marked with the diacritic Tr, requiring them to have a syllabic trochee (σ σ) at their right edge. The forms without final [ε] do not have this diacritic. The purpose of the epenthesis is thus not to break up the /C?/ cluster but to provide an unstressed final syllable so that the (σ σ) template can be met, as Bruening (1999) argued. The constraint hierarchy is shown in the tableaux in (39), on the basis of first declension (no diacritic) [hət̚] ‘sin’ and second declension (trochaic diacritic) [dɛʃε] ‘grass’.

(39) Treatment of final C? clusters

a. First declension (no diacritic) [hət̚] ‘sin’

/hət̚?/	*?] _σ	TROCHEE	STRESSRIGHT	MAX-IO(C)	DEP-IO(V)	*CC] _σ
(hət̚?)	* !					*
☞ (hət̚)				*		
(hɛ̌t̚é:)				*	* !	
(hét̚ε)			* !	*	*	

b. Second declension (trochaic diacritic) [dɛʃε] ‘grass’

/daʃ?/Tr	*?] _σ	TROCHEE	STRESSRIGHT	MAX-IO(C)	DEP-IO(V)	*CC] _σ
(daʃ?)	* !	*				*
(daʃ)		* !		*		
☞ (dɛʃε)			*	*	*	

The same tableaux would hold for /waC-yar?/ \mathfrak{R} [way-yár] ‘see (imperfect consecutive, no diacritic)’ and /waC-yed?/Tr \mathfrak{R} [way-yéd̚ε] ‘fly swiftly (imperfect consecutive, trochaic diacritic)’.

7.3 Spirantization after consonants

The third opacity of TH is that between syncope and spirantization, as illustrated by the relationship /malake:/ \Re [malxè:] ‘kings (construct)’ (Gen. 17:16). The normal application of spirantization can be attributed to the constraint interaction $*V^{\wedge}[-cont] \gg *NONSIBILANT\ FRICATIVE \gg IDENT-IO(cont)$. The constraint $*V^{\wedge}[-cont]$ prohibits postvocalic stops, the constraint $*NONSIBILANT\ FRICATIVE$ ($*NONSIBFRIC$) prohibits the feature cooccurrence $[-son, +cont, -sib]$ (penalizing [f θ x v ð ɣ] while permitting [s š z ʃ ʒ]¹⁶), and $IDENT-IO(cont)$ prohibits a mismatch in the feature [continuant] between the input and the output.¹⁷ In §2 it was mentioned that the principle of richness of the base predicts that the correct distribution of stops and fricatives will be achieved regardless of whether it is stops or fricatives that are present in the input. To show this, in the tableaux in this section, two inputs will be provided for each form, input (α) with stops and input (β) with fricatives. The constraint interaction responsible for the distribution of stops and fricatives is illustrated in (40) by [kɔ:θáv] ‘he wrote’ (Josh. 8:32). (The relationship between input /a/ and output [ɔ:] in the first syllable is not analyzed here.)

(40) $*V^{\wedge}[-cont] \gg *NONSIBFRIC \gg IDENT-IO(cont)$, from [kɔ:θáv]

(α) /katab/	$*V^{\wedge}[-cont]$	$*NONSIBFRIC$	$IDENT-IO(cont)$
(β) /xaθav/			
kɔ:táb	*!*		(α)
			(β) **
 kɔ:θáv		**	(α) **
			(β) *
xɔ:θáv		***!	(α) ***
			(β)

But in the opaque relationship /malake:/ \Re [malxè:], the fricative [x] surfaces even though a consonant precedes on the surface. This looks like a gratuitous violation of $*NONSIBFRIC$; the transparent candidate would be *[malkè:]. The constraint LAPSE prohibits a sequence of two unstressed light syllables in a row.

¹⁶ Cf. Benua’s (1998) constraint $*SPIR$.

¹⁷ The nonspirantization of [t̪ q] may be attributed to undominated markedness constraints against the “emphatic” nonsibilant fricatives [θ] and [χ].

(41) Transparent candidate *[malkè:] wins instead of [malxè:] ‘kings (construct)’

(α) /malake:/ (β) /malaxe:/	LAPSE	*V̄[-cont]	*NONSIBFRIC	MAX-IO(V)	IDENT-IO(cont)
mələkè:	* !	*			(α) (β) *
mələxè:	* !		*		(α) * (β)
⊗ malkè:				*	(α) (β) *
malxè:			* !	*	(α) * (β)

This overapplication of spirantization is found in only a few morphological contexts in TH.¹⁸ First, it is found in imperative verb forms in the second person singular feminine, the second person plural masculine, and the second person singular masculine when this is augmented by the emphatic suffix [-ɔ:]. Examples are shown in (42), where overapplication of spirantization is indicated by italics.¹⁹

(42) Overapplication of spirantization in imperatives

a. Second person feminine singular

šixví:	‘lie down!’	2 Sam. 13:11
šifxí:	‘pour out!’	Lam. 2:19
məlxí:	‘rule!’	Judg. 9:10

b. Second person masculine plural

kiθnú:	‘write!’	Deut. 31:19
wə-šivðú:	‘and serve!’	Exod. 10:11
məšxú:	‘draw!’	Ezek. 32:20
ħorvú:	‘be wasted!’	Jer. 2:12

c. Second person masculine singular + emphatic suffix [-ɔ:]

šɔzv-ó:	‘leave!’	Jer. 49:11
šixv-ó:	‘lie down!’	Gen. 39:7
šerx-ó:	‘set in array!’	Job 33:5

In nouns, overapplication of spirantization is found in the plural construct, in-

¹⁸ It is also found in a few foreign (Persian) names, such as [ħarvo:nó:], [biγθó:], [ʔävayθó:] in Esther 1:10.

¹⁹ There are occasional exceptions such as [ʔispi:] ‘gather! (2 f. sg.)’ (Jer. 10:17) beside expected [ʔisfú:] ‘gather! (2 m. pl.)’ (Ps. 50:5).

cluding forms with a possessive suffix. Examples are shown in (43).²⁰

(43) Overapplication of spirantization in plural construct forms

- a. Unsuffixed forms
- | | | |
|---------|-------------|------------------|
| biyðè: | ‘garments’ | Gen. 27:15 |
| zanvò:θ | ‘tails’ | Isa. 7:4 |
| ḥasðè: | ‘mercies’ | Isa. 55:3 |
| ḥorvò:θ | ‘ruins’ | Isa. 5:17 |
| yiqvè: | ‘vats’ | Zech. 14:10 |
| yiθðò:θ | ‘pins’ | Exod. 27:19 |
| kanfè: | ‘wings’ | Exod. 19:4 |
| kanfò:θ | ‘wings’ | Deut. 22:12 |
| kiθfò:θ | ‘shoulders’ | Exod. 28:12 |
| fiqvè: | ‘heels’ | Song of Sol. 1:8 |
- b. Forms with possessive suffix
- | | | |
|---------------|--------------------------|--------------|
| biyðo:θ-é:xó: | ‘your (m. sg.) garments’ | Ps. 45:19 |
| ḥorvo:θ-áyix | ‘your (f. sg.) ruins’ | Isa. 49:19 |
| kanfe:-hém | ‘their (m.) wings’ | 1 Kings 6:27 |
| kanfe:-hén | ‘their (f.) wings’ | Ezek. 17:3 |

The third context where overapplication of spirantization is consistently found is with the accusative/possessive suffixes [-xó:] ‘you/your (m. sg.)’, [-xém] ‘you/your (m. pl.)’ and [-xén] ‘you/your (f. pl.)’, which always begin with the fricative [x], regardless of whether a vowel or a consonant precedes. Examples of these suffixes occurring after a consonant are shown in (44).

(44) Accusative/possessive suffixes beginning with [x] even after a consonant

- a. 2nd person masculine singular
- | | | |
|--------------|-------------------------|------------------|
| ʔǎxól-xó: | ‘your eating’ | Gen. 2:17 |
| bin-xó: | ‘your son’ | Gen. 22:2 |
| u-ve:rax-xó: | ‘and he will bless you’ | Deut. 7:13 |
| go:ʔal-xó: | ‘your Redeemer’ | Isa. 48:17 |
| ḥibbəláθ-xó: | ‘she bore you’ | Song of Sol. 8:5 |

²⁰ Here too there are occasional exceptions such as [ʔospé:] ‘gatherings’ (Mic. 7:1), [ḥerpó:θ] ‘reproaches’ (Ps. 69:10), [ʔarpé:] ‘preys’ (Ezek. 17:9), [kaspe:-hém] ‘their (m.) silver coins’ (Gen. 42:25), [niske:-xém] ‘your (m. pl.) drink offerings’ (Num. 29:39), [šimdé:] ‘pairs’ (Isa. 5:10), [rišpé:] ‘flames’ (Song of Sol. 8:6, beside expected [rišfè:] at Ps. 76:4). Since spirantization (or lack of it) is indicated by the vowel pointing, which was not devised until several centuries after Hebrew had stopped being used as a spoken language, it is impossible to know to what extent exceptions such as these were genuinely present in the living language.

	wə-ʔeʔeś-xó:	‘and I make you become’	Gen. 12:2
	šim-xó:	‘your name’	Gen. 17:5
b.	2nd person masculine plural		
	ʔaðmaθ-xém	‘your ground’	Gen. 47:23
	ʔāxól-xém	‘your eating’	Gen. 3:5
	ʔeθ-xém	‘you (acc.)’	Gen. 47:23
	bi-vśar-xém	‘in your flesh’	Gen. 17:13
	dim-xém	‘your blood’	Gen. 9:5
	bə-yεð-xém	‘into your hand’	Gen. 9:2
	mišmar-xém	‘your ward’	Gen. 42:19
	ʔorlaθ-xém	‘your foreskin’	Gen. 17:11
c.	2nd person feminine plural		
	lə-qaðmaθ-xén	‘to your former state’	Ezek. 16:55

Finally, overapplication of spirantization is found when one of the proclitics [bə] ‘in’, [wə] ‘and’, [kə] ‘like, according to’, [lə] ‘to’ is attached to a word beginning [C₁əC₂-]. The result is [biC₁C₂-], [uC₁C₂-], [kiC₁C₂-], [liC₁C₂-] with spirantization of both C₁ (normal application) and C₂ (overapplication), if possible.

(45) Overapplication of spirantization after proclitics

a.	šəʔo:ʔó:	‘error’	Num. 15:25
	bi-šʔo:ʔó:	‘in error’	Lev. 4:2
b.	gəvó:l	‘a border’	Gen. 10:19
	u-ʔvó:l	‘and a border’	Num. 34:6
c.	ləvɔ:v-ó:	‘his heart’	Deut. 2:30
	ki-lvɔ:v-ó:	‘according to his heart’	1 Sam. 13:14
d.	bəθu:làθ	‘a virgin (construct)’	Deut. 22:19
	li-vθu:làθ	‘to a virgin (construct)’	Lam. 1:15

All of these cases of overapplication of spirantization can be attributed to paradigm uniformity/uniformity of exponence effects (Kenstowicz 1996, 1997, Buckley 1999, Steriade 2000, Downing et al. forthcoming).

Consider first the cases in (42). The imperative forms listed there are the only ones in the entire verbal paradigm where the third consonant of the verbal root follows a consonant. In all other forms, the third consonant follows a vowel and thus undergoes spirantization regularly. The spirantization in the imperative forms can thus be considered analogical, occurring in order to reduce variation within the paradigm. The first and second consonants of the root do alternate between stops and fricatives (provided they belong to the class of consonants allowing this alternation), but the third consonant is invariably spirantized. Tak-

ing as an example the form [kiθnú:] ‘write!’ (2 m. pl.) from (42)b, the paradigm of the qal binyan of the root *k-t-b* ‘write’ is given in (46); note that [k] and [θ] alternate with [x] and [t] in the imperfective, but [v] is invariant across all forms.

(46) Qal binyan of *k-t-b* ‘write’

	Perfective	Imperative	Imperfective
1 sg.	kɔ:θávti:		ʔextó:v
2 m. sg.	kɔ:θávto:	kəθónv/kiθvó:	tixtó:v
2 f. sg.	kɔ:θavt	kiθví:	tixtəví:
3 m. sg.	kɔ:θáv		yixtó:v
3 f. sg.	kɔ:θəvó:		tixtó:v
1 pl.	kɔ:θávnu:		nixtó:v
2 m. pl.	kəθavtém	kiθnú:	tixtənú:
2 f. pl.	kəθavtén	kəθóvno:	tixtóvno:
3 m. pl.	kɔ:θəvú:		yixtənú:
3 f. pl.	"		tixtóvno:
Infinitive construct		kəθón	
Infinitive absolute		kɔ:θó:v	
Active participle		ko:θé:v	
Passive participle		kɔ:θú:v	

There are many proposals for analyzing paradigm uniformity effects in OT (see references mentioned above), and this is not the place to discuss the merits and liabilities of each. Suffice it to say the reason the italicized sounds in (46) are fricatives rather than stops is not because of an underlying vowel before them which causes spirantization before disappearing, but rather because of the desire to reduce intraparadigm variation and allow the third consonant of the root to surface in a uniform manner. The stop/fricative alternations of the first and second consonants are widespread and systematic within the paradigm and therefore tolerated.

The fact that nouns in the construct plural show spirantization of the final consonant after another consonant, as shown in (43), can also be attributed to paradigm uniformity. Consider the paradigms of an iambic-template noun such as [zɔ:nó:v] ‘tail’ and a trochaic-template (i.e. segolate) noun such as [mélex] ‘king’. As shown in (47), the iambic-template noun has a vowel before the root-final consonant in all forms of the singular as well as in the plural absolute, so that a fricative is phonologically predicted here. It is only in the construct plural that the root-final consonant follows another consonant; phonologically a stop would be expected, but in fact a fricative occurs.

(47) Paradigm of [zɔ:nó:v] ‘tail’ (first declension)

	Singular	Plural
Absolute	zɔ:nó:v	zəno:vó:θ
Construct	zənav	zanvò:θ
Construct suffixed	zəno:v-ó:	zanvo:θ-e:hém

In segolate nouns, there is a stop/fricative alternation in the singular, but none in the plural, as shown in (48).

(48) Paradigm of [mélex] ‘king’ (second declension, i.e. segolate)

	Singular	Plural
Absolute	mélex	məlo:xí:m
Construct	mélex	malxè:
Construct suffixed	malk-ó:	malxe:-hém

The fact that the segolate paradigm contains both stop and fricative allophones led Benua (1998) to argue that overapplication of spirantization in plural construct forms cannot be attributed to output-output faithfulness.²¹ I disagree and would argue that in both (47) and (48), the spirantization of the root-final consonant in the construct plural may be analyzed as analogical to the absolute plural, where the spirantization is phonological. In the case of first-declension nouns like [zɔ:nó:v], the singular may be exerting analogical influence as well. Segolate (second-declension) nouns generally tolerate the stop/fricative alternation in the singular, but there are a few nouns like [béyεð] ‘garment’ that consistently retain the fricative even in singular suffixed forms: [biyð-í:] ‘my garment’ (Ezra 9:3), [biyð-ó:] ‘his garment’ (Gen 39:12); another example is [yiqv-é:xɔ:] ‘your (m. sg.) wine vat’ (Deut. 15:14). These cases show that the paradigmatic pressure for a consistently spirantized root-final consonant has begun to spread to the segolate construct singular as well.

As for the suffixes in (44), they may be analyzed as obeying a high-ranking uniformity of exponence constraint requiring them to surface with [x]. Another option would be to posit /x/ rather than /k/ in the input, but this analysis comes into conflict with richness of the base. In an alternation like [zɔ:xár] ‘he remembered’ (Gen. 40:23)/[yizkó:r] ‘he will remember’ (Hosea 8:13), richness of the base predicts that in allophonic variation, the correct allophones will surface regardless of which allophone is present in the input, i.e. even the input /zakar/ should surface as [zɔ:xár], and even the input /yizxo:r/ should surface as [yizkó:r]. So if the constraint ranking is such that the optimal output of /yizxo:r/ is [yizkó:r], then the optimal output of /binxo:/ ‘your son’ should likewise be *[binkó:], not [binxó:]. Therefore it is preferable to assume a constraint unique

²¹ Instead, she outlines a possible sympathy theory analysis.

to these three possessive suffixes requiring them to surface with [x], rather than attempting a purely phonological analysis of the facts of (44).

Finally, the data in (45) are also best explained as the result of uniform exponence. The forms [šəʔɑ:ʔó:], [gəvó:l], [ləvɑ:vó:], and [bəθu:làθ] have a phonologically expected spirant as their second consonant; this spirant remains in the forms [bi-šəʔɑ:ʔó:], [u-ʔvó:l], [ki-lvɑ:vó:], [li-vθu:làθ] with a proclitic. If a stop were to occur in the forms with the proclitics, the relationship between the basic form and the extended form would be made less transparent, damaging recoverability. The one exception to the generalization that words beginning [C₁əC₂-] retain the spirantization of C₂ when a proclitic precedes is when the proclitic [lə] attaches to an infinitive like [kəθó:v] ‘write’. In this case, the result is not *[lixθó:v] but rather [lixtó:v] ‘to write’ (Josh. 18:8) with a stop. With other proclitics, infinitives behave normally, e.g. [bixθó:v] ‘in writing’ (Ps. 87:6). Idsardi (1998), arguing against an output-output correspondence analysis, accounts for this difference by analyzing [lə] as being [+cyclic] and [bə] as [-cyclic] and arguing that [+cyclic] forms behave in such a way as to delete vowels by syncope before spirantization has a chance to apply, whereas [-cyclic] forms trigger syncope only after spirantization has already applied. But forms like [li-vθu:làθ] beside [bi-šəʔɑ:ʔó:] prove that the difference is not between [lə] and [bə], but rather between [lə] + noun and [lə] + infinitive. Gesenius (1910: 123–24, 348–51) has argued that [lə] + infinitive is a distinct grammatical form (he calls it “a kind of gerund”); if this is so, then a gerund like [lixtó:v] is sufficiently removed from the infinitive [kəθó:v] that the [t/θ] alternation is accepted between them, whereas [bi-xθó:v], being simply a proclitic + infinitive and not a distinct grammatical form, requires a greater degree of identity to [kəθó:v].

8 Conclusions

In this paper I have argued that the various alleged opaque relationships of TH do not refute OT, nor do they require any additional mechanism like sympathy theory. My analysis supports the hypothesis of Sanders (2003) that there is no instance of opacity in a purely phonological relationship (i.e. one that is free of morphological influence). All of the apparently opaque relationships in TH are influenced heavily by the morphology, in particular by the presence of trochaic templates for some words, including the declension class of segolate nouns. It is conformity to the trochaic template that causes the nonfinal stress discussed in §3 as well as the apparently unmotivated final epenthetic vowel discussed in §4. The overapplication of spirantization discussed in §5 is due to paradigm uniformity or uniformity of exponence considerations.

9 References

- Benua, L. 1998. Transderivational identity: Phonological relations between words. PhD diss., UMass. Pub. as *Phonological relations between words*, New York: Garland, 2000.
- Boersma, P. 1998. *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.
- Boersma, P. 2000. Learning a grammar in functional phonology. In Dekkers, J., F. van der Leeuw, and J. van de Weijer, eds. 2000. *Optimality theory: Phonology, syntax, and acquisition*. Oxford: Oxford Univ. Press., 465–523.
- Booij, G. 1996. Lexical phonology and the derivational residue. In *Current trends in phonology: Models and methods*, ed. J. Durand and B. Laks, 69–96. Salford: European Studies Research Institute, Univ. of Salford.
- Booij, G. 1997. Non-derivational phonology meets lexical phonology. In Roca 1997: 261–88.
- Brown, F., S. R. Driver, and C. A. Briggs. 1906. *The new Brown-Driver-Briggs-Gesenius Hebrew and English lexicon*. Boston: Houghton, Mifflin. Repr. Peabody: Hendrickson, 1979.
- Bruening, B. 1999. Morphological templates and phonological theory. Talk presented at the second Mediterranean Meeting on Morphology, Msida, Malta.
- Buckley, E. 1999. Uniformity in extended paradigms. In *The derivational residue in phonological optimality theory*, ed. B. Hermans and M. van Oostendorp, 81–104. Amsterdam: Benjamins.
- Bye, P. 2003. Opacity, transparency and unification in the phonology of Tiberian Hebrew. In *Proceedings of the North East Linguistic Society 33*, ed. M. Kadowaki & S. Kawahara. Amherst: GLSA, University of Massachusetts.
- Chomsky, N. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.
- Churchyard, H. 1999. Topics in Tiberian Biblical Hebrew metrical phonology and phonetics. PhD diss., Univ. of Texas.
- Coetsee, A. W. 1999. *Tiberian Hebrew phonology: Focussing on consonant clusters*. Assen: Van Gorcum.
- Downing, L. J., T. A. Hall, and R. Raffelsiefen, eds. Forthcoming. *Paradigms in phonological theory*. Oxford: Oxford Univ. Press.
- Eisner, J. 1997. FOOTFORM decomposed: Using primitive constraints in OT. In *Proceedings of the eighth Student Conference in Linguistics*, ed. B. Bruening. MITWPL 31. Cambridge, Mass.: MIT.
- Elmedlaoui, M. 1993. Gemination and spirantization in Hebrew, Berber, and Tigrinya: A fortis-lenis module analysis. *Al-Tawāṣul al-Lisānī (Linguistica Communicatio)* 5:121–76.
- Gesenius, F. H. W. 1910. *Gesenius' Hebrew grammar*. Edited and enlarged by E. Kautzsch; 2nd English edition by A. E. Cowley. Oxford: Clarendon.
- Green, A. D. In prep. *Phonology limited*. Book ms., University of Potsdam.
- Guerssel, M. 1977. Constraints on phonological rules. *Linguistic Analysis* 3:267–305.
- Hall, T. A. 2000. *Phonologie: Eine Einführung*. Berlin: Walter de Gruyter.
- Hayes, B. 1986. Inalterability in CV phonology. *Language* 62:321–52.
- Hayes, B. 1995. *Metrical stress theory: Principles and case studies*. Chicago: Univ. of Chicago Press.

- Hetzron, R. 1987. Hebrew. In *The world's major languages*, ed. B. Comrie, 686–704. New York: Oxford Univ. Press.
- Idsardi, W. J. 1997. Phonological derivations and historical changes in Hebrew spirantization. In *Roca 1997*: 367–92.
- Idsardi, W. J. 1998. Tiberian Hebrew spirantization and phonological derivations. *Linguistic Inquiry* 29:37–73.
- Idsardi, W. J. 2000. Clarifying opacity. *The Linguistic Review* 17:337–50.
- Inkelas, S., and Y. Y. Cho. 1993. Inalterability as prespecification. *Language* 69:529–74.
- Iverson, G. K. 1995. Rule ordering. In *The handbook of phonological theory*, ed. J. Goldsmith, 609–14. Oxford: Blackwell.
- Kager, R. 1993. Alternatives to the iambic-trochaic law. *Natural Language & Linguistic Theory* 11:381–432.
- Kager, R. 1999. *Optimality theory*. Cambridge: Cambridge Univ. Press.
- Kaye, A. S., ed. 1997. *Phonologies of Asia and Africa*. Winona Lake, Ind.: Eisenbrauns.
- Kenstowicz, M. 1982. Gemination and spirantization in Tigrinya. *Studies in the Linguistic Sciences* 12:103–22.
- Kenstowicz, M. 1995. Cyclic vs. non-cyclic constraint evaluation. *Phonology* 12:397–436.
- Kenstowicz, M. 1996. Base-identity and uniform exponence: Alternatives to cyclicity. In *Durand and Laks 1996*: 363–93.
- Kenstowicz, M. 1997. Uniform exponence: Exemplification and extension. In *Selected phonology papers from H-O-T 97*, ed. V. Miglio and B. Morén, 139–54. University of Maryland Working Papers in Linguistics 5. College Park: Univ. of Maryland. MS also available as ROA-218, Rutgers Optimality Archive, <http://roa.rutgers.edu/>. Revised version to appear in *Bollettino della Società Linguistica Italiana*.
- Kenstowicz, M., and C. Kisseberth. 1971. Unmarked bleeding orders. *Studies in the Linguistic Sciences* 1:8–28. Repr. in *Studies in generative phonology*, ed. C. Kisseberth, 1–13. Edmonton: Linguistic Research, 1973.
- Kenstowicz, M., and C. Kisseberth. 1977. *Topics in phonological theory*. New York: Academic Press.
- Khan, G. 1997. Tiberian Hebrew phonology. In *Kaye 1997*: 85–102.
- Kiparsky, P. 1968. Linguistic universals and linguistic change. In *Universals in linguistic theory*, ed. E. Bach and R. T. Harms, 171–202. New York: Holt, Rinehart & Winston. Repr. in *Kiparsky 1982b*, 13–43.
- Kiparsky, P. 1971. Historical linguistics. In *A survey of linguistic science*, ed. W. O. Dingwall. First edn., College Park: Univ. of Maryland, 576–649. Second ed., Stamford: Greylock, 33–62. Repr. in *Kiparsky 1982b*, 57–80.
- Kiparsky, P. 1973. Abstractness, opacity and global rules. In *Three dimensions of linguistic theory*, ed. O. Fujimora, 1–136. Tokyo: Taikusha.
- Kiparsky, P. 2003. Syllables and moras in Arabic. In *Féry and van de Vijver 2003*: 147–82.
- Kirchner, R. 2000. Geminate inalterability and lenition. *Language* 76:509–45.
- Malone, J. L. 1993. *Tiberian Hebrew phonology*. Winona Lake, Ind.: Eisenbrauns.
- McCarthy, J. J. 1979. Formal problems in Semitic phonology and morphology. PhD diss., MIT. Pub. New York: Garland, 1985.
- McCarthy, J. J. 1999. Sympathy and phonological opacity. *Phonology* 16:331–99.

- McCarthy, J. J. 2003. Sympathy, cumulativity, and the Duke-of-York gambit. In *The syllable in Optimality Theory*, ed. C. Féry and R. van de Vijver, 23–76. Cambridge: Cambridge Univ. Press.
- McCarthy, J. J., and A. S. Prince. 1993. Prosodic morphology: Constraint interaction and satisfaction. RUCCS Technical Report #3. Revised version (2001) available as ROA-482, Rutgers Optimality Archive, <http://roa.rutgers.edu/>.
- McMahon, A. 2000. *Change, chance, and optimality*. Oxford: Oxford Univ. Press.
- Noyer, R. 1997. Attic Greek accentuation and intermediate derivational representations. In *Roca 1997*: 501–27.
- Paradis, C. 1997. Non-transparent constraint effects in Gere: From cycles to derivations. In *Roca 1997*: 529–50.
- Prince, A. 1990. Quantitative consequences of rhythmic organization. In *Papers from the 26th Regional Meeting of the Chicago Linguistic Society*, ed. M. Ziolkowski, M. Noske, and K. Deaton, 2:355–98. Chicago: Chicago Linguistic Society.
- Rendsburg, G. A. 1997. Ancient Hebrew phonology. In *Kaye 1997*: 65–83.
- Roca, I., ed. 1997. *Derivations and constraints in phonology*. Oxford: Clarendon.
- Rubach, J. 1997. Extrasyllabic consonants in Polish: Derivational optimality theory. In *Roca 1997*: 551–81.
- Sanders, R. N. 2003. Opacity and sound change in the Polish lexicon. PhD diss., Univ. of California at Santa Cruz. ROA-603, Rutgers Optimality Archive, <http://roa.rutgers.edu/>.
- Schein, B., and D. Steriade. 1986. On geminates. *Linguistic Inquiry* 17:691–744.
- Spencer, A. 1996. *Phonology: Theory and description*. Oxford: Blackwell.
- Steiner, R. C. 1997. Ancient Hebrew. In *The Semitic languages*, ed. R. Hetzron, 145–73. London: Routledge.
- Steriade, D. 2000. Paradigm uniformity and the phonetics-phonology boundary. In *Acquisition and the lexicon*, ed. M. Broe and J. Pierrehumbert, 313–34. *Papers in Laboratory Phonology* 5. Cambridge: Cambridge Univ. Press.

Phonological Phrases in Xhosa (Southern Bantu)¹

Sabine Zerbian

ZAS Berlin

This paper investigates how syntax and focus interact in deriving the phonological phrasing of utterances in Xhosa, a Bantu language spoken in South Africa. Although the influence of syntax on phrasing is uncontroversial, a purely syntactic analysis cannot account for all the data reported for Xhosa by Jokweni (1995). Focus influences the phrasing in that it inserts a phonological phrase-boundary after the focused constituent. This generalization can account for the variation found in the phrasing of adverbials.

The findings are dealt with in an OT-based framework following Truckenbrodt's work on Chichewa (1995, 1999) which is extended to the phrasing of adjuncts.

1 Introduction

Work on vowel length in Chi Mwi:ni (Kisseberth and Abasheikh 1974) or on tone assignment in Xiamen Chinese (Chen 1987), to name just two, brought evidence for the assumption that certain phenomena are sensitive to prosodic domains, and that these domains stand in a certain relationship to syntactic domains. In both languages, Chi Mwi:ni and Xiamen Chinese, the correct domains can be predicted by aligning the right edges of syntactic XPs with prosodic boundaries.

In this paper I re-examine the interrelation of syntactic and phonological structure in phrasing in Xhosa based on the data provided in Jokweni (1995), a Southern African Bantu language spoken in South Africa. I follow the standard view that an approach in terms of prosodic structure domains defined in relation to syntactic structure provides a general rule for domain-related phonological processes such as the lengthening of vowels in penultimate position. However, reference to syntax alone in defining prosodic domains encounters problems with respect to the syntactic status of adverbs. Instead I show that focus has to

¹ I thank Hubert Truckenbrodt, Caroline Féry, David Odden and Laura J. Downing for discussing aspects of this work and valuable feedback.

be taken into consideration when accounting for phonological phrases in Xhosa. This view is in accord with research done for other Bantu languages for which focus is claimed to have considerable influence on phonological phrasing (e.g. Kanerva 1990, Truckenbrodt 1995, 1999, Jokweni 1995). In order to express the influence focus has on the phrasing derived by syntax, I will formalize the findings in an OT-based approach following Truckenbrodt (1995, 1999) who has accounted for the influence of focus on phrasing in Chichewa. Furthermore, I will extend his theory which has accounted for argument structures in Chichewa to adjunct structures in Xhosa.

The paper is organized as follows: In section 2 I will present phonological processes that apply in certain prosodic domains as evidence for the need of phonological phrases in Xhosa. In section 3 I present the data on phrasing in Xhosa and discuss how Jokweni (1995) accounts for the observed boundaries. In section 4 I show that focus has a systematic influence on phrasing in Xhosa. Section 5 deals with the theoretical treatment of conflicts of syntax and prosody following Selkirk (1986, 1995) and Truckenbrodt (1995, 1999). In section 6 the analysis for Xhosa is presented. It will be shown that the theory has to be extended to the phrasing of adjunct structures in order to account for the Xhosa data. The paper closes with a summary in section 7.

The paper presents a preliminary approach to these aspects in Xhosa as it is solely based on data available in the literature. Since the goal of the available literature is different from the goals pursued in this paper, revealing data on e.g. the phonology-syntax interface are missing in the original literature or negative data are not supplied. The issues touched upon on here will be pursued in further research with mother tongue speakers.

2 Phonological Processes at phonological phrase boundaries

Evidence for pP-boundaries in Xhosa comes from two phonological processes that indicate right edges of phonological phrases (Jokweni 1995). One is the lengthening of the vowel in penultimate position, the other process involves tonological changes. Both processes will be dealt with in turn.

2.1 Vowel lengthening

Jokweni (1995) shows that the process of Penultimate Vowel Lengthening in Xhosa provides evidence for the existence of the Phonological Phrase (pP)².

² I use a somewhat different notation for the prosodic units (pP, iP) in order to avoid confusion with the syntactic constituents (PP, IP).

In Xhosa as in many other Bantu languages there is no vowel length contrast in words. All vowels appear as short vowels. Nevertheless, depending on its position in the sentence an underlying short vowel can surface as long in the penultimate syllable of a word. Literature on the phonology of Bantu languages suggest that the long vowel in penultimate position is the result of a boundary-related phonological process. For Xhosa, Jokweni (1995) proposes that this process of penultimate vowel lengthening applies at the right edge of a pP. The size of the pP varies with respect to the number of constituents it comprises, as the examples in (1) show (pP-boundaries are marked by parentheses, acute marks high tones):

- (1) pPs in Xhosa, Jokweni (1995: 26f, 31)³
- (a) (íntaaka)_{pP} ,bird‘
(intákaana)_{pP} ,small bird‘
(ba-vúl’ íncwaadí)_{pP} ,They are opening a book.‘
SC-open book 9
- (b) (bá-ya-yi-vúúl’)_{pP} (íncwaadí)_{pP}, They are opening the book.‘
SC-YA-OC9-open book 9
- (c) (íncwaadí)_{pP} (bá-ya-yi-vúúla)_{pP}, As for the book, they open it.‘
book SC-YA-OC9-open

The examples in (1) show variations in pP-boundaries that Jokweni (1995) uses to argue for the necessity of the prosodic unit pP. The example in (1b) shows that Penultimate Lengthening cannot be related to the fact that a word appears at the boundary of an utterance as could be imagined from the examples in (1a). The variation in length of the penultimate vowel of the verb given in (1b) shows that the lengthening process is related to other factors than utterance finality as also utterance-medial words can be lengthened. Jokweni (1995) proposes that lengthening is sensitive to pPs and that it appears at pP-boundaries. Consequently, a pP-boundary intervenes between the verb and object in (1b) and between the object and verb in (1c). The rule of Penultimate Lengthening is given in (2).

³ The abbreviations used in the glosses are the following:

SC subject concord
OC object concord with numbers indicating noun classes
YA morpheme of the short form of the present tense
nouns are given followed by their noun class number
bold printing indicates focus

- (2) Penultimate Lengthening in Xhosa, Jokweni (1995: 25)
 $\emptyset \rightarrow V / _ CV)_{pP}$

The rule in (2) states that a short vowel becomes long in penultimate position of a phonological phrase.

2.2 *Word-final High Tone Deletion*

With respect to tonal evidence Jokweni's work (1995) provides examples that show that the rule of word-final H deletion only applies pP-medially. When occurring at the edge of pP a final high tone of a word is kept, as shown in the examples in (3).

- (3) Word-final H deletion, Jokweni (1995: 117, 128)
- (a) (kuutyá)_{pP} 'to eat'
 - (b) (bá-tya ínyaama)_{pP} 'They eat meat.'
 SC-eat meat
 - (c) (bá-ya-yii-tyá)_{pP} (ínyaama)_{pP} 'They eat it, meat.'
 SC-YA-OC9-eat meat 9

Example (3a) shows that the verb has a high tone on the final syllable when it appears in isolation. This high tone is not realized when the verb form appears phrase-medially, as shown in (3b). In (3c), the right-dislocated object constitutes its own pP. Therefore the H appears on the verb as it ends a pP.

Both phonological processes, vowel lengthening and word-final H deletion, are diagnostics for pP-boundaries in Xhosa. I will only refer to lengthening phenomena in the following as they are easily indicated in the transcription by a doubled vowel.

3 **Phrasing in Xhosa: Jokweni's proposal (1995)**

In the preceding section I have shown that a right pP-boundary can be detected in Xhosa by a lengthened penultimate vowel. This section discusses the relation between syntax and phrasing in Xhosa. On the basis of the data presented in this section, Jokweni (1995) proposes a language-specific algorithm in order to account for the derivation of phonological phrase boundaries in Xhosa. Jokweni (1995: 52ff) shows that algorithms proposed as universal in the literature (Selkirk (1980), Selkirk (1986), and Nespor & Vogel (1986)) do not account for the data in Xhosa, nor that language-specific algorithms (McHugh (1990), Chen (1987), Hale & Selkirk (1987)) can be transferred to Xhosa. Selkirk's proposal

(1980, 1986) will be discussed in section 5 in more detail. In the following I will have a look at the phrasing of VO- and of V(O)Adv-structures and how they are accounted for in Jokweni (1995).

3.1 The phrasing of VO-structures

First, complement structures will be considered, that display the basic word order (SVO), i.e. that object(s) follow the verb. Consider the data in (4).

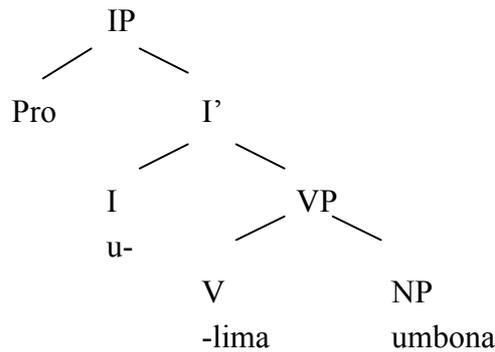
- (4) Phrasing in Xhosa, Jokweni (1995: 31f)
- (a) (abántwana ba-kháb' íbhóóla)_{pP}
children SC-kick ball
'Children are kicking the ball.'
- (b) (ba-vúl íncwaadí)_{pP}
SC-open book
'They open the book.'

The sentences in (4) show simple transitive sentences ((4a) with, (4b) without explicit subject). They all display the basic word order SVO, in which the complements directly follow the verb. Furthermore, they all have in common that they are phrased into one single pP as indicated by the lengthened vowel only in penultimate syllable of the last word in the sentence. In the cases in (4), the pP is coextensive with the utterance. A pP-boundary is inserted at the right edge of the utterance. The syntactic structure following Haegemann (1991)⁴ is illustrated in (5).

- (5) Phrasing in Xhosa, Jokweni (1995: 38f)
- (a) (u-lím' úmbóóna)_{pP}
SC-cultivate maize
'He/ she cultivates maize.'

⁴ The syntactic structure above the VP is a matter of controversy in Bantu languages. Whereas Jokweni (1995) and Spuy (1993) adopt the generative framework, Bresnan (1994) and Morimoto (2000) e.g. do not assume two functional phrases above the VP. However, it shall not matter here, how the superior syntactic phrase is labelled. It is only important that it is no lexical phrase but a functional phrase.

- (b) syntactic structure following Jokweni 1995: 39
- ⁵



In (5), a transitive verb surfaces with a substantival complement. Evidence for the complement status is that it is obligatory (**Uliima.*) and that no constituent can intervene between the verb and the object (see also van der Spuy (1993) for a similar argument).

In order to account for this phrasing in Xhosa, Jokweni proposes the algorithm in (6) for the construction of pPs.

- (6) Preliminary version of the PP-algorithm for Xhosa (Jokweni 1995:24)
 A head X of a syntactic phrase XP phrases with all elements within the maximal projection XP.

The algorithm in (6) assumes that all postverbal constituents in (5) are phrased together with the verb in a VP. It can therefore correctly account for the observed phrasing in (5)⁶.

Now consider cases where a comparable sentence is divided into more than one pP, as shown in (7).

- (7) Phrasing in Xhosa, Jokweni (1995: 31f, 55)
- (a) (bá-ya-yi-vúúl')_{pP} (íncwaadí)_{pP}
 SC-YA-OC9-open book 9
 'They open it, as for the book.'
- (b) (ba-m-ník' úkuuty')_{pP} (úmaamá)_{pP}
 SC-OC1-give food mother 1
 'They give her food, mother.'

⁵ The substantival object can also be expressed by an object marker which appears left to the verb as in (u-ya-wu_i-liima pro_i)- 'He/ she cultivates it.'. In this case Jokweni assumes that the position of the substantival NP is occupied by an object *pro*. I will not consider these cases here.

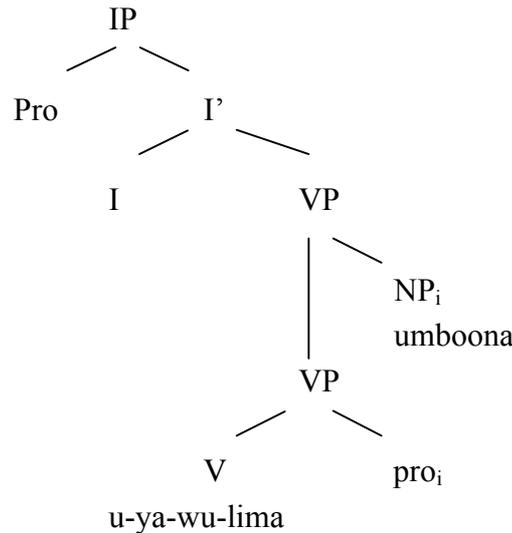
⁶ The phrasing of the subject will be dealt with separately in section 6.4.

In the examples in (7) the sentence is divided into two pPs. The syntax of the construction in (7) is exemplified in (8).

(8) Phrasing in Xhosa, Jokwani (1995: 39f)

- (a) (ú-ya-wu_i-líim' *pro*_i)_{PP} (úmbóona_i)_{PP}
 SC-YA-OC-cultivate maize
 'He/ she cultivates it, maize.'

(b) syntactic structure from Jokwani (1995: 40)



In (8), a transitive verb surfaces with both a pronominal marker and a substantival NP. The object is optionally included in the sentence as the object pronominal already fulfills the function of the complement of the transitive verb (with respect to *pro* see footnote 5). The substantival object is not part of the VP but is adjoined to the VP instead⁷. A comparable analysis has also been proposed by Bresnan (1987) for similar data in Chichewa, and by Spuy (1993) for Zulu. Besides the optionality of the object NP and the presence of the object marker there is further evidence for the adjunction of the NP outside the VP: With adjoined object NPs movement and intervening constituents are possible. The examples in (9) illustrate that syntactic structures as in (8) allow a wider range of word order variations than the basic SVO order.

(9) Phrasing in Xhosa, Jokwani (1995: 49, 50, 47)

- (a) (íncwaadí)_{PP} (bá-ya-yi-vúúla)_{PP}
 book 9 SC-YA-OC9-open
 'The book, they open it.'

⁷ Jokwani assumes the adjoined object to be in the upper VP projection in order to be in line with the phrase structure as outlined in Haegemann (1991: 44). The question where the dislocated NP is adjoined to is of no importance at the moment.

- (b) (Lé ndooda)_{pP}(abántwaana)_{pP} (í-ya-ba-béétha)_{pP}
 this man children 2 SC-YA-OC2- beat
 ‘This man, children, he beats them.’
- (c) (bá-ya-yi-vúúl’)_{pP} (ébusuuk’)_{pP}(íncwaadí)_{pP}
 SC-YA-OC9-open at-night book 9
 ‘They open it at night, the book.’

The examples in (9) all show word orders that are only possible when the object NP cooccurs with an object concord on the verb. (9a) shows the object left-dislocated, preceding the subject, (9b) left-dislocated following the subject, and (9c) right-dislocated with an adverbial intervening between verb and object.

Jokweni (1995) accounts for the separate phrasing by modifying the phrasing algorithm, as in (10).

- (10) Final PP-algorithm for Xhosa as proposed by Jokweni (1995: 24)
- (a) a head X of a syntactic phrase XP phrases with all elements within the maximal projection XP provided those elements are not adjoined.
- (b) Adjuncts or adjoined elements each constitute a pP of their own.

To phrase it differently, Jokweni (1995) proposes that a head is phrased together with all elements within a maximal projection. Adjuncts to a maximal projection are phrased separately.

Jokweni’s (1995) analysis similarly applies to ditransitive constructions, as shown in (11).

- (11) Phrasing of ditransitive verbs in Xhosa, Jokweni (1995: 55)
- (ba-ník’ úmam’ úkuutyá)_{pP}
 SC-give mother food
 ‘They give mother food.’

In (11), the whole sentence is contained in one pP. Both objects are complements to the verb and are phrased within the maximal projection of the verb. The derivation of pP-boundaries by the algorithm in (10) makes Jokweni’s analysis superior to earlier analyses such as Nespor & Vogel (1986), and Selkirk (1980, 1986), as in Xhosa both object NPs phrase together. Algorithms by both Nespor & Vogel (1986), and Selkirk (1980, 1986) predict a separate phrasing of the second object (see demonstration and detailed discussion of Selkirk (1980, 1986) in section 5).

To sum up, in the cases in (4) the object is complement to the verb and is phrased together into one VP, and consequently, according to (10), into one pP. In the cases in (7), the object NP is represented by an object marker on the verb,

and the optional object (or objects) is (are) no longer complement but adjunct to VP. As the head V is phrased with all elements in its VP except adjoined elements, the adjoined object NPs constitute their own pP. We will have a look at the phrasing for adjuncts in the next section.

3.2 Phrasing of V-(O)-Adv-structures

In Xhosa, adverbials appear at the right sentence periphery following the verb and its objects in basic word order. Consider (12) to see the variation with respect to phrasing in adverbials.

(12) Phrasing of adverbials, Jokweni (1995: 60, 64)

- (a) (ba-líma ngo-Mqgibéèlo)_{pP}
SC-plough on-Saturday
'They plough on Saturday.'
- (b) (ba-yá-liima)_{pP} (ngoMqgibéèlo)_{pP}
SC-YA-plough on-Saturday
'They plough on Saturday.'

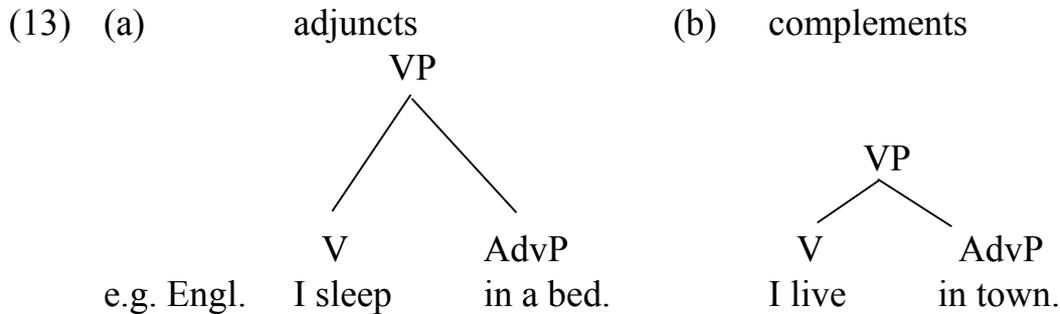
In (12a), the adverbial phrases together with the verb as can be seen from the missing long vowel in penultimate position in the verb. In (12b), the sentence is split up into two pPs. The first pP ends after the verb, and the adverbial is phrased in a separate pP.

A syntactic account assumes two different syntactic structures for the two different phrasings in (12), see e.g. Jokweni (1995). Adverbs are therefore assigned ambiguous syntactic structure, sometimes inferring being phrased together with the verb, sometimes inferring being phrased separately. This is problematic as there is no further evidence that supports such an ambiguous syntactic status in minimal pairs as (12)⁸. Adverbials are always optional (except for verbs that subcategorize for an adverbial) and can consequently be dropped without the sentence becoming ungrammatical.

A purely syntactic approach to the difference in phrasing weakens the classical notions of complements and adjuncts. In the following section, I will show that the data in (12) can be accounted for when focus is taken into consideration. With regard to the syntactic position of adverbials discussed in this section, I will assume that adjuncts are optional by definition as

⁸ Jokweni (1995) claims that the distribution of the verbal particle *-ya-* shows further evidence for the ambiguous status of adverbs. According to Jokweni (1995) the absence of the verbal particle *-ya-* in (12a) indicates that the verb subcategorizes for an overt complement. But see Zerbian (2003) for a different analysis concerning the distribution of *-ya-* which makes no reference to grammatical relations but to pPs instead.

obligatoriness means being subcategorized for by the verb which adverbials are not. Furthermore, I will assume that also optional elements can be phrased within the VP (see (13)). So far there is no evidence for a structural difference between complements and adjuncts in Xhosa.



In (13a) the adverbial phrase is an adjunct as the verb *sleep* does not subcategorize for a local adverbial. The verb *live* in (13b), however, does categorize for a local adverbial.

4 Prosodic marking of focus

In the preceding section I have shown that a purely syntactic analysis of the data on phrasing in Xhosa assigns adverbials an undesirable ambiguous status. In this section I want to show that taking focus into consideration can explain the phrasing of adverbials. The influence of focus was mentioned in Jokweni (1995:14) but was not integrated in his algorithm for phonological phrasing in Xhosa.

In 4.1. I present some theoretical background on focus, in 4.2. I will present the Xhosa data in a way that show the influence of focus.

4.1 Focus

In Western European languages, intonational accent is the main way focus is expressed, while focus position and focus morphology play more minor roles. However, like most of the Bantu languages Xhosa is a tone language, where intonational accent is not reported to correlate with focus. Instead, tone languages are reported to use alternative means of expressing focus: non-intonational prosodic phrasing, focus morphemes on the verb and particles, and fixed clause position.

In this paper I will concentrate on the use of non-intonational prosodic phrasing as a means of expressing focus, although also morpho-syntactic changes are of importance as can be easily seen from the examples. For the explanation of the Xhosa data presented so far a distinction of two different

kinds of focus is necessary: sentence focus (or all-new/ wide focus) and term focus (narrow focus). Sentence focus means that all the information conveyed in the sentence is new to the hearer. Term focus means that only a piece of information is new to the hearer. Focus structures can be easily and unambiguously elicited by means of questions. The interrogative elicits the new information, so that in the answers the constituent corresponding to the interrogative word is automatically placed in focus. Chomsky (1971: 199) and Jackendoff (1972: 230) express the general rule that the question and answer should have equivalent focus structures. Wide focus turns up in answers to questions such as ‘*What happened?*’, or ‘*What are they doing?*’ where all the information given is new. Narrow focus appears in answers to Wh-questions.

4.2 Influence of focus in Xhosa

The focus structure of the examples presented so far has been elicited by Jokweni (1995) by means of questions. Rearranging the Xhosa data according to the kind of focus they express show a striking difference in phrasing. The sentences in (14) all display wide focus/ sentence focus (as indicated by the question).

(14) Phrasing under wide focus, Jokweni (1995: 31, 47, 60, 67, 74)

- (a) (ba-vúl’ íncwaadí)_{pp} (Q: What are they doing?)
SC-open book
‘They are opening a book.’
- (b) (ba-ník’ úmam’ úkuutyá)_{pp}
SC-give mother food
‘They give the mother food.’
- (c) (ba-líma ngo-Mqgibèèlo)_{pp}
SC-plough on-Saturday
‘They plough on Saturday.’
- (d) (ba-líla kákuubí)_{pp}
SC-cry badly
‘They cry badly.’
- (e) (ndi-tyél’ ínyam’ émbizééni)_{pp}
SC-eat meat in.pot
‘I eat meat in the pot.’

Syntactically, the sentences in (14) display the unmarked word order SVO and are phrased together in one pP. This is true for all constructions considered so far: (14a) shows a transitive verb, (14b) a ditransitive verb, (14c) a verb followed by a temporal, (14d) by a modal adverbial, and finally (14e) a verb

followed by an object and a local adverbial. In terms of focus one can conclude that in sentences displaying wide focus the VP is phrased into one pP.

The data in (15) show that it is due to narrow focus that the phrasing is split up.

- (15) Phrasing under narrow focus, Jokweni (1995: 65, 69, 64, 134)
- (a) (bá-zaku-liima)_{pP} (ngezáándla)_{pP}
 SC-Fut-plough by-hands
 ‘They are going to plough (not do something else) by hands.’
- (b) (bá-ya-fudúúka)_{pP} (ngoLwésihláánu)_{pP}
 SC-YA-emigrate on-Friday
 ‘They do emigrate on Friday.’
- (c) (ba-yá-liima)_{pP} (ngoMqgibéèlo)_{pP}
 SC-YA-plough on-Saturday
 ‘They plough on Saturday.’
- (d) (ba-ya-kuutya)_{pP} (ukuutya)_{pP} (ngezaandla)_{pP}
 SC-YA-eat food by.hands
 ‘They do eat food by hands.’

The data in (15) show that narrow focus induces a phrase boundary following the focused constituent. In (15a-d), the verb is focused in contrast to the sentences in (14). Note, that there is no evidence for a change in syntactic structure. A phrase-boundary occurs after the focused verb. The examples show that focus influences the phrasing in that the focused constituent triggers the insertion of a pP-boundary. One can further observe that in verb forms in the Present Tense (15b-d) an additional morpheme *-ya-* occurs on the verb. This morpheme has been analysed as a focus marker, marking predicate focus on the verb (Güldemann 1996). The status of this morpheme is unclear (for a semantically oriented analysis for comparable facts in Sotho see Kosch (1988), for a syntactic analysis see Spuy (1993), for a prosodic analysis see Zerbian (2003)). However, the insertion of prosodic boundaries also occurs in other tenses than the Present Tense, see (15a), justifying a more general analysis.

To sum up, the data in (14) and (15) show that a given sentence differs in phrasing according to its information structure, i.e. the constituents that are in focus. Whereas in (14) under wide focus all postverbal elements are phrased together, in (15) where the verb is in focus, the phrasing is split up. However, focus alone cannot predict the phrasing neither. Consider the examples in (16).

- (16) Phrasing in Xhosa, Jokwani (1995: 48, 49)
- (a) (ba-kháb' íbhóól')_{pP} (ébusuuk')_{pP} (abántwaana)_{pP}
SC-kick ball at night children 2
'They kick the ball, at night, as for children.'
- (b) (íncwaadí)_{pP} (ébusúuku)_{pP} (bá-ya-yi-vúúla)_{pP}
book 9 at night SC-YA-OC 9-open
'As for the book, at night, they open it.'

Though focus determines the phrase boundary after the focused argument *ball* in (16a), the phrase boundary after the following adjunct is not determined by focus but by syntax as the following subject is right-dislocated. Similarly, in the example in (16b) syntax alone determines the phrasing of the adjunct *at night* as it is not in focus but left-dislocated for other pragmatic reasons. Therefore, in the examples in (16) morphosyntactic changes play a crucial role and determine the phonological phrasing. Consequently, it is the interaction of pragmatic (focus) and syntactic (dislocation) considerations that determines phonological phrasing in Xhosa.

In the following I will account for the competing forces of syntax and focus in phrasing by presenting Truckenbrodt's (1995, 1999) approach to comparable data on phrasing in Chichewa (Kanerva 1990). Acknowledging the influence focus has on phrasing correctly predicts the pP-boundaries in Chichewa.

5 Conflicting constraints in syntax and prosody

In section 3 I have shown that the phrasing in Xhosa cannot be accounted for by reference to syntax alone, instead I have shown in section 4 that focus is necessary to be taken into consideration as well.

Truckenbrodt's (1995, 1999) work investigates the relation of syntax, focus, and phrasing in Chichewa, a Bantu language spoken in Malawi, basing his analysis on original work by Kanerva (1990) who showed that phrasing is crucially determined by focus in Chichewa. In his work on the relation of syntactic XPs and prosodic structure Truckenbrodt (1995, 1999) shows that patterns of phrasing can better be understood when the syntax-prosody mapping is formalized in terms of ranked and violable constraints as in the framework of Optimality Theory (Prince and Smolensky 1993). He improves on Selkirk's (1986, 1995) end-based theory of the syntax-prosody relation in accounting for data in Tohono O'odham, Kimatuumbi, and Chichewa, which have not been accounted for in the end-based theory.

In this section I want to develop an analysis for the data in Xhosa presented in section 3 in the OT-framework following the work of Truckenbrodt

(1995, 1999) for Chichewa. I want to show how the account proposed here can account for the phrasing in Xhosa. I will first show the cases whose phrasing is syntax-based such as the complement structures, also in connection with focus. I will then proceed to the crucial cases of adverbials and show that the OT-based analysis that considers focus can account for the phrasing if extended to adverbials. Finally, I will present the phrasing of the subject in Xhosa, and show that further research is needed to provide independent evidence for the syntactic structure of the subject.

5.1 *The prosody-syntax interface*

Based on the observations that certain phenomena are sensitive to prosodic domains and that these domains stand in a certain relationship to syntactic domains (e.g. Kisseberth & Abasheikh (1974), Chen (1987)), Selkirk (1986) proposes a universal theory of phrasing which formulizes the edge alignment of phonological phrases with syntactic XPs. This theory is called end-based theory as it inserts a prosodic phrase boundary at the end of syntactic phrases. Selkirk (1995) reformulates her former analysis in the format of General Alignment (McCarthy & Prince 1993). The constraint for aligning syntactic XPs with prosodic boundaries in order to derive prosodic domains following Selkirk (1995) reads as in (17).

- (17) ALIGN-XP,R: ALIGN (XP, R; P, R) (following Truckenbrodt (1999))
For each XP there is a P such that the right edge of XP coincides with the right edge of P.
ALIGN-XP,L: ALIGN (XP, L; P, L)
For each XP there is a P such that the left edge of XP coincides with the left edge of P.

According to Selkirk (1995), every language makes a paradigmatic choice between right- or left alignment. In the following I will restrict myself to right-alignment of syntactic phrases with prosodic boundaries as this will be the relevant parameter for the analysis of Chichewa and Xhosa.

In his work on phonological phrases and their relation to syntax, focus, and prominence Truckenbrodt (1995, 1999) introduces the constraint WRAP-XP. He claims that this constraint exists besides Selkirk's edge alignment, and that its relation to the alignment constraints can explain the variation in phrasing found in the languages of the world. Moreover, it can explain the data on phrasing found in Kimatuumbi and Chichewa which have not been accounted for before. The constraint WRAP-XP reads as in (18).

- (18) WRAP-XP
Each XP is contained in a phonological phrase.

One effect of the constraint in (18) is that it reduces prosodic structure. In contrast to ALIGN (XP, R; P, R) as stated by Selkirk (1995) it only ensures that every XP is contained in one pP, but it does not require every XP to be aligned with a prosodic boundary, thus constituting its own phrase. As WRAP-XP is claimed to exist besides the alignment constraint, these two constraints are in conflict and the stronger influence of one constraint or the other yields at different phrasings. The different phrasings are shown in (19) and (20).

(19) Predicted phrasing when ALIGN is dominant (following Selkirk 1995)

- | | | | | | | | |
|-----|------------------|---------------------------------|----------------------------------|------------------|----------------------------------|------------------|-----------------------------------|
| (a) | [XP ₂ | X ₁] _{XP1} | (b) | [X ₁ | XP ₂] _{XP1} | | |
| | () _P | () _P | | () _P | () _P | | |
| (c) | [X ₁ | XP ₂ | XP ₃] _{XP1} | (d) | [XP ₁ | Fct | XP ₂] _{FctP} |
| | () _P | () _P | () _P | () _P | () _P | () _P | |

(20) Predicted phrasing when WRAP-XP is dominant (following Truckenbrodt 1999)

- | | | | | | | | |
|-----|------------------|---------------------------------|----------------------------------|------------------|----------------------------------|------------------|-----------------------------------|
| (a) | [XP ₂ | X ₁] _{XP1} | (b) | [X ₁ | XP ₂] _{XP1} | | |
| | () _P | () _P | | () _P | () _P | | |
| (c) | [X ₁ | XP ₂ | XP ₃] _{XP1} | (d) | [XP ₁ | Fct | XP ₂] _{FctP} |
| | () _P | () _P | () _P | () _P | () _P | () _P | |

When comparing the different predictions in terms of phrasing made by the dominance of one of the two constraints we can see that in some cases they coincide (b, d) but in others they differ (a, c). (20a, c) shows that WRAP-XP is a structure-preventing constraint as it reduces the number of pPs in these constructions.

The examples from Chichewa and Xhosa illustrate the effects of the two constraints in the following section.

6 The case of Xhosa

6.1 Phrasing under wide focus

As shown by the examples in (14) in Xhosa sentences show a strong tendency to be phrased into one pP in the pragmatically unmarked case. Under wide focus, complements are phrased into one phrase in transitive and ditransitive constructions. Tableau 1 shows the phrasing of simple transitive structures in Xhosa (example 14a, repeated as (21)).

- (21) (ba-vúl' íncwaadí)_{pp} (Q: What are they doing?)
 SC-open book
 'They are opening a book.'

Tableau 1 : Basic phrasing in Xhosa

[V NP] _{VP}	WRAP-XP	ALIGN-XP, R	NONREC
a. ()P			
b. () ()P	*!		
c. (()P)P			*!

The tableau in (1) lists possible candidates for the phrasing of simple transitive structures. Candidate a is the winner: The whole VP is phrased into one phonological phrase. As shown in (19b, 20b) this can be accounted for by either ranking WRAP-XP above ALIGN-XP, R or vice versa since the transitive structure violates none of these constraints. Simple transitive structures do not provide the necessary data to decide between these two rankings. However, tableau 1 shows that NONREC, a constraint against recursive structures, must be ranked high enough to prevent the corresponding candidate (c) to emerge as the winning candidate. As recursive structures have been shown not to be allowed in Chichewa (Truckenbrodt 1995, 1999) and as there is no evidence for recursive structures in Xhosa neither, candidates displaying recursive structures will not be considered in the following tableaux⁹.

The case of double object constructions is the crucial case that motivates the constraint WRAP-XP for Chichewa and also for Xhosa.

- (22) (a) Chichewa (Kanerva 1990)
 (tinapatsa mwana njiinga)_{pp}
 we-gave child bicycle
 'We gave the child a bicycle.'
- (b) Xhosa (Jokweni 1995)
 (ba-ník' úmam' úkuutyá)_{pp}
 SC-give mother food
 'They give the mother food.'

In Xhosa as in Chichewa, ditransitive structures like (22) are phrased into one phonological phrase. They thereby provide evidence that a constraint like WRAP-XP is ranked above ALIGN-XP,R as both Xhosa and Chichewa prefer the

⁹ Neither will the constraint EXHAUSTIVITY (Selkirk 1995: 443) be considered that requires that phrasing on every prosodic level is exhaustive. As words are always parsed into some pP, no candidates with unparsed words will be included in the following tableaux.

whole VP being wrapped in one pP instead of phrasing each XP individually. The latter phrasing would be predicted by the end-based approach (cf. also 19c, 20c).

The analysis for examples as in (22) is shown in tableau 2.

Tableau 2: Basic phrasing in Chichewa and Xhosa

[V NP YP] _{VP}	WRAP-XP	ALIGN XP-R
a. ()P()P	VP!	
b. $\left[\begin{array}{c} \text{ } \\ \text{ } \end{array} \right]_P$		NP

Candidate a in which the VP is phrased into two separate phonological phrases is ruled out by violation of the high-ranked constraint WRAP-XP that favours the VP being wrapped into one phonological phrase. Candidate b emerges as winner. Though it violates ALIGN XP-R by not aligning the first NP with a phonological phrase boundary the violation is not fatal as ALIGN XP-R is ranked below the other relevant constraints.

The phrasing of structures containing an adverbial in Xhosa can be captured by the same account. Relevant examples already mentioned in the data exposition in section 3, 4 are repeated in (23).

- (23) Xhosa, Jokweni (1995 : 60, 67, 74)
- (a) (ba-líma ngo-Mqgibèèlo)_{pP}
 SC-plough on-Saturday
 ‘They plough on Saturday.’
- (b) (ba-líla kákuubí)_{pP}
 SC-cry badly
 ‘They cry badly.’
- (c) (ndi-tyél’ ínyam’ émbizééni)_{pP}
 SC-eat meat in.pot
 ‘I eat meat in the pot.’

The account for the observed phrasing is illustrated in tableau 3 for example (23c).

Tableau 3 : Basic phrasing in Xhosa

[V NP AdvP] _{VP}	WRAP-XP	ALIGN XP-R
a. () _P () _P	VP!	
b. $\left[\begin{array}{c} \text{ } \\ \text{ } \end{array} \right]_P$		NP

Similarly to complements, candidate a in which the VP is phrased into two separate phonological phrases is ruled out by violation of the high-ranked constraint WRAP-XP that favours the VP being wrapped into one phonological phrase. Candidate b emerges as winner. Though it violates ALIGN XP-R by not aligning the object NP with a phonological phrase boundary the violation is not fatal as ALIGN XP-R is ranked below the relevant constraint.

The discussion of the phrasing of transitive, ditransitive and structures containing an adverbial under wide focus in Xhosa supports the need for a constraint like WRAP-XP as proposed by Truckenbrodt (1995, 1999). Neither the end-based approach nor the relation-based approach (Nespor & Vogel 1986) can account for the observation that under narrow focus the verb is phrased into one phonological phase together with the constituents that follow it.

6.2 Phrasing under narrow focus

As shown in section 4, narrow focus leads to syntactic changes in complement structures in Xhosa. The object(s) representing old information are moved out of the VP, leaving an object clitic on the verb¹⁰. The relevant examples are repeated in (24).

- (24) Xhosa, Jokweni (1995: 31f, 55, 78)
- (a) (**bá-ya-yi-vúúl'**)_{VP} (íncwaadí)_{NP}
 SC-YA-OC9-open book 9
 'They open it, as for the book.'
- (b) (ba-m-ník' **úkuuty'**)_{VP} (úmaamá)_{NP}
 SC-OC1-give food mother 1
 'They give her food, mother.'
- (c) (**ndi-ya-m-phékeel'**)_{VP} (úmaam')_{NP} (ínyaama)_{NP}
 SC-YA-OC1-cook.for mother 1 meat
 'I cook for her, mother, meat.'

In (24a) the verb is focused in a simple transitive structure. The object clitic appears on the verb and the verb is phrased separately. In (24b), the patient NP is focused in a ditransitive structure. The object clitic of the benefactive NP appears on the verb and the benefactive NP appears outside the VP. In (24c), the verb is focused in a ditransitive structure. As there is only one slot in the TMA-

¹⁰ The question which phrase the dislocated NPs are adjoined to is not important for the analysis here. Spuy (1993) assumes that they are adjoined to IP, Jokweni (1995) for Xhosa, and Truckenbrodt (1995, 1999) for Chichewa assume that they are adjoined to the VP. Truckenbrodt (1995:235) shows that only the deepest embedded VP is subject to WRAP-XP, not the doubled VPs resulting from the adjunction of dislocated phrases.

template of the verb to accommodate an object clitic the object clitic of the beneficiary NP appears on the verb. However, both objects appear outside the VP, as indicated by prosody and word order.

The analysis for these structures is given in tableau 4. Tableau 4 exemplifies the structure in (24c).

Tableau 4 : Phrasing under narrow focus in Xhosa

	[V _{FOC}] VP [NP] [PP]	WRAP-XP	ALIGN XP-R
a.	() _P		VP !, NP
b.	() _P () _P		NP!
c.	☞ () _P () _P () _P		

As the objects are moved out of the VP, they are no longer under the influence of WRAP-XP. As they are not headed by a shared lexical category the conditions for WRAP-XP to apply are not met. Consequently they are phrased separately in order to maximally satisfy ALIGN-XP, R, as does the winning candidate c.

As has been demonstrated in tableau 4, in double object constructions with the verb in focus, the phrasing is determined by syntax in Xhosa and can be accounted for by the constraints WRAP-XP and ALIGN-XP, R. A parallel analysis along these lines applies for the other examples in (24).

The effect of focus on argument structures reveals interesting variation among Bantu languages. In Xhosa, it results in morphosyntactic changes whereas a different structure is found in Chichewa. In Chichewa, a pP-boundary is inserted after the focused constituent without changes in syntactic structure. Examples are given in (25).

(25) Phrasing in double object constructions in Chichewa (Kanerva 1990)

- (a) Q: What did they do in Mavuto's house?
 (**anagoona**)_{pP} (mnyumba ya mavuuto)_{pP}
 they-slept in-house of Mavuto
 'They *slept* in the house of Mavuto.'
- (b) Q: What did he hit with the rock?
 (anamenya **nyumba**)_{pP} (ndi mwaala)_{pP}
 he-hit house with rock
 'He hit *the house* with the rock.'

Evidence for the emerging pP-boundaries right after the focused constituent is provided by the lengthened vowels in penultimate position (see Kanerva 1990). The effect of boundary insertion after a focused constituent is accounted for by

the constraint ALIGN-FOC which is given in (26) following Truckenbrodt (1995, 1999).

- (26) ALIGN-FOC = ALIGN (FOC, R; P, R)
 ‘Each focused constituent is right-aligned with a p-boundary.’

A constraint similar to (26) has also been proposed for other languages (such as Bengali, Japanese, and Korean). How ALIGN-FOC is integrated into the theory is illustrated in tableau 5 for Chichewa.

Tableau 5: Phrasing under narrow focus in Chichewa

[V _{FOC} NP PP] _{VP}	ALIGN-FOC	WRAP-XP	ALIGN XP-R
a. ()P	*!		NP
b. ()P()P		*	NP!
c. ↵()P()P()P		*	

In contrast to Xhosa, unfocused objects in Chichewa are not moved out of the VP. They therefore fall under the influence of the constraint WRAP-XP (for the need of WRAP-XP in Chichewa see the examples in (22a)). By ranking ALIGN-FOC above WRAP-XP the influence of WRAP-XP is neutralized. Note that this assumption only holds if WRAP-XP is not a gradient constraint, i.e. if WRAP-XP does not differentiate between how many phrase boundaries split up the VP. Because of the neutralization of WRAP-XP the constraint ALIGN-XP has crucial influence on the phrasing of the complements and will therefore prefer each complement in a separate phonological phrase, as in the winning candidate in c.

6.3 Analysis of adverbial structures

As shown in the preceding section, there is variation in languages concerning the question if they show purely phonological reflects of information structure as in Chichewa or (morpho-)syntactic reflects of information structure as in Xhosa¹¹. In the last section the discussion was mainly restricted to arguments. In this section I want to show that the same constraint responsible for prosodic changes in argument structures in Chichewa, namely ALIGN-FOC, can account for focus in adverbial structures in Xhosa. Adverbial constructions in Xhosa show a crucial influence of focus on phrasing in Xhosa. The data have been

¹¹ The question that is not adressed here is what triggers the morphosyntactic changes in Xhosa. One could speculate if the same constraint, namely ALIGN-FOC, is responsible for the observed changes.

presented in examples (12, 15) already. Remember that for a sentence displaying wide focus the same analysis applies as for transitive sentences, see tableau 3.

The same mechanism as in Chichewa complement structures applies to Xhosa adverbial structures when the verb is narrowly focused. Relevant examples are provided in (27).

- (27) Xhosa, Jokweni (1995: 64, 65, 69)
- (a) (ba-yá-liima)_{PP} (ngoMqgibéèlo)_{PP}
 SC-YA-plough on-Saturday
 ‘They plough on Saturday.’
 - (b) (bá-zaku-liima)_{PP} (ngezáándla)_{PP}
 SC-Fut-plough by-hands
 ‘They are going to plough (not do something else) by hands.’
 - (c) (bá-ya-fudúúka)_{PP} (ngoLwésihláánu)_{PP}
 SC-YA-emigrate on-Friday
 ‘They do emigrate on Friday.’

For the analysis compare tableau 6 for adverbials under verb focus in Xhosa with tableau 5 for complement structures under verb focus in Chichewa.

Tableau 6 : Phrasing under narrow focus in Xhosa

	[V _{FOC} AdvP] _{VP}	ALIGN-FOC	WRAP-XP	ALIGN XP-R
a.	()P	*!		
b.	()P()P		*	

Although the winning candidate b violates the constraint WRAP-XP this violation is not fatal. Candidate a that satisfies this constraint is ruled out because it violates the higher-ranked constraint ALIGN-FOC.

In tableau 6 I expand Truckenbrodt’s (1995, 1999) theory in order to account for focus in adjunct structures. Assuming that adjuncts are part of the VP his analysis can account for the phrasing of adverbs in unmarked structure but also under focus. ALIGN-FOC is necessary in order to account for the prosodic structure as there is no independent evidence for the syntactic structure.

To sum up, in the preceding section I have presented Truckenbrodt’s approach to account for the Chichewa data. Since Selkirk’s end-based theory does not account for the data, he proposes an additional constraint WRAP-XP that can explain why in double object constructions no phrase boundary is inserted after the first complement. Nevertheless, if WRAP-XP is neutralized e.g. by means of focus the influence of ALIGN-XP, R shines through by phrasing individually each complement that follows a focused verb.

6.4 Data on subjects

A further refinement of the end-based theory was initiated by Selkirk & Shen (1990) in their work on Shanghai Chinese. It is also integrated into the format of Selkirk (1995). The basic idea is that lexical phrases (and their projections) and functional phrases (and their projections) behave differently in terms of relating syntactic and prosodic structure. Truckenbrodt (1999) calls this the Lexical Category Condition.

- (28) Lexical Category Condition (LCC, Truckenbrodt 1999)
 Constraints relating syntactic and prosodic categories apply to lexical syntactic elements and their projections, but not to functional elements and their projections.

The LCC predicts that alignment constraints will apply to lexical categories such as NPs, VPs, PPs, and APs, but not to functional syntactic categories such as DPs, IPs, or CPs. Consequently, the constraint ALIGN (XP, R; P, R) stated by Selkirk (1995) will exclude the insertion of prosodic phrase boundaries at edges of functional syntactic phrases.

For the discussion of Chichewa and Xhosa data in this section functional phrases at a higher syntactic level are of importance. Since it is controversial how many functional phrases above the VP exist in Bantu languages (see e.g. Spuy (1993) for IP and CP in Zulu; Bresnan (1994) for only one functional phrase), I assume the more minimal view and only consider one higher syntactic phrase, called IP here.

This section deals with problems encountered when expanding Truckenbrodt's analysis also to the phrasing of the subject in Xhosa. The data on subject phrasing reveal a basic difference between Xhosa and Chichewa. Whereas in Chichewa the subject is phrased separately from the verbal complex, the data on subject phrasing in Xhosa show that in the pragmatically unmarked case it is phrased together with the verbal complex. Examples are given in (29).

- (29) Phrasing of the subject in Xhosa (Jokweni 1995: 31, 50)
- (a) (abántwana ba-kháb' íbhóóla)_{PP}
 children SC-kick ball
 'Children are kicking the ball.'
- (b) (Lé ndod' i-béth' abántwaana)_{PP}
 this man SC-beat children
 'This man beats children.'

Under the premises that the subject and verb complex is unified in a functional phrase, the IP in standard generative grammar, the phrasing in (29) cannot be accounted for in neither the end-based theory nor the OT-based analysis. Selkirk predicts a pP-boundary after every XP, Truckenbrodt's analysis ends in the same result as the subject and verb complex are unified in a functional category, thus WRAP-XP does not apply and ALIGN-XP, R shows its influence. Only the parametric approach (Hale & Selkirk (1987) on Papago) can account for the phrasing in (29) given that the choice is such that only functional categories trigger the insertion of pP-boundaries (this will not be pursued here further).

I will have a closer look at the crucial point that makes the OT-based approach fail the analysis. As mentioned above, the premise that subject and verb phrase are unified in the functional phrase IP predicts a pP-boundary following the subject, as WRAP-XP does not apply and ALIGN-XP, R right-aligns every XP with a pP-boundary. There are two different ways to save the analysis for the Xhosa data; either to change the ranking in an OT-based account, or to revise the syntactic assumptions. Both possibilities will be presented and discussed in turn.

*6.4.1 Reranking of *P-Phrase*

A further deviation in the analysis by Truckenbrodt that distinguishes it from its antecedent by Selkirk is the assumption that alignment is not subject to language-specific parametrization (left-/ right-alignment), but to differences in ranking. Besides the two alignment constraints shown in (17) every language has an additional constraint that can render one or both of the alignment constraints inactive. This additional constraint is given in (30).

- (30) *P-PHRASE
Avoid p-phrases altogether.

Depending on the ranking of the constraint in (30) with respect to the alignment constraints the winning candidate will either show a uniform direction of alignment or no alignment at all. The data on phrasing in Xhosa presented so far suggest that under wide focus there is a single pP. Neither subjects nor objects trigger the insertion of a pP-boundary. An example that supports this generalization is given in (31). The sentence displays no sentence-medial pP-boundaries although it has a highly complicated and long structure.

- (31) Xhosa, Jokweni (1995: 98)
 (si-fún' úkú-xélel' abántwana úkubama-bá-ty' ínyaama)_{pp}
 SC-want to-tell children that HORT-SC-eat meat
 'We want to inform the children that they should eat meat.'

The observation that in pragmatically unmarked structures pPs are avoided can be expressed by ranking the constraint *P-PHRASE high in an analysis of the Xhosa data. The possibility to avoid pPs in a given language is implied in OT-based approaches (typologies by different rankings of the constraints). Such an analysis works well for pragmatically unmarked transitive and ditransitive structures, and naturally makes the right predictions concerning the phrasing of the subject.

The relevant data from Xhosa are given in (32).

- (32) Phrasing of the subject in Xhosa (Jokweni 1995: 31, 50)
- (a) (abántwana ba-kháb' íbhóola)_{pp}
 children SC-kick ball
 'Children are kicking the ball.'
- (b) (Lé ndod' i-béth' abántwaana)_{pp}
 this man SC-beat children
 'This man beats children.'

Tableau 7: Phrasing of the subject in Xhosa

[S [V XP] _{VP}] _{IP}	*P-PHRASE	WRAP-XP	ALIGN XP-R
a. ()P	*		**
b. ()P()P	**!		*

Both candidates a and b satisfy the constraint WRAP-XP since WRAP-XP only requires lexical XPs to be phrased into one phonological phrase. Therefore it requires the VP in the structures in (32) to be phrased together. The subject, however, constitutes a functional phrase which does not fall in the scope of WRAP-XP. On the contrary, ALIGN XP-R becomes relevant here which requires every XP to be phrased separately. Candidate a violates this constraint once more than candidate b, exactly because the subject is not phrased separately. Therefore, *P-PHRASE must be ranked above ALIGN XP-R in order to assure candidate a as the winning candidate.

However, the analysis proposed in tableau 7 based on the phrasing of the subject runs into problems when faced with the data on verb focus in ditransitive constructions, as will be shown in tableau 8. To remind the reader of the relevant structures an example is repeated in (33).

- (33) Xhosa, Jokweni (1995)
(ndi-ya-m-phékeel')_{pP} (úmaam')_{pP} (ínyaama)_{pP}
 SC-YA-OC1-cook.for mother 1 meat
 'I cook for her, mother, meat.'

Tableau 8: Phrasing of ditransitive structures in Xhosa

[S V _{FOC}] VP [XP] [XP]	ALIGN-FOC	*P-PHRASE	WRAP-XP	ALIGN XP-R
a. ()P	*!	*		**
b. ● [*] ()P()P		**		*
c. ()P()P()P		***!		

Tableau 8 shows that ranking *P-PHRASE high but still below ALIGN-FOC predicts pP-boundaries induced by focus but no other pP-boundaries. It correctly excludes candidate a where no focus-induced boundary occurs. Under this ranking, however, the observable additional pP-boundary following the first dislocated complement remains unexplained, as in candidate c which is the truthful winner. The wrong candidate is chosen as the winner (indicated by the bomb).

Turning the constraint *P-PHRASE into a non-gradient (=categorical) constraint (McCarthy 2003) does not save the analysis. Non-gradedness of a constraint means that it can only be violated once. Therefore in tableau 9 every candidate having more than one pP violates the constraint *P-PHRASE only once, independent of the number of pPs it actually displays.

Tableau 9: Phrasing of ditransitive structures in Xhosa

S V _{FOC}] VP [XP] [XP]	ALIGN-FOC	*P-PHRASE	WRAP-XP	ALIGN XP-R
a. ()P	*!	*		***
b. ()P()P		*		**!
c. ()P()P()P		*		* !
d. ● [*] ()P()P()P()P		*		

Although candidate b is ruled out in the analysis in tableau 8 because it crucially violates ALIGN-XP,R, the right output c is ruled out too. Candidate d satisfies the constraints best when phrasing the subject separately. Though the necessary data are missing, the prediction based on the data available is, that this is not the case in Xhosa. The subject is only phrased separately when in focus (see section 6.5.).

Dominance of the constraint *P-PHRASE is consequently not a possible solution as it only allows pP-boundaries induced by focus. The data from Xhosa, however, show evidence that pP-boundaries show up for syntactic reasons. Therefore, the next possibility that is going to be discussed is assuming a change in syntactic structure.

6.4.2 Subject-in-VP- hypothesis

Data on subject phrasing that are parallel to Xhosa have been discussed by Hale & Selkirk (1987) for Papago. The solution offered by Hale & Selkirk (1987) is that in Papago, subject and verb are both constituents of the VP. Their parametric approach to phrasing provides the possibility for a language to set the phrasing parameters in such a way that either functional or lexical phrases trigger the insertion of pP-boundaries. For Papago they assume lexical categories to trigger the insertion of pP-boundaries. Consequently, the VP would trigger a pP-boundary. Truckenbrodt (1995) adopts this analysis when discussing the Papago data in his framework. As subject and objects all belong to VP, WRAP-XP applies and favours wrapping the whole complex into one pP. This is exactly what happens in the Xhosa data. Why not accept such an analysis for Xhosa?

Whereas Papago and Chichewa are unrelated languages, Chichewa and Xhosa belong to the same language family, namely Bantu languages. There might be independent evidence to assume the subject-in-VP hypothesis for Papago. For Chichewa, Truckenbrodt (1999:227) assumes at least such a complex structure of the VP as proposed by Larson (1988). This means that the subject is base-generated in Spec-V position where it gets its theta-role assigned by the verb as it is within the minimal m-command. Here, the subject is internal to VP. But under subject-verb agreement the subject is moved out of the VP to the next higher syntactic phrase, which is a functional phrase.

To sum up this section, the OT-based analysis makes the right predictions when assuming the subject to be in VP-internal position. However, as long as there is no independent evidence for the syntactic structure regarding the subject in Xhosa it seems arbitrary to determine the syntax of the subject based on phrasing alone.

6.5 Subject and focus

In the preceding subsections I have shown that the phrasing of the subject in constructions displaying wide focus cannot be accounted for yet. However, under narrow focus on the subject the same generalizations apply as with verbs

in focus. A pP-boundary is inserted after the focused constituent. The data are given in (34).

- (34) Xhosa, Jokweni (1995: 27, 32)
- (a) (abántwana ba-kháb' íbhóola)_{pP}
 children SC-kick ball
 'Children are kicking the ball.' (What is going on?)
- (b) (**abántwaana**)_{pP} (ba-kháb' íbhóola)_{pP}
 children 2 SC-kick ball
 'The children kick the ball.'

The sentences in (34) show a near minimal pair that differs only in focus and pP-boundaries. Note that there is no evidence to assume a cleft-like construction in example (34b) although cleft constructions are often used in Bantu languages to express subject focus. (34b) supports the view that also in the case of subject focus a boundary is inserted after the focused constituent.

Tableau 10 shows that the effect of focus on the phrasing of subjects can be accounted for by the analysis presented so far.

Tableau 10: Phrasing under subject focus in Xhosa

[[S _{FOC} [V AdvP] _{VP}] _{IP}	ALIGN-FOC	WRAP-XP	ALIGN XP-R
a ()P	*!		*
b. $\left[\begin{array}{c} \text{ } \\ \text{ } \end{array} \right] P \left(\begin{array}{c} \text{ } \\ \text{ } \end{array} \right) P$			

Though candidate a is preferred under wide focus, candidate b is the winner for narrow focus on the subject. This can be accounted for by ranking the constraint ALIGN-FOC high which is crucially violated by candidate a.

This section has shown the problems that arise in connection with the phrasing of the subject. Further research and more data are needed in order to account for the relation of syntax and prosody with regard to the subject.

7 Summary

In this paper I have shown how syntax and focus interact in deriving the phonological phrasing of utterances in Xhosa. Although the influence of syntax on phrasing is uncontroversial, a purely syntactic analysis cannot account for the data. I have shown how one could argue for a systematic influence of focus. The influence is such that it inserts a pP-boundary after the focused constituent

unless there is one triggered by syntax already. This generalization can account for the variation found in the phrasing of adverbials.

I have formalized the findings in an OT-based framework in order to theoretically account for the impact focus-related constraints have on phrasing as derived by syntax-related constraints. In order to account for the variation found in the phrasing of adverbials I have transferred Truckenbrodt's analysis (1995, 1999) to the phrasing of adjuncts. Furthermore, I have pointed out problems in connection with the phrasing of the subject.

References

- Bresnan, J. (1994) Locative inversion and the architecture of universal grammar. *Language* 70 (1): 72-131.
- Bresnan, J. & S. A. Mchombo (1987) Topic, pronoun, and agreement in Chichewa. *Language* 63 (4): 741-782.
- McCarthy, J. (2003) OT-constraints are categorical. *Phonology* 20: 75-138.
- Chen, M. Y. (1987) The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4: 109-149.
- Chomsky, N. (1971) Deep structure, surface structure, and semantic interpretation. *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*. D.D. Steinberg & L. A. Jakobovits (eds.) Cambridge, Cambridge University Press: 183-216.
- Güldemann, T. (1996) *Verbalmorphologie und Nebenprädikation im Bantu*. Bochum, Universitätsverlag Dr. N. Brockmeyer.
- Haegemann, L. (1991) *Introduction to Government & Binding Theory*, Blackwell.
- Hale, K. & E. Selkirk (1987) Government and tonal phrasing in Papago. *Phonology Yearbook* 4: 151-183.
- Jackendoff, R. (1972) *Semantic Interpretation in Generative Grammar*. Cambridge (Mass.), Massachusetts Institute of Technology.
- Jokweni, M. W. (1995) Aspects of Isixhosa Phrasal Phonology. Doctoral thesis, Urbana-Champaign, University of Illinois.
- Kanerva, J. (1990) *Focus and Phrasing in Chichewa Phonology*. New York: Garland Publishing.
- Kisseberth, C. W. & M. I. Abasheikh (1974) Vowel length in Chi-Mwi:ni- A Case Study of the Role of Grammar in Phonology. *Papers from the Parasession on Natural Phonology*. A. Bruck et al. (eds.), CLS: 193-209.
- Kosch, I. M. (1988) Imperfect tense -a' of Northern Sotho revisited. *South African Journal of African Languages* 8 (1): 1-6.
- Larson, R. (1988) On the double object construction. *Linguistic Inquiry* 19: 335-391.
- McCarthy, J. & A. Prince (1993) Generalized Alignment. *Yearbook of Morphology 1993*. G. Booij & J. v. Marle (eds.). Dordrecht, Kluwer: 79-153.
- McHugh, B. D. (1990) The Phrasal Cycle in Kinvunjo Chaga Tonology. In: *The Phonology-Syntax Connection*. S. Inkelas & D. Zec (eds.). Stanford, CSLI.
- Morimoto, Y. (2000) Discourse Configurationality in Bantu Morphosyntax, Doctoral thesis, Stanford University.
- Nespor, M. and I. Vogel (1986) *Prosodic Phonology*. Dordrecht, Foris Publications.

- Selkirk, E. (1980) *On Prosodic Structure and its Relation to Syntactic Structure*. Bloomington, Indiana: Indiana University Linguistics Club.
- Selkirk, E. (1986) On Derived Domains in Sentence Phonology. *Phonology Yearbook* 3: 371-405.
- Selkirk, E. (1995) The prosodic structure of function words. In: J. Beckman, L. Wash Dickey & S. Urbanczyk (eds.). Amherst, GLSA, University of Massachusetts. *Papers in Optimality Theory*: 439-469.
- Selkirk, E. & Shen, T. (1990) Prosodic Domains in Shanghai Chinese. In: *The Phonology-Syntax Connection*. S. Inkelas & D. Zec (eds.). Stanford, CSLI: 313- 337.
- Spuy, A. v. d. (1993) Dislocated noun phrases in Nguni. *Lingua* 90: 335-355.
- Truckenbrodt, H. (1995) Phonological phrases: Their relation to syntax, focus, and prominence. Doctoral thesis, Massachusetts Institute of Technology.
- Truckenbrodt, H. (1999) On the Relation between Syntactic Phrases and Phonological Phrases. *Linguistic Inquiry* 30 (2): 219-255.
- Zerbian, S. (2003) Morphological Marking of Phonological Phrase Boundaries in Xhosa. *Proceedings of Prosodic Interfaces*, Nantes: 209-214.

What African Languages Tell Us About Accent Typology*

Laura J. Downing

ZAS Berlin

The goal of this paper is to survey the accent systems of the indigenous languages of Africa. Although roughly one third of the world's languages are spoken in Africa, this continent has tended to be underrepresented in earlier stress and accent typology surveys, like Hyman (1977). This one aims to fill that gap. Two main contributions to the typology of accent are made by this study of African languages. First, it confirms Hyman's (1977) earlier finding that (stem-)initial and penult are the most common positions, cross-linguistically, to be assigned main stress. Further, it shows that not only stress but also tone and segment distribution can define prominence asymmetries which are best analyzed in terms of accent.

1 Introduction

According to Heine & Nurse (2000), citing the *Ethnologue*, over 2,000 languages – roughly one-third of the world's languages – are currently spoken in Africa by some 750 million people. (In this paper, no discussion is included of recently introduced European colonial languages, like English, French, Portuguese or Afrikaans.) African languages tend to be underrepresented in surveys of stress and accent systems like Hyman (1977), however, as scholars seeking to make generalizations about accent in African languages must tackle a number of problems. First, one has to confront the sheer number and variety of

* Many people generously shared their expertise with me as I wrote this paper, providing information and feedback, sharing their own unpublished work and pointing me to the work of others. I would especially like to thank Akin Akinlabi, Lee Bickmore, Simon Donnelly, Knut Felberg, Leoma Gilley, Rob Goedemans, Tom Güldemann, Silke Hamann, John Harris, Omar Ka, Michael Kenstowicz, Will Leben, Brian McHugh, Amanda Miller-Ockhuizen, Paul Newman, David Odden, Bert Remijsen, Sharon Rose, Eno-Abasi Urua and Valentin Vydrine. This paper could not have been written without the research assistance of Sabine Zerbian, who guided me through the Berlin library system and spent many hours helping me locate much of the information surveyed here.

African languages. These 2,000+ languages have been classified into four large families: Niger-Congo (1,436 languages), Afro-Asiatic (371), Nilo-Saharan (196) and Khoisan (35). Further, one member of the Austronesian family, Malagasy, is spoken on the island of Madagascar. As each of the four major families includes a large number of languages further grouped into diverse subfamilies, it can be difficult to make generalizations about accent even within a single subfamily.

Another problem in surveying African accent systems is that, as Heine & Nurse (2000) point out, most African languages remain poorly documented. Of the four large families, Afro-Asiatic is the best analyzed, as it has the longest history of scholarship carried out by the largest number of scholars. The majority of the languages in the other families are not as well described and analyzed due to the relatively small number of scholars working on an extremely large number of languages for a much shorter period of time. As a result, for most African languages, the only documentation often consists of a grammatical sketch (often many years old using idiosyncratic transcription systems), an analysis of part of the language, or a couple of short journal articles. Accent (along with many other grammatical features) is not always mentioned or reliably described and documented in these sources. As a result, this survey, like all comparative work on African languages, comes with the caveat that the generalizations are based on the information available which might be so limited as to be unrepresentative in some cases.

A final problem in surveying African accent systems arises from both descriptive and theoretical controversies in defining metrical accent. As background to discussing these controversies, it is important to note that the languages of sub-Saharan Africa are overwhelmingly tonal (Welmers 1973). Of the non-tonal languages, many of the best documented – the African Arabic dialects, Berber, North Atlantic (Niger-Congo) languages – uncontroversially have stress accent. Many of the uncontroversially tonal languages, notably many Bantu languages, also uncontroversially have non-tonal stress accent. (All of these stress systems are discussed below.) However, for many African tonal languages, scholars disagree about the role of accent. As Welmers (1973) notes, tone has been notoriously underdescribed for African languages as it is often considered “too hard to hear” by non-native speakers of tone languages. (Most descriptions of African languages have been written by native speakers of Indo-European languages rather than by native speakers of the languages themselves.) One easily finds examples of languages which were described as accentual or non-tonal by earlier scholars which are currently accepted as having tone (e.g., Somali (Hyman 1981) or Jita (Downing 1996)). This means that not all descriptions of accent are reliable.

To be charitable, it is, in fact, often not straightforward to decide whether a particular pitch system is best described as tonal or accentual. Pitch (especially pitch change) is a principal phonetic correlate of stress, with duration and amplitude or intensity acting as the main secondary cues (Hyman 1977, Odden 1999, and references cited therein). Since raised pitch, especially when it coincides with vowel length, makes a syllable perceptually more prominent, it can often require detailed phonetic and phonological analysis to disentangle whether pitch is playing a more stress-like or a more tone-like role in a particular language. (See Odden 1999 for detailed discussion of this point.) Unfortunately, few phonetic studies of the correlates of stress in African tone languages have been carried out. Most authors reporting on stress do not even make explicit what impressionistic phonetic correlates of stress guide their descriptions. As a result of these problems, it is not uncommon to find conflicting opinions among current scholars about whether perceived prominence in particular languages is best described as stress-accent. Some of these controversies are highlighted in the discussion below.

Another source of controversy arises in languages where the distribution of tone or other features is restricted in such a way as to create prominence asymmetries within morphemes reminiscent of prominence realization in stress-accent systems. What is at issue in such cases is whether accent/metrical structure should be appealed to formally to account for these asymmetries. I review here the principal arguments for accentual structure in conditioning the realization of tone and the distribution of segmental features.

Since the work of McCawley (1970, 1978), it has been recognized that some tone systems might better be characterized as pitch-accent systems because tone realization shares properties with stress-accent realization. First, as work like Hayes (1995), Hyman (1977, 1978), Odden (1988, 1999) argue, a typological characteristic of stress is that it is culminative: only one (main) stress may occur within a particular domain and every word must be assigned stress. Similarly, tonal systems are said to be accent-like if they have properties that lead to H tone culminativity, such as:

- lexical tonal contrasts are defined on morphemes (not moras or syllables), so that at most one H tone (or tone melody) is found per morpheme;
- H tone is assigned predictably to every word/lexical morpheme;
- H tones on adjacent syllables or morphemes are eliminated (this is analogous to stress reduction to resolve clash, a metrical process motivated by culminativity).

Secondly, as Hyman (1977) argues, an important function of stress is to demarcate major morpheme edges (stem or word). This explains why main stress in the languages of the world is overwhelmingly located at a word or stem edge: initial, penult or final position. Similarly, tone systems are said to be

accent-like if H tones are restricted to occur in these same positions: stem/word-initial, stem/word/phrase-penult, stem/word/phrase-final (see, e.g., Odden 1988). Finally, tone has been argued to be accent-like if tone and stress interact (van der Hulst & Smith 1988, Downing 1996). There are two main types of tone-stress interaction found in African languages. The position of High tone realization is conditioned by the position of the stressed syllable, or the position of main stress is conditioned by (High) tone.

It should be pointed out that while there is agreement among tone scholars that these tonal properties are analogous to stress-accent, there is widespread disagreement about whether accent or metrical constituency is required as a formal device to distinguish pitch-accent systems from other tone systems. This problem is, in fact, a recurring theme in critical surveys of tone and accent, especially for Bantu languages (see, e.g., Hyman (1989), Downing (1996), Odden (1988, 1995, 1999) for discussion). To avoid misunderstandings, then, when languages are characterized as having pitch-accent systems in the sections that follow, this means that their tone systems clearly have one or more of the properties described above: culminativity, positional restrictions or tone-stress interactions. These properties are considered accentual for the purposes of this paper, as they lead to tonal prominence asymmetries within morpho-syntactic domains, however they are to be formally analyzed.

Tone is not the only feature with asymmetrical distribution in African languages. In Khoisan languages as well as in a number of West African Niger-Congo languages, roots tend to be maximally bimoraic. (Detailed discussion and references are found below.) Although the root contains two consonant positions, in these languages one finds that the full range of consonant contrasts can typically only be found in C1 position. In C2 position, contrasts are often reduced drastically, so that only sonorants or some other very restricted subset of consonants can occur. Vowel features show similar asymmetries in some languages, with the full range of contrasts found only in the initial syllable, while contrasts are reduced in the other syllable(s). As Harris (2004) argues particularly forcefully, these asymmetries in the distribution of segmental contrasts within a bimoraic (or bisyllabic) string cannot easily be motivated by appealing to syllable structure. Word-final and prevocalic, non-initial (C2) positions do not form a natural class in terms of syllable structure or features that distinguishes them systematically from C1. However, if the bimoraic string is a foot, C2 is in foot-medial position and C1 is in foot-initial position. In languages that have stress-accent, like English or Danish, there is a strong tendency for full consonantal contrast to be found only in foot-initial position, with contrasts reduced or neutralized in foot-medial position. Likewise, it is common for vocalic contrasts to be realized in the strong syllable of the foot and reduced in the weak syllable (see, e.g., Hayes 1995). Even though African

languages with contrast distribution asymmetries do not have the usual stress-accent motivating the asymmetries, it is plausible to propose that, universally, these kinds of distributional restrictions are to be accounted for in terms of accent, as they create prominence asymmetries akin to stress accent.

In the remainder of the paper, the accentual systems of African languages are briefly surveyed. The paper is organized by language family: AfroAsiatic, Nilo-Saharan, Khoisan, Malagasy and Niger-Congo. For each family, the emphasis will be not only on examples of stress accent, but also on pitch-accent and accent-like segment distributional asymmetries. It is this diversity of prominence asymmetries that in fact make African languages particularly interesting for research on the typology of accent.

2 Afro-Asiatic

The Afro-Asiatic languages are spoken in Northern Africa, in north central Africa as far south as northern Nigeria and in eastern Africa through the Horn. According to Hayward (2000), the modern languages are divided into 5 major branches: Semitic, Berber, Cushitic, Omotic and Chadic. This section will discuss accent patterns in each of these branches in turn.

2.1 Semitic

The modern Semitic languages spoken in Africa belong to two major subgroups, Central Semitic (which includes the African dialects of Arabic) and Southern Semitic. Arabic is one of the most widely spoken languages in Africa. It is a major language of Egypt, Libya, Algeria, Tunisia, Morocco and Sudan, and a minority language in Sub-Saharan African countries like Chad and Nigeria. According to the *Ethnologue* (2000), over 110,000,000 people speak Arabic as a first or second language in these countries. Southern Semitic languages are spoken mainly in Eritrea and Ethiopia. Major languages in this group include Amharic, Gurage, Tigrinya and Tigre.

2.1.1 African Arabic dialects (Central Semitic)

In most African Arabic dialects, stress falls on one of the last three syllables in a word, with syllable weight playing an important role in determining which of the last three syllables is stressed (Fischer & Jastrow 1980; Angoujard 1990; Mitchell 1993: chapter 4). Often the rightmost heavy syllable within this window is stressed, with final syllables only counting as heavy if they are “superheavy”: CVVC or CVCC. If there are no heavy syllables, then stress falls on the penult or the antepenult, depending on the other parameters of stress

assignment in the particular dialect. African Arabic dialects fitting this rightmost, quantity-sensitive pattern include: Tunis Arabic (Angoujard 1990), Egyptian Arabic (Harrell 1957, Hayes 1990, McCarthy 1979; Mitchell 1993), Moroccan Arabic (Kenstowicz & Kisseberth 1979, citing Harrell 1962; Mitchell 1993), Chadian Arabic (Abu-Absi 1995) and Nigerian Arabic (Owens 1993). Morphological and lexical factors also play a role in conditioning stress in most dialects. For example, the root syllable may be stressed in preference to a non-root syllable of equal weight; or particular morphemes are always stressed; or some borrowed words keep the stress of the original language. Some of these prosodic and morphological conditions are illustrated in the Egyptian Arabic data below:

- (1) Egyptian Arabic stress (Harrell 1957, pp 15-16; Mitchell 1993: 197-198)
 - (a) Antepenult of 3-syllable word is stressed if all syllables are light
búxala ‘misers’ kátaba ‘clerks’
 - (b) Penult of 3-syllable word is stressed if it is heavy
ʔaxáttu ‘I took it (m.)’ maxtába ‘library’
 - (c) Final syllable is stressed if it is super-heavy
fanagáin ‘cups’ fihímt ‘I/you s.m. understood’
 - (d) Penult of 3-syllable word is stressed if the initial syllable is a prefixal particle
bi+ʔálam ‘with a pencil’

Even though all the African Arabic dialects share the general parameters of stress assignment outlined above, there is considerable variation in the details of the stress systems among the dialects, as the reader will quickly discover in consulting the works cited.

2.1.2 *Amharic and Tigrinya (Southern Semitic)*

In contrast to Arabic, where word stress-accent is a salient, well-described feature of all dialects, one finds much vaguer descriptions of stress-accent in Southern Semitic languages. For example, in Amharic, a major language of Ethiopia and the most widely spoken Southern Semitic language, stress placement is described in terms of the following tendencies (Leslau 1995; Richter 1987). In general, the penultimate syllable is stressed (arängwáde ‘green’). However, in three-syllable words, the first syllable is often stressed rather than the penult if its vowel is more sonorous than the penult vowel (sámuna ‘soap’). Vowels preceding geminate consonants are also likely to be stressed (amäs’aggänä ‘he thanked’). And stress is also more likely to fall on root vowels than affixal vowels. According to Cowley et al. (1976), certain

phrasal positions also attract stress, for example, the vowel immediately preceding the copula *nāw* 'he is' is usually stressed (*tāmarí nāw* 'he is a student'). These authors all suggest that the main correlates of stress are vowel quality and syllable duration, not pitch and loudness, which might be the reason the position of stress-accent is hard to determine.

Similarly, Bender et al. (1976) state that stress-accent is not particularly salient in Tigrinya, another major Southern Semitic language spoken in Ethiopia and Eritrea. They suggest, however, that stress is typically word-final, and that phrasal stress plays the most important role in determining the rhythm of the sentence.

Recent sketches of individual Southern Semitic languages in Hetzron (1997) confirm that stress in these languages is poorly understood.

2.2 Berber

The Berber languages, spoken over a wide area of northern Africa and Saharan Africa, are diverse enough to be broken down into four subfamilies (Hayward 2000). Of these, the languages of Morocco and Algeria appear to be the best described.

Tamazight and Tachelhit (ShilHa) are two major Berber dialects spoken in southern Morocco and Algeria. In both of these dialects, stress is described as falling on the last vowel in the word (Abdel-Massih 1971; Applegate 1958). In Tachelhit, affixes preceding the stem which contain a vowel are also said to be stressed (Applegate 1958). The data in (2) illustrates the final stress pattern of Tamazight:

- (2) Tamazight Berber stress (Abdel-Massih 1971, pp. 17-18; capitalized consonants are emphatics)
- (a) ndawá 'we cured'
 - (b) tfafád 'you (sg) woke up'
 - (c) dayt:HaDáR 'he is present'

2.3 Cushitic and Omotic

The Cushitic and Omotic languages are spoken mainly in Ethiopia, Eritrea and Somalia. Major Cushitic languages include Somali, Sidamo, Afar, Oromo, Arbore and Dahalo. Major Omotic languages include Dime and Wolaitta. Omotic was formerly considered a branch of Cushitic, but now is recognized as a distinct subgrouping by most Afroasiatic scholars (Hayward 2000, Amha 1996). Similarities in their accentual systems, however, justify discussing them together.

Almost all of the Cushitic (and Omotic) languages are tonal, with tone placement conditioned morpho-syntactically (Sasse 1981). Many of these tone systems have the sorts of metrical restrictions on tone realization discussed in Section 1 which have led recent authors to characterize them as pitch-accent systems. For example, tone is subject to culminativity in Somali (Hyman 1981) and Wolaitta (Amha 1996). Only one High tone is allowed per word, and in compounds only the first member retains its High tone. (However, according to Sasse (1981), in some Cushitic languages – see e.g. Iraqw (4), below - not all words must bear High tone, which would be an unusual property in a stress system. And in others, more than one High tone per word can be found.) Also, tone is subject to a positional restriction, occurring generally on the penult or final syllable of the word or stem in languages like Somali, Wolaitta, and Beja (Bedawi) (Hudson 1973). (However, it is not phonologically predictable where the High tone will fall, though there are morphologically predictable patterns.) The Wolaitta data in (3) illustrates these points:

- (3) Wolaitta pitch-accent (Amha 1996): Only one H per word, on the penult or final of word or stem

zaré	‘lizard’	
záre	‘relative’	
zígir-etta	‘to gossip’	/zígir + ettá/
k’aar-ettá	‘to uncover’	/k’aar + ettá/
hayttá	‘spicy coffee made from coffee	(hayttá ‘leaf’ + tukké ‘coffee’)
tukke	leaves’	

Iraqw, a South Cushitic language spoken in Tanzania, is described as having a very different sort of accent system, with quantity-sensitive stress, as well as tone (Mous 1993). High tones pattern with long vowels in attracting stress. As shown by the data in (4), below, the penult is stressed if it has a long vowel. If the penult has a short vowel and the final syllable has a High tone, then the final syllable is stressed. Otherwise, the first syllable is stressed. (Not all words in this language are assigned High tone.)

- (4) Iraqw stress (Mous 1993); accents indicate tone; the stressed vowel is underlined

(a)	ba?e <u>eso</u>	‘bushbucks’
(b)	h <u>l</u> ooró	‘grasshopper’
(c)	mugúl	‘collarbone’
(d)	w <u>a</u> witmo	‘king’
(e)	b <u>a</u> sa	‘south’

As we can see from comparing Wolaitta (3) with Iraqw (4), there is considerable variation in the accent systems found in the Cushitic and Omotic languages. One cannot easily generalize about shared parameters of accent assignment or even accent realization in these languages, like we could in discussing Arabic (section 2.1.)

2.4 *Chadic*

There are approximately 140 Chadic languages, spoken mainly in northern Nigeria, Chad, Cameroon, Central African Republic and Niger. Hausa is the most widely spoken, with an estimated 38,000,000 first and second language speakers (*Ethnologue* 2000). In fact, when second language speakers are included, Hausa ranks as one of the largest African languages excluding Arabic (Hayward 2000).

Chadic languages are all tonal, and it remains controversial whether stress is also a feature of these languages. As Wolff (1993) points out, even though stress has been reported for Hausa, for example, this is likely due to the fact that the correlation between vowel-length and High tone on particular syllables often strikes European ears as stress on that syllable.¹ (As noted in Section 1, since pitch change and duration are the important phonetic correlates of stress, syllables with these properties are more prominent.) Phonetic work by Dresel (1977) confirms that there is, in fact, no accent independent from High tone in Hausa.

However, Kwami, a Chadic language spoken in Nigeria, is described by Leger (1994) as having stress in words that lack vowel length and High tone. According to Leger, while syllables with a long vowel are always stressed and initial High toned syllables or closed syllables also attract stress (when there is no long vowel), stress is also heard on the final syllable of words lacking High tones or long vowels. These points are illustrated by the data below (accents indicate tones; the stressed syllable is underlined):

- (5) Kwami stress (Leger 1994, pp 91-92)
- | | |
|--------------------|------------------|
| háakù | ‘yawn (noun)’ |
| gàringá <u>alè</u> | ‘bicycle, wheel’ |
| ká <u>bá</u> | ‘inside’ |
| kòmkáy | ‘pluck’ |
| kùrè | ‘fence’ |

¹ Thanks to Paul Newman (p.c.) for bringing this to my attention. It is telling that Newman’s (2000) truly encyclopedic grammar of Hausa contains no discussion of stress, although the phonology, tone and intonation of Hausa receive detailed coverage.

Clearly more phonetic studies of the correlates of stress in Chadic languages are necessary to disentangle this controversy.

3 Nilo-Saharan Languages

Nilo-Saharan remains the most diverse and controversial of the four African language families (see Bender 2000 for an up to date discussion). It contains a dozen sub-families whose speakers are scattered across the northern Democratic Republic of the Congo (DRC), southern Sudan, Uganda and Kenya.

Many of the languages have not been described in detail, making it difficult to ascertain what role, if any, stress-accent plays in this language family. Most of those which have been described are said to have tone systems, with no mention of stress. However, Dholuo (E. Nilotic) is described by Tucker (1993) as having stress as well as tone. Stress is assigned to the stem-initial vowel. (That is, stress placement is determined by morphological, not phonological, position.) When a word is in phrase-final position, the stressed vowel is noticeably lengthened, at least in some dialects. (Tucker unfortunately does not describe other correlates of stress, though his discussion implies length is not the only one.) Since stress is morphologically determined, words of identical lengths but different morphological structures will be stressed on different syllables, as shown in (6a). And adding suffixes does not change the position of stress away from the stem vowel, as shown in (6b).

(6) Dholuo stress (Tucker 1993, p 18); stressed vowels are underlined; accents indicate high tone

- | | | | | | |
|-----|----------|--------------|-----|------------------|--------------|
| (a) | ra+hî | ‘cobra’ | vs | r <u>awo</u> | ‘hippo’ |
| (b) | ra+hî+nì | ‘this cobra’ | vs. | r <u>awo</u> +ni | ‘this hippo’ |

Similarly Lango (W. Nilotic), is described by Noonan (1992) as having stress as well as contrastive tone. As in Dholuo, stress (realized as increased amplitude and length) is assigned to the root-initial vowel, as shown in (7a). (Only borrowed words have stress on other root syllables.) And in compounds, the rightmost word of the compound receives primary stress, as shown in (7b).

(7) Lango stress (Noonan 1992, pp 42-43); stressed vowels are underlined; accents indicate high tone

- | | | | | | |
|-----|------------|-------------------|-----|-------|------------|
| (a) | à+bòlò | ‘plaintain’ | vs. | màkâc | ‘scissors’ |
| (b) | píg á+bólò | ‘plaintain juice’ | | | |

Dinka and Shilluk, two other Western Nilotic languages, are described by Gilley (2003) as also having contrastive tone and grammatical (morphologically- or syntactically-conditioned) stress. Surprisingly, even though vowel length is contrastive in these two languages, stressed vowels are said by Gilley (2003) to be shorter than unstressed vowels (recall from section 1 that lengthening is the usual correlate of stress) but have greater intensity than unstressed vowels.

While these Nilotic languages have morphemic stress, some Nilo-Saharan languages apparently have positional stress. For example, both Hyman (1977) and Ruhlen (1975, 1976) list Mangbetu (Central Sudanic) as having dominant penultimate stress (as well as tone). Finally, Nicolai & Zima (1997) and Heath (1999) note that the Western dialects of Songhay have a stress-accent system, rather than the tonal one found in Proto-Songhay and the other modern Songhay dialects. Unfortunately, no description of the stress system is provided in these works.

4 Khoisan languages²

The Khoisan languages represent the smallest of the four large African language groups. According to Güldemann & Vossen (2000), it is estimated that there might have been as many as 100 Khoisan languages in the past, spoken over much of Southern Africa. Today, barely 30 Khoisan languages survive, with the largest communities concentrated in Namibia and Botswana. Very few of these languages have been described in any detail, and it seems likely that many of the remaining Khoisan languages will become extinct before they can be carefully documented.

All the Khoisan languages are described as having tone, and I could find no discussion of stress in the existing Khoisan literature. However, recent work points to other evidence for the importance of foot structure in particular Khoisan languages. One sort of evidence comes from the skewed distribution of consonants in initial vs. medial and final position within stems (which are minimally bimoraic). Traill (1985) reports that most of the 116 contrastive consonants of !Xóõ are only found morpheme-initially. Morpheme-medially (intervocally), only 6 consonants occur, all arguably sonorants. Morpheme-finally, only 3 consonants, all nasals, occur. As noted in Section 1, intervocalic and final consonant-position do not form a syllabic natural class. Instead, the restrictions found on consonants in these positions are best explained by proposing they are foot-bound. Foot-initial position licenses all consonants; non-

² Thanks to Amanda Miller-Ockhuizen for helpful discussion of the points raised in this section. See Güldemann & Vossen (2000) for an up to date overview of the sociolinguistic status, classification and grammatical structure of Khoisan.

initial positions license only sonorants. Miller-Ockhuizen (2001b) shows that a similar asymmetry in consonant distribution is also found in Ju|'hoansi, and that marked vocalic phonation is only found in foot (or root) initial syllables.

Another piece of evidence for foot structure comes from minimality requirements on roots. In both !Xóǀ (Traill 1985) and Ju|'hoansi (Miller-Ockhuizen 2001a, b) roots are minimally bimoraic. As Miller-Ockhuizen (2001a, b) shows, roots described as CV by other authors actually contain significantly lengthened vowels, so that they are similar in overall duration to CVCV roots. This requirement is best explained if roots must minimally contain a single bimoraic foot. Finally, in both !Xóǀ (Traill 1985) and Ju|'hoansi (Miller-Ockhuizen 2001a, b), the possible tone patterns for roots are very limited. Traill (1985) argues these can best be explained if the entire root morpheme (or root foot), not each mora of the root, is associated with a root tone melody. As noted in Section 1, above, it is considered diagnostic of pitch-accent languages for tone melodies to take a morpheme as their domain rather than the syllable. (However, see Miller-Ockhuizen (1998) and Haacke (1999) for alternative analyses of these restricted tonal patterns.)

Traill (1985) argues that adaptations of borrowings into !Xóǀ provide evidence that the consonant distribution and minimality requirements on stems (roots) are not static lexical generalizations but rather productive processes. For example, Afrikaans lorrie 'lorry' and koppie 'cup' both have short initial vowels in the original language, but lorrie is borrowed as nòli (with a short vowel) while koppie is borrowed as kóopi (with a long vowel). The explanation for this is that lorrie has a medial sonorant, so fits the consonant distribution restrictions on root-second consonants. However, koppie has a medial stop, so the initial vowel is lengthened to make the initial syllable a possible bimoraic root. The second syllable can now be interpreted as a suffix licensing a morpheme-initial obstruent.

The restrictions on the distribution of consonants and marked vocalic phonation suggest that it is the morpheme-initial (or foot-initial) syllable that counts as strongest in Khoisan languages. However, there are also restrictions on the distribution of vocalic features that point to the second syllable as stronger, at least in Ju|'hoansi. In this language, some reduced vowels only occur in the first syllable of bisyllabic roots, never in the second syllable (Miller-Ockhuizen, p.c.; see, too, Miller-Ockhuizen (1999) for a more detailed discussion of vowel distribution restrictions). Obviously, more work needs to be done on the phonology of these languages to reach a better understanding of how these conflicting foot-bound prominence asymmetries could be resolved.

5 Malagasy

Malagasy, spoken on the island of Madagascar, is the westernmost Austronesian language and the only Austronesian language spoken in Africa. According to Beaujard (1998) and Keenan & Polinsky (1998), main stress in Malagasy is always assigned to one of the last three syllables of a word. Main stress generally falls on the penult, as shown in (8a). However, lexical and morphological factors interfere with this basic stress pattern. Words ending in *-tra*, *-ka*, *-na*, and *-na* generally receive main stress on the antepenult. The minimal contrast between (8a) and (8b) shows that the antepenult stress pattern is lexically conditioned. The final syllable can also be stressed in certain lexical and morphological contexts. As shown in (8c), some roots ending in /e/ are stressed on the final syllable. Morphologically-conditioned final stress is illustrated by the contrast in (8d). In words of four or more syllables, secondary stress is assigned to every second syllable preceding the syllable with main stress, as shown in (8e):

- (8) Malagasy stress patterns (Beaujard 1998: 23; Keenan & Polinsky 1998: 570); vowels with main stress are underlined
- | | | | |
|-----|------------------------|---|---|
| (a) | Regular penult stress: | tan <u>ana</u> | ‘village’ |
| (b) | Antepenult stress: | ta <u>na</u> na
sar <u>i</u> taka | ‘hand’
‘disorder’ |
| (c) | Final stress 1: | ome <u>e</u> | ‘act of giving’ |
| (d) | Final stress 2: | mil <u>a</u> za
vs.
milaza <u>a</u> | ‘say, active’
‘say, passive’ |
| (e) | Secondary stress: | sòmar <u>i</u> taka
pàtal <u>o</u> ha | ‘preoccupied’
‘pants’ (<French ‘pantalon’) |

6 Niger-Congo³

The Niger-Congo language family is the largest in Africa – in fact, one of the largest in the world – containing 1436 languages grouped into some 20 subfamilies spoken by 360+ million speakers distributed over most of sub-Saharan Africa (Williamson & Blench 2000).⁴ The languages in this family are overwhelmingly tonal. Only the North Atlantic subfamily is generally non-tonal

³ The information about the distribution of the subfamilies and languages discussed in this section comes from Williamson & Blench (2000) and the *Ethnologue* (2000).

⁴ See, too articles in Bendor-Samuel (1989) for a detailed, current discussion of the classification and important grammatical features of the Niger-Congo languages.

(though individual languages in other subfamilies have been described as lacking tone). Several of the tone languages are described as also having stress-accent. In others, we find evidence for pitch-accent, as tone placement is subject to the metrical conditions, like culminativity and attraction to word-peripheral position, discussed in Section 1. And several show accent-like asymmetries in consonant distribution.

Since this is a large, diverse language family with many subfamilies, this section is organized by subfamily, beginning with Atlantic. Next, other West African subfamilies are discussed. (If a subfamily is not discussed, that is because no discussion of stress or other evidence for accent in languages of that subfamily was found.) Bantu languages are discussed last. Since this is the largest subfamily and the accentual properties of Bantu languages have been extensively studied, this lengthy subsection is further broken down by topic: stress-accent, accent-like distribution asymmetries and pitch-accent.

6.1 *Atlantic*

The Atlantic languages are spoken mainly along the North Atlantic coast of Africa, from the mouth of the Senegal River (and along the river to the interior) south to Liberia. Major languages in this group include Ffulde, Wolof, the Diola group and Serer.

Within the Atlantic subfamily, only most North Atlantic languages are non-tonal. One finds a surprising range of stress systems within the North Atlantic group. Sapir (1965) shows that Diola-Fogny, spoken in Gambia and Senegal, has morphologically-determined stress. The first syllable of a bound root receives primary stress. Some final clitics, like possessives, the definite article and the third person object particle, also receive primary stress, even when this leads to adjacent primary stresses. These points are illustrated below:

- (9) Diola-Fogny stress (Sapir 1965, p. 9); [‘ marks the edge of a bound root; stressed vowels are underlined
- (a) ε-[ja-mɛn ‘goat’
 - (b) ε-[ja-mɛn-ney ‘the goat’
 - (c) ε-[ja-mɛn-nɔm ‘my goat’
 - (d) pa-ni-[ka-tɔ ‘I will leave him’
 - (e) ni-tɛ-[tɛk ‘I hit’
 - (f) ni-[ka-nɔ ‘I dress myself’
 - (g) ni-[ka-nɔ ‘I dress him’

However, in other North Atlantic languages, phonological factors determine stress placement. In Senegalese Wolof, for example, stress is assigned to the

initial syllable of the word (Ka 1988). Non-initial heavy syllables (CVV) are assigned secondary stress (except a heavy syllable adjacent to an initial syllable receives main stress). In the Pulaar variety of Ffulde, the heaviest syllable in the word (CVVC>CVV>CVC>CV) is assigned main stress (except the final syllable is never stressed). If the heaviest syllables are of equal weight (Niang 1997), the leftmost receives main stress and the other(s) secondary stress.⁵ These points are illustrated in (10), below:

(10) (a) Wolof stress patterns (Ka 1988)

w <u>o</u> newu	‘to show off’
do <u>x</u> antuji	‘to go for a walk’
tera <u>a</u> nga	‘hospitality’
ju <u>m</u> tuka <u>a</u> y	‘tool’

(b) Pulaar stress patterns (Niang 1997)

<u>a</u> duna	‘world’
<u>a</u> samaan	‘sky’
da <u>d</u> orde	‘waist’
ta <u>l</u> lorde	‘place for rolling’
ha <u>k</u> aa <u>d</u> e	‘to perish’
ha <u>a</u> lpula <u>a</u> r [?] en	‘speakers of Pulaar’

In Noon (Cangin), spoken in Senegal, the penult is regularly assigned stress, except in three syllable words where stress instead is on the initial syllable (Soukka 2000)⁶. Neither syllable weight nor morphological boundaries play a major role in stress assignment. Interestingly, Noon also has high pitch assigned to the penult of every word. As shown in (11), pitch and stress generally coincide. However, in three-syllable words, high pitch is assigned to the penult while stress is on the initial syllable, showing that pitch and stress are independent in this language.

⁵ See Anyanwu (2002) for discussion of the phonetic correlates of stress in the Fuuta Jalloo variety of Pular, spoken in Guinea.

⁶ As Hyman (1977) notes, it is cross-linguistically common for languages with regular penult stress to assign stress to the initial syllable of three-syllable words or stems. (Additional examples of this pattern in African languages are illustrated in (12) and (16), below.) Hyman (1977) proposes that a plausible functional explanation for this alternation is that the most important phonetic cue to stress is a pitch fall over either the first two or the last two syllables of the word. When the two edges are in conflict, as when penult stress is assigned to a three-syllable word, stress tends to shift to the initial syllable. It is, however, problematic for this general explanation that High tone remains on the penult in three syllable words in Noon, even when stress is on the initial syllable. It would be interesting to know what phonetic cues there are to stress on the initial syllable.

- (11) Noon pitch and stress (Soukka 2000; p. 42); stressed vowel is underlined; accent indicates High tone

kédik	‘tree’
lĕ́gísoh	‘disentangle, explain’
kĕ́wdí:ri:	‘the pot’
hotubitéli:	‘the eucalyptus’

The question that arises is how, within the same subfamily, one could find languages with main stress at the opposite edges of a word: root-initial in Diola-Fogny vs. penult in Noon. As Hyman (1977) shows, this range of stress patterns is actually not unusual in families that include languages with quantity sensitive stress assignment. For example, Proto-Altaic is reconstructed with initial stress, but Turkish has regular final stress while in Khalka Mongolian stress placement is quantity sensitive. Hyman (1977) suggests the following plausible diachronic scenario relating these diverse systems. In the Proto-language, stress would be assigned to one edge of a morpheme. Since duration is a phonetic cue for stress, it is a natural change for stress to be attracted away from a morpheme edge towards a long syllable. If this quantity-sensitivity condition is lost in a subsequent change, then stress would again be assigned to a word or morpheme edge. At this stage of historical development, it would be equally natural for stress to be assigned either to the same edge as in the Proto-language or to the opposite edge.

6.2 Adamawa-Ubangi

The Adamawa-Ubangi languages are spoken in north-central Africa from northwest Nigeria as far as the DRC and Sudan. The Ubangi language Zande, spoken in the DRC and Sudan, has been described by Tucker (1959) as having stress, as well as tone. As shown in (12), stress (‘accent d’intensité’) falls on the first syllable of two and three syllable words (prefixes are ignored) and the penult of longer words. (Notice, this is the same pattern found in Noon, as shown in (11), above.) The first syllable of each half of a compound is also stressed.

- (12) Zande stress patterns (Tucker 1959, p. 48); stressed vowel is underlined; accents indicate tone

ngìrimò	‘flood’
a- ngìrimò	‘floods’
fùfùr`áfù	‘butterfly’
kpòto-ngbâ	‘rabbit (skin-mouth)’
rungudi-rúngúdí	‘red-hot’

Tone and stress are apparently independent. However, the correlates of stress are said to be vowel lengthening, more energy and a raise in pitch. As Tucker (1959) notes, the interaction between stress and pitch makes the Zande tone system difficult to establish.

6.3 Kru

Kru languages are spoken throughout Liberia and in adjacent areas in the Ivory Coast. According to Marchese (1989), Kru languages are tonal. Stress is reported for one language in this group, Grebo. It is controversial whether Grebo actually has stress, however. According to Innes (1966), stress and tone are independent in Grebo, with some words distinguished only by stress (the stressed vowel is underlined): eg., sono (21 tone pattern) ‘sting (noun)’ vs. son (21 tone pattern) ‘smell (noun)’. Newman (1986) reports, though, that he was unable to identify or confirm the stress patterns described by Innes. Indeed, Innes gives no clear description of the correlates of stress and suggests that contrastive stress can be neutralized in certain emphatic pronunciation styles. More work is certainly needed to resolve this controversy.

6.4 Mande

Mande languages are spoken widely over the western half of West Africa, with major communities found from Senegal to the Ivory Coast, along the coast, and towards the interior as far as Mali and Burkina Faso. Manding is the major language in this subfamily.

According to Dwyer (1989), Mande languages are tonal. Stress-accent has also been described for some languages in this group, like Kpelle, a Northwest Mande language. According to Welmers (1973), in Kpelle words with a High tone on the initial syllable or words that are Low-toned throughout, stress is assigned to the initial syllable. Words with non-initial High tones, e.g., with Mid-High-Low tone pattern, have stress assigned to the High-toned syllable. Words with a Mid tone throughout have no syllable appreciably more prominent. If High tones can be considered the equivalent of heavy syllables, this stress pattern is, then, analogous to the quantity-sensitive, leftmost stress pattern of Pulaar (10b), discussed above. What makes this an unusual stress system is that it lacks culminativity, since Mid-toned words are not stressed.

The phonotactics of some West Mande languages also display accent-like properties. According to Vydrine (to appear), most roots are maximally bisyllabic, and the distribution of segments is not symmetrical in bisyllabic words. In C2 position some phonological contrasts are neutralized or reduced, and Vydrine notes a tendency for C2 consonants to be lenited or deleted within

the West Mande languages. In C1 position a much fuller range of consonant contrasts is realized. Vowel contrasts are also often reduced in the second syllable: long vowels (in the languages where they exist) cannot occur word-finally in many of these languages. Finally, the tone system of many West Mande languages shows culminativity in the sense that the morpheme, not the syllable, is the domain of tone contrast.⁷ As argued in Section 1, these asymmetries in the distribution of consonants, vowel length and tonal contrasts are plausibly considered accentual phenomena since they increase the prominence of some syllables at the expense of others in the word.

6.5 *Gur*

The Gur languages are spoken in a belt from Mali through Burkina Faso to northern Togo and Benin. Mò:ré, with 1,700,000 speakers, is the major language in this subfamily.

Gur languages are tonal, and it is apparently controversial whether or not stress-accent is a feature of these languages. Naden (1989) suggests that what gives the impression of syllable prominence in Gur languages is a combination of High tone, intonation peak and vowel length. (See the discussion in Section 1, above.) However, Olawsky (1999) describes Dagbani as having regular stress on the penult independent of tone or vowel length. And Koromfe (Rennison 1997), Sopyire (Carlson 1994) and Mò:ré (Cann 1976) have all been described as having stress on the word-initial or root-initial syllable. As Cann shows, in Mò:ré stress is independent of tone: sídá ‘spouse’ vs s̀ibrì ‘fruit of the *Lanea oleosa*’. Note that the initial syllable of both words is stressed, even though one has a High tone and the other a Low tone.

6.6 *Delta-Cross*

As the name implies, the Delta-Cross languages are spoken in the Cross River area of southeastern Nigeria. These languages are all tonal with no reported stress-accent. However, Faraclas (1989) notes that several languages in this group have asymmetrical distributions of segmental contrasts within an initial bisyllabic or bimoraic string. As argued in Section 1, these asymmetries are best

⁷ Leben (2002) argues that a bisyllabic constituent, the tonal foot, is the domain for tone association in the Mande language, Bambara. As both syllables within the foot can bear High tone, we do not necessarily find the sort of tonal prominence asymmetry which is considered diagnostic of accent in this paper. The interested reader can consult Leben (2002) for arguments that the tonal foot provides non-accentual evidence for the interaction of metrical constituency and tone.

explained by appealing to foot structure, as they result in accent-like prominence asymmetries.

For example, in Gokana (Hyman 1990), inflected stems can have up to two consonants. The stem-initial (C1) consonant can be any one of the 15+ contrastive consonants of the language. However, only three of these consonants can occur in C2 position. As Hyman (1990) shows, this restriction on C2 holds whether it is in onset or coda position, so that it is not determined by syllable structure. A better explanation of the asymmetry is to propose that full consonant contrasts can only be realized in foot initial position. In foot medial position, contrasts are drastically reduced.

Similar distributional asymmetries between C1 and C2 are found in Efik (Hyman 1990, citing Cook 1985). Other evidence for the importance of the bimoraic foot in Efik comes from an alternation found in the negative suffix. As Hyman (1990) shows, the negative suffix is realized as [-kɛ́] after a CVC or a CVCV verb (13a) but as [-ɣɛ́] after a CV verb (13b):

(13) Efik alternations in the negative suffix (Hyman 1990; figs. (32), (33))

(a) After CVC and CVCV verbs

m̀-b̀òm-kɛ́	‘I am not breaking’
ɲ̀-ẁàt-kɛ́	‘I am not paddling’
ń̀-d̀òr̀ĩ-kɛ́	‘I am not putting on top (of)’
ń̀-s̀àṅ̀á-kɛ́	‘I am not walking’

(b) After CV verbs⁸

ń̀-d̄ĩ-ɣɛ́	‘I am not coming’
ń̀-sé-ɣɛ́	‘I am not looking at’
m̀-b̀ḍ-ɣó	‘I am not receiving’
ɲ̀-k̀à-ɣá	‘I am not going’

The context for the alternation is best captured by proposing that [k] occurs in foot-initial position (CVC and CVCV verbs form a bimoraic foot on their own and the negative suffix begins a new foot), while gamma occurs in foot-medial position. (Gamma is one of the three consonants which can occur in C2 – foot medial – position in Efik.)

As Akinlabi & Urua (2002), Harris (2004) and Urua (1999) show, Ibibio phonology also provides several motivations for foot structure. Distributional asymmetries between C1 and C2 are found, similar to those described for Efik and Gokana. There are also contrast asymmetries between V1 and V2, confirming the initial CV as the strong position within the bisyllabic foot. As in

⁸ Notice that the vowel of the negative suffix also harmonizes with the root vowel when it is within the same foot as the root vowel.

Efik, various allomorphic processes take as their conditioning context the segment's position within the foot. Finally, the foot defines a prosodic template for some verb paradigms. For example, Akinlabi & Urua (2002), Harris (2004) and Urua (1999) show that the vowel of CV roots is lengthened and that of CVVC roots is shortened when certain consonant-initial suffixes are added:

(14) Ibibio root vowel length alternations before the negative/reversive suffix

CVVC roots

fííp	‘suck’	fííp-pé	‘remove sucked object from mouth’
kóɔŋ	‘hang on hook’	kóɔŋ-ŋó	‘remove from hook’
wúúk	‘drive in’	wáák-kó	‘remove an object driven in’

CV roots

kǎ	‘go’	ń-kàà-γá	‘I am not going’
sé	‘look’	ń-séé-γé	‘I am not looking’
dí	‘come’	ń-díí-γé	‘I am not coming’

They argue these vowel length alternations are best explained by proposing the output of the root+suffix combination must be parsed as a heavy-light trochee.

Accent-like distribution asymmetries like those found in Delta Cross seem to be fairly widespread in West African Niger-Congo languages. Elsewhere, they have been reported for Ijoid (Williamson 1978), Kainji and Platoid (Gerhardt 1989) and Adamawa (Boyd 1989) languages. (Sections 4, 6.4, and 6.7.2. in this paper discuss similar asymmetries for Khoisan, Mande and Bantu languages, respectively.)

1.7 Bantu

The Bantu languages represent the largest group within Niger-Congo, comprising 500+ languages, spread over most of Africa south of a line extending from southern Cameroon in the west to southern Somalia in the east. While most Bantu languages have been described as tonal (and Proto-Bantu is reconstructed as having two tones), almost all are also described as having accentual properties. Many of the languages have stress-accent independent from tone, while a few have lost tone to retain only stress. Further, many Bantu tone systems have been described as having pitch accent features. And several languages show evidence of accent-like distribution asymmetries. This section surveys the role of accent in Bantu, beginning with the relatively uncontroversial stress patterns and segment distribution asymmetries and ending with a discussion of the more controversial role of accent in defining properties of Bantu tone systems.

1.7.1 *Stress-accent in Bantu languages*

Stress-accent is widely reported for languages from all over the Bantu-speaking area. Some of the earliest work on Bantu prosody, like Werner (1919), observes that the penult syllable is the most likely to be stressed. The table below, which lists a geographically diverse set of Bantu languages, shows that this is indeed the overwhelmingly most common pattern:

(15) Summary table of stress-accent in Bantu languages

<u>Language</u>	<u>Stress position</u>	<u>Tone?</u>	<u>Source</u>
Bondei	penult	yes	Philippson 1998
Chichewa	penult	yes	Odden 1999
Chimwi:ni	penult	no	Kisseberth & Abasheikh 1974; Kisseberth 2000
Chizigula	penult	yes	Philippson 1998
Digo	penult	yes	Philippson 1998
Giryama	penult	yes	Cassimjee & Kisseberth 2000
Kinande	penult	yes	Mutaka 1994; Mutaka & Hyman 1990
Kirufiji	penult	yes	Odden 1999
Kishambaa	penult	yes	Odden 1999; Philippson 1998
Kituba (Kongo véhiculaire)	penult	no	Odden 1999
Lingala, eastern dialects	penult	yes	Guthrie & Carrington 1988
Lingala, western dialects	stem-initial	yes	Guthrie & Carrington 1988
Luvale	penult (and verb stem initial)	yes	Horton 1949
Makonde	penult	yes	Odden 1999, Liphola 2001
Namwanga	penult	yes	Bickmore 2001
Nguni (Zulu, Xhosa, Swati, Ndebele)	penult	yes	Doke 1954
Nyakyusa	penult	no	Felberg 1996
Nyooro-Tooro	penult	no	Taylor 1985
Pogolo	penult	no	Odden 1999
Runyankore, Low	penult	no	Johnson 1976
Shona	penult	yes	Doke 1954
Sotho-Tswana	penult	yes	Doke 1954
Swahili	penult	no	Ashton 1969
Taita	penult	yes	Odden 1999
Tsonga	penult	yes	Doke 1954
Tumbuka	penult	no	Odden 1999
Venda	penult	yes	Doke 1954

The other common position for stress is stem-initial, as reported for western dialects of Lingala by Guthrie & Carrington (1988). As shown in the contrasting pair below, the penult bears main stress in the east, while the first syllable of the stem is stressed in the Lower Congo River area. When the stem contains more than three syllables, the penult is also stressed in the Western area.⁹

- (16) Lingala stress (Guthrie & Carrington 1988); [‘ marks stem edge; stressed syllable is underlined
- (a) stress on penult in the East: to-ko-[kanisa ‘we shall think’
 - (b) stress on stem-initial in West to-ko-[kanisa ‘we shall think’
 - (c) stress also on penult in West ba-ko-[palanganisa ‘they will scatter’
for longer stems

The most commonly reported phonetic correlate of penult stress, for both the tonal and non-tonal languages listed above, is vowel lengthening, with sources often noting that lengthening is most noticeable when words are in phrase-final position.¹⁰ However, Chimwi:ni presents an interesting exception to the generalization that stress and positionally restricted vowel length tend to coincide.¹¹ As Kisseberth & Abasheikh (1974) and Kisseberth (2000) show, the surface distribution of long vowels in Chimwi:ni is restricted: they can only occur in the penultimate or antepenultimate syllable of a phrase. Since these positions are familiar from languages like Latin as typical ones for the realization of stress, one might propose that length correlates with stress-accent (Selkirk 1986). However, Kisseberth & Abasheikh (1974) and Kisseberth (2000) make clear that the vowel length distribution pattern does not correlate with accent. Accent is realized as a higher pitch, assigned in the default case to the penult syllable of a phrase, or to the final syllable in some grammatical constructions, never to the antepenult. As intonational pitch is aligned with these accents, not with syllables containing long vowels, accent is arguably the more

⁹ As Rob Goedemans (p.c.) points out, Western Lingala (16b, c) very plausibly illustrates yet another example of the pattern found in Noon (11) and Zande (12), where regular stress assignment to the penult in words/stems of more than three syllables alternates with stress assignment to the initial syllable in shorter words/stems.

¹⁰ Liphola (2001) shows that unstressed mid vowels preceding the lengthened penult reduce to [a] in his dialect of Makonde. That is, reduction in unstressed position enhances the prominence of stressed vowels. It is an interesting research question whether similar stress-conditioned vowel reduction is found in other Bantu languages.

¹¹ Chimwi:ni is a non-tonal dialect of Swahili spoken on the coast of southern Somalia. As Abasheikh & Kisseberth (1974) and Kisseberth (2000) show, Chimwi:ni has an underlying vowel length contrast as well as several contexts where only long vowels predictably occur.

stress-like form of positional prominence in this language. The data below provide examples of the mismatch between positionally restricted vowel length realization and positional accent in Chimwi:ni:

- (17) Chimwi:ni (Kisseberth 2000); accents indicate stress accent, colons indicate vowel length
- | | | | | |
|-----|--------------|--------------|------------------------|-----------------------|
| (a) | ku-lokóta | | ‘to pick up’ | |
| (b) | x-furahikíla | | ‘to be pleased for’ | |
| (c) | ku-da:rána | | ‘to touch one another’ | |
| (d) | x-sa:řira | | ‘to travel’ | |
| (e) | fí:le | ‘s/he died’; | fíle: n-dála | ‘s/he died of hunger’ |
| | n-dála | ‘hunger’ | | |
| (f) | m-zé:le | ‘old man’; | m-zele úje | ‘that old man’ |
| | úje | ‘that’ | | |

1.7.2 Accent-like segment distribution asymmetries in Bantu languages

In most Bantu languages there is a marked asymmetry between the segmental contrasts found in the first syllable of the stem compared with subsequent stem syllables. As Bennett (1978; cited in Hyman 1989) shows, the full Proto-Bantu seven-vowel contrast is only realized in the stem-initial syllable. Elsewhere, a reduced number of contrasts is found. Similarly, only the stem-initial vowel contrasts systematically for tone. Noun prefixes and verb derivational suffixes do not contrast for tone. Vowel length contrasts are also mainly confined to the stem-initial syllable. Bennett (1978) argues that these asymmetries can all be best accounted for if the stem-initial syllable is accented. As argued in Section 1 (and subsequent sections), we expect, cross-linguistically, a fuller distribution of features under accent and reduced contrasts in unaccented syllables. It is, in fact, a traditional claim about Bantu languages (see e.g. Werner 1919) that the stem-initial syllable bears what she calls “etymological stress.” (As we saw in the previous section, in Western Lingala the stem initial syllable receives stress-accent, and in the next section we shall see that it also is accented for tone realization in some Bantu languages.) It is not clear in much of the traditional literature what is meant by stress, but certainly the tone and vowel quality asymmetries make the stem-initial syllable more perceptually prominent.

Other evidence for accentual prominence of the stem-initial syllable comes from some NW Bantu languages where we find the same sorts of consonant distribution asymmetries familiar from Khoisan and other West African Niger-Congo languages. For example, Hyman (1990, 2000) shows that in Basaa, roots have the shape C1V(C2), and stems are maximally three

syllables. All 22 consonants of the language contrast in C1 position. In C2 position, only 12 consonants can occur, and even fewer are possible in C3. Particularly striking evidence that these restrictions are accent based, not syllable based, comes from restrictions on voicing contrasts in obstruent stops. As Hyman (1990, 2000) shows, C1 is voiceless while C2 is voiced (and continuant between vowels) whether C2 is in the onset or coda:

(18) Basaa voicing complementarity (Hyman 1990, pp 184-185)

(a) P,T,K in C1 position are voiceless ([li-] is a noun class prefix preceding the root)

póó	‘to borrow’	li-pàn	‘forest’
táj	‘to deny’	li-tàm	‘fruit’
kém	‘to groan’	li-kùŋ	‘owl’

(b) P,T,K in C2 position are voiced (three dots indicate phrase-medial form)

(bi) Onset		(bii) Coda	
híŋ-îl	‘to breathe for/at’	héb...	‘to breathe’
kér-êl	‘to pick for/at’	kéd...	‘to pick fruit’
ǂúy-ûl	‘to break for/at’	ǂóg...	‘to break’

Similar restrictions are found in other NW Bantu languages like Kukuya (Teke) and Tiene. According to Hyman (1989), stems in Kukuya are maximally three syllables. The first stem consonant can be any of the 30+ contrastive consonants of the language, while only six consonants are found in C2 and C3 position. Likewise, in Tiene (Hyman & Inkelas 1997), stems are maximally three syllables. As in Basaa and Kukuya, all consonantal contrasts are found in C1 position. Hyman & Inkelas do not discuss restrictions on C2 in bisyllabic stems. However, they show that in trisyllabic stems C2 and C3 must agree in nasality, and C2 must be coronal while C3 must be non-coronal. These asymmetries in consonant distribution fall out if the stem-initial syllable is accented, and if, cross-linguistically, lack of accent is the best explanation for reductions in contrast.

1.1.3 Pitch-accent features of Bantu tone systems

That tone realization in Bantu languages, especially in verbs, has accentual properties has been recognized since McCawley (1970). As this work emphasizes, what makes it difficult to characterize Bantu tone systems as accentual is that many languages with distinctive accentual properties also have many properties characteristic of more ‘tonal’ systems. As a result, much theoretical work on Bantu tone has focussed on how the interaction between the

accentual and tonal properties should best be formalized, and even whether the distinction should be formalized. (See Downing (1996), Hyman (1989) and Odden (1988, 1995, 1999) for critical surveys of accent in Bantu. And see Cassimjee & Kisseberth (1998) for a current prosodic approach to Bantu tone.¹²) This section naturally focuses on the accentual properties of Bantu tone systems, and ignores the more tonal properties. Because Bantu accent is such a theoretically controversial topic, two caveats are in order. First, it should be made clear that pitch-accent is used here to refer to particular stress-like properties of tone realization discussed in Section 1 – culminativity, positional restrictions and tone-(stress) accent interactions - and not to any particular formal theory of accent. The interested reader should also consult the works cited to gain an appreciation of how the more tonal properties of the system interact with the pitch-accent properties highlighted here.¹³

1.1.3.1 Culminativity effects

As noted in Section 1, a distinctive characteristic of stress-accent is that it is culminative. In many Bantu languages, we find culminativity restrictions on High tones within stems or words that make these tone systems resemble pitch-accent systems. First of all, just as in stress systems only one (main) stress is found per word or stem, in many Bantu languages only one High tone is found per word or stem. As noted in the preceding section (see, too, Kisseberth & Odden 2003), only the stem-initial syllable of verbs contrasts for tone. In many Bantu languages (see, e.g., Philippson 1998), noun stems also have only one (underlying) High tone per stem, not one on every syllable, even though High tones can be found contrastively on all syllables of the stem.¹⁴ These points are illustrated in the Jita data below:

¹² As these authors make clear, it can be quite confusing to read through the literature on Bantu accent because the theoretical definition of accent and the tone systems motivating accentual formalisms have changed so much and so rapidly. Frequently the same authors writing on the same languages have changed their minds about whether the tone system of the language should be analyzed as accentual. As a result, when reading papers on Bantu accent, it should be borne in mind that a language is being characterized as “accentual” based on the theory of accentual representation prevailing at the time, rather than on some atheoretical conception of pitch-accent.

¹³ In some Bantu languages, the surface High tones are realized in an alternating pattern reminiscent of bounded stress systems. Work like Bickmore (1995), de Lacy (2002) and Pulleyblank (1983) has argued for a metrical (accentual) analysis of these tone patterns. As these analyses assume a different definition of accent from the one adopted in this paper, they are not considered in this survey of the accentual properties of Bantu tone realization. The interested reader should consult the works just cited for further discussion.

¹⁴ In the data in the remainder of this paper, acute accent indicates High tone and Low tone is unmarked.

(19) Jita High tone culminativity in lexical tone contrasts (Downing 1996)

(a) Verbs (infinitive form is shown; oku- is the infinitive prefix)

<i>High-toned stems</i>		<i>Low-toned stems</i>	
oku-lyá	‘to eat’	oku-sya	‘to grind’
oku-ǫ́na	‘to see’	oku-ǫ́uma	‘to hit’
oku-í:ga	‘to look for’	oku-i:ga	‘to pass a test’

(b) Nouns (stem follows hyphen)

omu-gási	‘woman’	omu-sa:ni	‘friend’
omu-tu:ngá	‘rich person’		
ji:-ng’okóra	‘knees’	omu-lamusi	‘judge’
li-dariná	‘tangerine’		

Most Bantu languages appear to be like Jita, imperfectly culminative since they preserve a contrast between stems which have a High tone and those which are toneless. However, some Bantu languages have taken a step further towards having an accentual system in that all words have a single High tone, consistently assigned to a limited number of positions in the word. For example, Odden (1988) shows that in Hibena, every noun must have a High tone, on either the penult or the pre-stem vowel, and most verb forms require a High tone on the penult:

(20) Hibena (Odden 1988; p 236, figs (19) and (20))

(a) Nouns

mú-goosi	‘man’
hí-fuva	‘chest’
mu-guúnda	‘field’
lu-fwiili	‘hair’
li-fulúha	‘cloud’

(b) Verbs

kwaamíle	‘put to pasture’ (subjunctive)
ndi-líma	‘I will cultivate’ (near-future)
ndaa-limága	‘I used to cultivate’
ndaa-limiige	‘I was cultivating’
ndihaa-limíle	‘I cultivated’ (intermediate past)
ndaa-limíle	‘I cultivated’ (far past)
hu-limíla	‘to cultivate for’

Other languages like Hibena are discussed in Odden’s (1988, 1999) critical overviews of what he calls “predictable” tone systems in Bantu languages.

Besides these lexical culminativity effects, many Bantu languages also have processes eliminating sequences of High tones. One of the most common of these is Meeussen's Rule (Goldsmith 1984, Clements & Goldsmith 1984), a process which deletes all but one High tone in a sequence.¹⁵ The Jita data in (21) illustrates this process. Three consecutive syllables – the Yesterday Past prefix /-má-/ , the object prefix immediately preceding the stem, and the verb stem-initial syllable – arguably contain High tones in the input. (See Downing (1996) for a detailed analysis of the Jita tone system.) However, only one High tone, the one on the object prefix, surfaces (and shifts one syllable to the right unless the following syllable is phrase-final):

(21) Jita examples of Meeussen's Rule (Downing 1996, p 220) – High-toned input vowels are underlined; '[' indicates the stem edge

- (a) a:-ma-gá-[lya] 's/he ate them (class 6)'
- (b) a:-ma-mu-[óna] 's/he saw him (class 2)'
- (c) a:-ma-ku-[gósorera] 's/he visited you'
- (d) a:-ma-ci-[sí:ndika] 's/he pushed us'

As Clements & Goldsmith (1984) and Philippson (1998) argue, Meeussen's Rule is one of the important synchronic and diachronic processes lending an accentual character to Bantu tone systems, because it results in culminative prominence of High tone.

Another common process resulting in culminativity is tone shift. The Jita data in (22) provide examples of this process. Notice that an input High tone is systematically realized one syllable to the right of its input sponsor unless the sponsor is in the penult or final syllable.

(22) Jita tone shift (High-toned input vowels are underlined; '[' indicates the stem edge)

- (a) oku-[óna] 'to see'
- (b) oku-[onána] 'to see each other'
- (c) oku-[uma] 'to hit'
- (d) oku-mu-[úma] 'to hit him/her (class 1)'

¹⁵ As Cassimjee (1998), Cassimjee & Kisseberth (1998) and Kisseberth & Odden (2003) argue, Meeussen's Rule is one of a family of processes found in Bantu languages which are motivated by the Obligatory Contour Principle (OCP). The interested reader can consult these works for discussion and exemplification. See, too, work like Goldsmith (1984) and Bickmore (1992) for analyses casting Meeussen's Rule in terms of accent clash resolution.

- | | | | | |
|-----|-------------------|----------|-----------------------|----------------|
| (e) | oku-[fwá | ‘to die’ | kumugera | ‘by the river’ |
| | oku-[fwa kúmugera | | ‘to die by the river’ | |

As Kenstowicz (1993), Cassimjee & Kisseberth (1998) and Kisseberth & Odden (2003) show, tone shift is clearly related to the assimilatory process of tone spread. What is puzzling is why the High tone should delink from its sponsor syllable, since other types of feature assimilation do not typically involve such delinking. Kenstowicz (1993) argues that delinking is best explained if High tone in Jita is analogous to stress accent. Then delinking can be seen as a process leading to High tone culminativity. Since it eliminates the sequence of High tones on adjacent syllables derived by tone spread, it makes one syllable of the string originally involved in tone spread more prominent than the others.¹⁶

1.1.1.2 Positional restrictions on tone realization

As noted in Section 1, main stress cross-linguistically tends to occur on syllables at the edge of a stem or word. In many Bantu languages, we find that tones are extremely mobile, spreading or shifting to positions several syllables from their sponsor to syllables at word or stem edges. The target for spread or shift in many languages is the penult syllable (see, e.g., Philippson 1998). For example, Cassimjee & Kisseberth (2000) show that in Giryama, if a verb word contains a single High tone, it surfaces on the (stressed) penult syllable, no matter which syllable in the word sponsors the High tone:

- (23) Giryama High tone shift to penult (Cassimjee & Kisseberth 2000, fig (4)); High-toned sponsor vowels are underlined
- | | | | | |
|-----|------------------|------------------------------|----------------------|---------------|
| (a) | ku-zi:ra | ‘to hate’ | <u>a</u> -na-zí:ra | ‘s/he hates’ |
| (b) | ku-lamu:sa | ‘to greet’ | <u>a</u> -na-lamú:sa | ‘s/he greets’ |
| (c) | ku-kalángí:ra | ‘to fry for’ | | |
| | ku-kalángirá:na | ‘to fry for each other’ | | |
| (d) | ni-kí-rí:ma | ‘if I cultivate’ | ni-kí-lamú:sa | ‘if I greet’ |
| (e) | hu-dza-lamusá:na | ‘we have greeted each other’ | | |

As Cassimjee & Kisseberth (1998), Odden (1999) and Kisseberth & Odden (2003) argue, one reason the penult is a likely target of tone spread or shift is that High tone realization is subject to the metrical principle of Nonfinality. If a High tone spreads or shifts as far to the right edge of the word as possible,

¹⁶ I am abstracting away from the theoretical details of Kenstowicz’s (1993) analysis. The interested reader can consult Kenstowicz (1993) and Cassimjee & Kisseberth (1998) for discussion of how delinking can be formalized in terms of metrical/prosodic constituency.

avoiding the final syllable, then the (rightmost or only) High tone will surface on the penult.

The penult is also a common position for main stress in stress-accent systems, often favored over the final syllable. Hyman (1977) argues that there is a perceptual reason for this tendency. The pitch fall that is one of the most salient perceptual cues to stress-accent is more noticeable if it is realized over two syllables (e.g., the penult and final) than only on the final syllable. The tendency for High tones to be realized on the penult and avoid final syllables in Bantu languages plausibly has a similar motivation: it allows High tones to be realized in a more perceptually salient position.

Cassimjee & Kisseberth (1998) and Kisseberth & Odden (2003) argue that the principle of Nonfinality unifies other processes found in Bantu languages which result in a High tone on the penult instead of the expected final syllable. For example, as shown in (21) and (22) above, High tones regularly shift one syllable rightward from their sponsors in Jita. A regular exception, shown in (21a) and (22a), is that they never shift from the penult to a phrase-final syllable. (See Downing (1996) for detailed discussion.) And in languages like High Runyankore High tones sponsored by final syllables shift to the penult phrase-finally (Poletto 1998). All of these processes motivated by Nonfinality reinforce the prominence of the penult syllable, as they converge on making it the target of High tone realization.¹⁷

Besides the penult, the stem-initial syllable is also a common target for High tone shift or spread. For example, in Giryama, if the verb word contains two High tones, the rightmost one shifts to the penult (see (23), above) while the leftmost shifts to the initial syllable:

(24) Giryama High tone shift to stem-initial syllable and penult (Cassimjee & Kisseberth 2000, fig. (18)); [‘ indicates the left stem edge

- (a) ku-[tsagatsá:ga ‘to break, damage’
 a-na-[tságatsá:ga ‘s/he breaks’
- (b) ku-[kubaliyá:na ‘to agree’ a-na-[kúbaliyá:na ‘s/he agrees’
- (c) ku-[tsundzulú:ka ‘to become sober’
 a-na-[tsúndzulú:ka ‘s/he becomes sober’

Other examples of shift of some High tones to the stem-initial syllable are found in Digo (Kisseberth 1984) and Chizigula (Kenstowicz & Kisseberth 1990, Kisseberth 1992).

¹⁷ Another frequent Nonfinality effect is found in languages like Bondei where High tones are deleted phrase-finally. This process has what might be called an “anti-accentual effect” since it eliminates High tones altogether from the stem.

1.1.1.3 *The interaction of tone and stress-accent*

A final property of Bantu tone languages which lends them an accentual character is the interaction of tone with (stress-)accent. As shown above, the penult is the most common locus for stress-accent in Bantu languages. It is no accident, then that, as Kisseberth & Odden (2003) and Philippson (1998) note, one finds a number of “penult effects” on tone in a number of Bantu languages. First, as illustrated by the Giryama data in (23) and (24), in at least some of the languages where High tones shift to the penult, the penult also has stress-accent. (This is indicated by length on the penult in Giryama.) As Philippson (1998) argues, one motivation for High tone shift to the penult is that the High tone is “attracted” to the stressed syllable. As a result, pitch change, a main correlate of stress-accent, is realized on the same syllable as other correlates of stress-accent (notably, duration).

Another penult effect on tone is found in Chaga (McHugh 1990, 1999). Chaga has a Super-High tone which results from raising High tones which occur on the phrase penult syllable (if it is not followed by a High-toned syllable in the phrase). Low tones on the phrase-penult syllable are also raised in the same context. As McHugh (1990, 1999) plausibly argues, tone raising is best explained as the realization of an accent on the phrase penult syllable.

Paradoxically, a final common penult effect is for High tones to target the syllable before the penult (i.e., the antepenult). Perhaps the most well-known example of this is found in Nguni languages like Xhosa. As shown in the data below, the penult is lengthened under stress-accent while High tones regularly spread or shift to the antepenultimate syllable:

(25) Xhosa antepenult shift (Cassimjee & Kisseberth (1998), p 77, fig (61));
High-toned sponsor vowels are underlined

- (a) ba-ya-chukúmi:sa ‘they provoke’
- (b) ba-ya-móne:la ‘they are jealous’
- (c) ba-ya-xóle:la ‘they forgive’
- (d) ba-ya-qononóndi:sa ‘they emphasize’

Cassimjee & Kisseberth (1998) argue that High tones in Xhosa (and other languages like this) target the antepenult instead of shifting further right (to the penult) to avoid the syllable which is prominent for stress-accent. More research is necessary to understand why High tones are attracted to the stressed syllable in some languages, while High tone and stress-accent remain independent in others. Independence of tone and stress are especially intriguing in languages like Xhosa where both are positionally conditioned.

As we can see, several factors converge to make the stem-initial syllable and the penult more prominent in Bantu languages. The stem-initial syllable is the locus of tone and vowel contrast in verb stems and a target of tone shift or spread, and it has stress-accent in Western Lingala. The penult has stress-accent in many Bantu languages and is a frequent site of High tone realization, either through shift or spread or to avoid a High tone on the phrase-final syllable. And the penult is the locus of tone raising in languages like Chaga. All of these properties lend credence to traditional claims (see e.g. Werner 1919) that the initial syllable and the penult are accented in Bantu languages.

7 Conclusion

As we have seen, African languages exhibit a rich variety of accentual systems. Many have familiar stress-accent systems, with the penult and/or initial syllable being especially common loci for stress-accent. These results confirm Hyman's (1977) finding that penult and initial positions are most likely to be assigned main stress, cross-linguistically. Many of the languages also have asymmetries in feature distribution or positional restrictions on tone realization which I have argued are best considered accentual in nature. The variety of accentual systems we find in African languages, indeed usefully highlights how elusive the notions of prominence and accent are, so that this survey leaves us with a number of questions. How is stress-accent phonetically realized in languages where vowel length and tone are contrastive? How can one reliably distinguish the impression of prominence given by vowel length and tone from stress-accent? How are apparently competing positional prominence asymmetries (like positionally restricted vowel length and positionally restricted tone) coordinated (or distinguished)? And do languages with featural distribution asymmetries within a bimoraic or bisyllabic constituent have other correlates of stress-accent within that same domain? It is hoped that this paper will stimulate more research on accent in African (and other) languages, and especially of the phonetic correlates of accent, to help resolve questions like these on the nature of accent and prominence systems.

References

- Abdel-Massih, Ernest T. 1971. *A Reference Grammar of Tamazight (Middle Atlas Berber)*. University of Michigan, Ann Arbor: Center for Near Eastern and North African Studies.
- Abu-Absi, Samir. 1995. *Chadian Arabic*. Munich: LINCOM EUROPA.
- Akinlabi, Akinbiyi & Eno Urua. 2002. Foot structure in the Ibibio verb. *JALL* 23, 119-160.
- Amha, Azeb. 1996. Tone-accent and prosodic domains in Wolaitta. *SAL* 25, 111-138.
- Angoujard, Jean-Pierre. 1990. *Metrical Structure of Arabic*. Dordrecht: Foris.
- Anyanwu, Rose-Juliet. 2002. On the manifestation of stress in African languages. In Ulrike Gut & Dafydd Gibbon, eds. *Proceedings, Typology of African Prosodic Systems*. Bielefeld Occasional Papers in Typology 1, 149-157.
- Applegate, Joseph R. 1958. *An Outline of the Structure of Shilha*. ACLS Program in Oriental Languages, Publications Series B – Aids – No. 11. New York: ACLS.
- Ashton, E.O. 1969. *Swahili Grammar*. London: Longmans.
- Beaujard, Philippe. 1998. *Dictionnaire Malgache-Français*. Paris: L'Harmattan.
- Bender, Lionel M. 2000. Nilo-Saharan. In Heine & Nurse, eds., 43-73.
- Bendor-Samuel, John, ed. 1989. *The Niger-Congo Languages*. Lanham, MD: University Press of America.
- Bennett, Patrick R. 1978. The role of stress and intonation in early Proto-Bantu. ms. University of Wisconsin. (Paper presented at the 9th Annual Conference on African Linguistics, Michigan State University.)
- Bickmore, Lee S. 1992. Multiple phonemic stress levels in Kinyambo. *Phonology* 9, 155-198.
- Bickmore, Lee S. 1995. Tone and stress in Lamba. *Phonology* 12, 307-341.
- Bickmore, Lee. 2001. Downstep and Fusion in Namwanga. *Phonology* 17.3, 297-333.
- Boyd, Raymond. 1989. Adamawa-Ubangi. In Bendor-Samuel, ed.
- Cann, Gaston. 1976. *La langue Mò:ré*. Paris: SELAF.
- Carlson, Robert. 1994. *A grammar of Supyire*. Berlin: Mouton de Gruyter.
- Cassimjee, Farida. 1998. *IsiXhosa Tonology: an Optimal Domains Theory Analysis*. LINCOM Studies in African Linguistics 29. Munich: LINCOM EUROPA.
- Cassimjee, Farida & Charles W. Kisseberth. 1998. Optimality Domains Theory and Bantu tonology: a case study from IsiXhosa and Shingazidja. In Larry M. Hyman & Charles W. Kisseberth, eds. *Theoretical Aspects of Bantu Tone*. Stanford, Calif.: CSLI, 33-132.
- Cassimjee, Farida & Charles W. Kisseberth. 2000. Opaque and depressed: Giryama tonology. Paper presented at the Tone Symposium, Tromsø, Norway, June 2000.
- Clements, G.N. & John Goldsmith. 1984. Introduction. *Autosegmental Studies in Bantu Tone*. Dordrecht: Foris, 1-17.
- Cook, Thomas L. 1985. *An integrated phonology of Efik*. Volume I. University of Leiden.
- deLacy, Paul. 2002. The interaction of tone and stress in Optimality Theory. *Phonology* 19, 1-32.
- Doke, C.M. 1954. *The Southern Bantu Languages*. London: Oxford University Press for the International African Institute.
- Downing, Laura J. 1996. *The Tonal Phonology of Jita*. LINCOM Studies in African Linguistics 05. Munich: LINCOM EUROPA.

- Dresel, Linda. 1977. *An Experimental Study of Hausa Tone*. PhD dissertation, Indiana University.
- Dwyer, David. 1989. Mande. In Bendor-Samuel, ed.
- Ethnologue*. 2000. Edited by Barbara F. Grimes. Consulting Editors, Richard S. Pittman & Joseph E. Grimes. Internet version, <http://www.sil.org/ethnologue/>
- Faraclas, Nicholas G. 1989. Cross-River. In Bendor-Samuel, ed.
- Felberg, Knut. 1996. *Nyakyusa-English-Swahili and English-Nyakyusa Dictionary*. Dar es Salaam, Tanzania: Mkuki na Nyota Publishers.
- Fischer, Wolfdietrich & Otto Jastrow, eds. 1980. *Handbuch der Arabischen Dialekte*. Wiesbaden: O. Harrassowitz.
- Gerhardt, Ludwig. 1989. Kainji and Platoid. In Bendor-Samuel, ed.
- Gilley, Leoma. 2003. The feature of stress in Nilotic. In Rose-Juliet Anyanwu (ed.), *Stress and Tone – the African Experience*. *Frankfurter Afrikanistische Blätter* 15. Köln: Rüdiger Köppe Verlag, 99-120.
- Goldsmith, John. 1984. Meeussen's rule. In M. Aronoff & R.T. Oehrle, eds. *Language Sound Structure*. Cambridge, Mass.: MIT Press, 245-259.
- Güldemann, Tom & Rainer Vossen. 2000. Khoisan. In Heine & Nurse, eds., 99-122.
- Guthrie, Malcolm & John F. Carrington. 1988. *Lingala Grammar and Dictionary*. London: Baptist Missionary Society.
- Haacke, Wilfrid H. 1999. *The tonology of Khoekhoe (Nama/Damara)*. Köln: Rüdiger Köppe Verlag.
- Harrell, Richard. 1957. *The Phonology of Colloquial Egyptian Arabic*. ACLS Program in Oriental Studies, Publications Series B – Aids – No. 9. New York: ACLS.
- Harrell, Richard. 1962. *A Short Reference Grammar of Moroccan Arabic*. DC: Georgetown University Press.
- Harris, John. 2004. Release the captive coda: the foot as a domain of phonetic interpretation. In John Local, Richard Ogden & Rosalind Temple (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology* 6. Cambridge: Cambridge University Press, 103-129.
- Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.
- Hayward, Richard J. 2000. Afroasiatic. In Heine & Nurse, 74-98.
- Heath, Jeffrey. 1999. *A Grammar of Koyraboro (Koroboro) Senni. The Songhay of Gao, Mali*. Berlin: Walter de Gruyter.
- Heine, Bernd & Derek Nurse. 2000. Introduction. In Heine & Nurse, eds., 1-10.
- Heine, Bernd & Derek Nurse, eds. 2000. *African Languages: An Introduction*. Cambridge: Cambridge University Press.
- Hetzron, Robert, ed.. 1997. *The Semitic Languages*. London: Routledge.
- Horton, A.E. 1949. *A Grammar of Luvale*. Johannesburg: Witwatersrand University.
- Hudson, R. A. 1973. Syllables, moras and accents in Beja. *Journal of Linguistics* 9, 53-63.
- Hyman, Larry M. 1977. On the nature of linguistic stress. In L.M. Hyman, ed. *Studies in Stress and Accent*. *SCOPIL* 4, 37-82.
- Hyman, Larry M. 1978. Tone and/or accent. In Donna Jo Napoli, ed. *Elements of Tone, Stress, and Intonation*. Washington, DC: Georgetown University Press, 1-20.
- Hyman, Larry M. 1981. Tonal accent in Somali. *SAL* 12, 169-203.
- Hyman, Larry M. 1989. Accent in Bantu: an appraisal. *Studies in the Linguistic Sciences* 19, 115-134.

- Hyman, Larry M. 1990. Non-exhaustive syllabification: evidence from Nigeria and Cameroon. *CLS* 26 (2), 175-195.
- Hyman, Larry M. 2000. Cyclicity and templatic morphology in the Basaa verb stem. Paper presented at the 2nd Round Table in Phonology of the GDR 1954, Bordeaux, June 8, 2000.
- Hyman, Larry M. & Sharon Inkelas. 1997. Emergent templates: the unusual case of Tiene. In Viola Miglio & Bruce Morén, eds. *University of Maryland WPL* 5, 92-116.
- Innes, Gordon. 1966. *An Introduction to Grebo*. SOAS. Printed by The Hague: Mouton.
- Johnson, Lawrence. 1976. Devoicing, tone, and stress in Runyankore. *SCOPIL* 3, 207-216.
- Ka, Omar. 1988. *Wolof Phonology and Morphology: a Non-linear Approach*. PhD dissertation. University of Illinois at Urbana-Champaign.
- Kaye, Alan S., ed. & Daniels, Peter T., technical advisor. 1997. *Phonologies of Asia and Africa: including the Caucasus*. Winona Lake, Ind.: Eisenbrauns.
- Keenan, Edward L. & Maria Polinsky. 1998. Malagasy (Austronesian). In Andrew Spencer & Arnold M. Zwicky, eds. *The Handbook of Morphology*. Oxford: Blackwell, 563-623.
- Kenstowicz, Michael. 1993. Evidence for metrical constituency. In Kenneth Hale & Samuel Jay Keyser, eds. *The View from Building 20: Essays in Linguistics in honor of Sylvain Bromberger*. Cambridge, Mass.: The MIT Press, 257-273.
- Kenstowicz, Michael & Charles W. Kisseberth. 1979. *Generative Phonology: Description and Theory*. New York: Academic Press.
- Kenstowicz, Michael & Charles W. Kisseberth. 1990. Chizigula tonology: the word and beyond. In Sharon Inkelas & Draga Zec, eds. *The Phonology-Syntax Connection*. Chicago: The University of Chicago Press, 163-194.
- Kisseberth, Charles W. 1984. Digo tonology. In G.N. Clements & John Goldsmith, eds. *Autosegmental Studies in Bantu Tone*. Dordrecht: Foris, 105-182.
- Kisseberth, Charles W. 1992. Metrical structure in Zigula tonology. Derek F. Gowlett, ed. *African Linguistic Contributions*. Pretoria: Via Afrika Limited, 227-259.
- Kisseberth, Charles W. 2000. The phonology-syntax interface: Chimwiini revisited. ms., Tel Aviv University.
- Kisseberth, Charles W. & Mohammed I. Abasheikh. 1974. Vowel length in ChiMwi:ni – a case study of the role of grammar in phonology. In A. Bruck, A. Fox & M. W. La Galy, eds. *CLS Papers from the Parasession on Natural Phonology*, 193-209.
- Kisseberth, Charles W. & David Odden. 2003. Tone. In Derek Nurse & Gérard Philippson, eds. *The Bantu Languages*. London: Routledge, 59-70.
- Leben, William R. 2002. Tonal feet. In Ulrike Gut & Dafydd Gibbon, eds. *Proceedings, Typology of African Prosodic Systems*. Bielefeld Occasional Papers in Typology 1, 27-40.
- Leger, Rudolf. 1994. *Eine Grammatik der Kwami-Sprache (Nordost-nigeria)*. Köln: Rüdiger Köppe Verlag.
- Leslau, Wolf. 1995. *Reference Grammar of Amharic*. Wiesbaden: Harrassowitz.
- Liphola, Marcelino M. 2001. *Aspects of Phonology and Morphology of ShiMakonde*. Ph.D. dissertation, The Ohio State University.
- Marchese, Lynell. 1989. Kru. In Bendor-Samuel, ed.
- McCarthy, John. 1979. *Formal Problems in Semitic Phonology and Morphology*. PhD dissertation, MIT.

- McCawley, James. 1970. Some tonal systems that come close to being pitch accent systems but don't quite make it. *CLS* 6, 526-532.
- McCawley, James. 1978. What is a tone language? In V. Fromkin, ed. *Tone: A linguistic survey*. New York: Academic Press, 113-131.
- McHugh, Brian. 1990. The phrasal cycle in Kivunjo Chaga tonology. In Sharon Inkelas & Draga Zec, eds. *The Phonology-Syntax Connection*. Chicago: The University of Chicago Press, 217-242.
- McHugh, Brian. 1999. *Cyclicity in the Phrasal Phonology of Kivunjo Chaga*. LINCOM Studies in African Linguistics 03. Munich: LINCOM EUROPA.
- Miller-Ockhuizen, Amanda. 1998. Towards a Unified Decompositional Analysis of Khoisan Lexical Tone. In Mathias Schladt, ed. *Language, Identity and Conceptualization among the Khoisan*. (Research in Khoisan Studies 15). Köln: Rüdiger Köppe, 217-244.
- Miller-Ockhuizen, Amanda. 1999. C-V coarticulations and complex consonants: New evidence for the ordering of place gestures in click consonants. In O. Fujimura, B. D. Joseph & B. Palek, eds. *Proceedings of LP '98*. Prague: Charles University, 301-330.
- Miller-Ockhuizen, Amanda. 2001a. Contrastive vowel-length and variable weight reduplicative templates in Ju|'hoansi. In A. Bell & P. Washburn, eds. *Cornell Working Papers in Linguistics* 18, 104-125.
- Miller-Ockhuizen, Amanda. 2001b. *Grounding Ju|'hoansi Phonotactics: The Phonetics of the Guttural OCP and other Acoustic Modulations*. Ph.D. dissertation. The Ohio State University.
- Mitchell, T. F. 1993. *Pronouncing Arabic 2*. Oxford: Clarendon Press.
- Mous, Maarten. 1993. *A Grammar of Iraqw*. Cushitic Studies 9. Hamburg: Helmut Buske.
- Mutaka, Ngessimo M. 1994. *The Lexical Tonology of Kinande*. LINCOM Studies in African Linguistics 01. Munich: LINCOM EUROPA.
- Mutaka, Ngessimo M. & Larry M. Hyman. 1990. Syllables and morpheme integrity in Kinande reduplication. *Phonology* 7, 73-119.
- Naden, Tony. 1989. Gur. In Bendor-Samuel, ed.
- Newman, Paul. 1986. Contour tones as phonemic primes in Grebo. In Koen Bogers, Harry van der Hulst & Maarten Mous, eds. *The Phonological Representation of Suprasegmentals*. PALL 4. Dordrecht: Foris, 175-193.
- Newman, Paul. 2000. *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven: Yale University Press.
- Niang, Mamadou O. 1997. *Constraints on Pulaar Phonology*. Lanham, MD: University Press of America.
- Nicolai, Robert & Petr Zima. 1997. *Songhay*. Munich: LINCOM EUROPA..
- Noonan, Michael. 1992. *A Grammar of Lango*. Berlin: Mouton de Gruyter.
- Odden, David. 1988. Predictable tone systems in Bantu. In Harry van der Hulst & Norval Smith, eds. *Autosegmental Studies on Pitch Accent*. Dordrecht: Foris, 225-251.
- Odden, David. 1995. Tone: African Languages. In John A. Goldsmith, ed. *The Handbook of Phonological Theory*. Cambridge, Mass.: Blackwell, 444-475.
- Odden, David. 1999. Typological issues in tone and stress in Bantu. In Shigeki Kaji, ed. *Proceedings of the Symposium on Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics*. Tokyo: ILCAA, 187-215.
- Olawsky, Knut J. 1999. *Aspects of Dagbani Grammar*. Munich: LINCOM EUROPA.
- Owens, Jonathan. 1993. *A Grammar of Nigerian Arabic*. Wiesbaden: Harrassowitz.

- Philippon, Gérard. 1998. Tone reduction vs. metrical attraction in the evolution of Eastern Bantu tone systems. In Larry M. Hyman & Charles W. Kisseberth, eds. *Theoretical Aspects of Bantu Tone*. Stanford, Calif.: CSLI, 315-329.
- Poletto, Robert E. 1998. *Topics in Runyankore Phonology*. PhD dissertation, The Ohio State University.
- Pulleyblank, Douglas. 1983. Accent in Kimatuumbi. In Jonathan Kaye, Hilda Koopman, Dominique Sportiche & André Dugas, eds. *Current Approaches to African Linguistics* (vol. 2). Dordrecht: Foris Publications, 195-215.
- Rennison, John R. 1997. *Koromfe*. London: Routledge.
- Richter, Renate. 1987. *Lehrbuch der amharischen Sprache*. Leipzig: VEB Verlag.
- Ruhlen, Merritt. 1975, 1976. *A Guide to the Languages of the World*. [self-published].
- Sapir, J. David. 1965. *A Grammar of Diola-Fogny*. Cambridge: Cambridge University Press.
- Sasse, Hans-Jürgen. 1981. Die kuschitischen Sprachen. In Bernd Heine, Thilo C. Schadeberg & Ekkehard Wolf, eds. *Die Sprachen Afrikas*. Hamburg: Helmut Buske Verlag, 187-215.
- Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3, 371-405.
- Soukka, Maria. 2000. *A Descriptive Grammar of Noon: A Cangin Language of Senegal*. Munich: LINCOM EUROPA.
- Taylor, Charles. 1985. *Nkore-Kiga*. London: Croom Helm.
- Traill, Anthony. 1985. *Phonetic and Phonological Studies of !Xóö Bushman*. Hamburg: Helmut Buske Verlag.
- Tucker, A.N. 1993. *A Grammar of Kenya Luo (Dholuo)*. Ed. by C. A. Creider, with the collaboration of T. Okelo Odongo, E.D. Jakeyo Ang'ina & J. N. Olum Oludhe. Köln: Rüdiger Köppe Verlag.
- Tucker, A.N. & P.E. Hackett. 1959. Le groupe linguistique Zande. Tervuren: *Annales du Musée Royal du Congo Belge*, Série in-8°, Sciences de l'Homme, Linguistique, vol. 22.
- Urua, Eno-Abasi E. 1999. Length and syllable weight in Ibibio. *SAL* 28, 241-266.
- Van der Hulst, Harry & Norval Smith. 1988. Introduction. *Autosegmental Studies on Pitch Accent*. Dordrecht: Foris, ix-xxiv.
- Vydrine, Valentin. to appear. Areal and genetic features in West Mande and Mani-Bandama (East Mande) phonology: In what sense did Mande languages evolve? *JWAL*.
- Welmers, William E. 1973. *African Language Structures*. Berkeley, Calif.: University of California Press.
- Werner, Alice. 1919. *Introductory Sketch of the Bantu Languages*. London: Kegan, Paul, Trench, Trubner and Co.
- Williamson, Kay. 1978. Consonant distribution in Ijo. In Mohammad Ali Jazayery et al, eds. *Linguistic and Literary Studies in Honor of Archibald A. Hill*, vol. 3. The Hague: Mouton, 341-353.
- Williamson, Kay & Roger Blench. 2000. Niger-Congo. In Heine & Nurse, eds., 11-42.
- Wolff, H. Ekkehard. 1993. *Referenzgrammatik des Hausa*. Münster: LIT Verlag.

(Un)markedness of trills: the case of Slavic r-palatalisation*

Marzena Żygis

Zentrum für Allgemeine Sprachwissenschaft, Berlin

This paper evaluates trills [r] and their palatalized counterparts [r^j] from the point of view of markedness. It is argued that [r]s are unmarked sounds in comparison to [r^j]s which follows from the examination of the following parameters: (a) frequency of occurrence, (b) articulatory and aerodynamic characteristics, (c) perceptual features, (d) emergence in the process of language acquisition, (e) stability from a diachronic point of view, (f) phonotactic distribution, and (g) implications.

Several markedness aspects of [r]s and [r^j] are analyzed on the basis of Slavic languages which offer excellent material for the evaluation of trills. Their phonetic characteristics incorporated into phonetically grounded constraints are employed for a phonological OT-analysis of r-palatalization in two selected languages: Polish and Czech.

1 Introduction

It has often been observed that trills are marked sounds in comparison to other natural classes, e.g. stops or fricatives, which has been especially attributed to the articulatory complexity of these sounds (see e.g. Šimáčková 2003). Similarly, the articulatory complexity has appeared essential for the claim that palatalized trills are marked with respect to other palatalized coronals, e.g., [t^j], cf. Hall (2000).

In this paper the focus is placed on the markedness relation between trills and their palatalized counterparts. It will be shown that plain trills are unmarked with respect to their palatalized counterparts. The conclusion is drawn from an evaluation of [r]s and [r^j]s against several aspects of markedness such as (a) frequency of occurrence, (b) articulatory and aerodynamic characteristics, (c) perceptual features, (d) production in the process of language acquisition, (e)

* I would like to thank Susanne Fuchs and Silke Hamann for their comments on this paper.

stability from a diachronic point of view, (f) phonotactic distribution and (g) implications.

A direct contribution of the present analysis to a markedness concept is the involvement of aerodynamic, acoustic and perceptual evidence, mostly neglected in the literature.

Several aspects of [r] and [r^ɨ] will be examined on the basis of Slavic languages which provide excellent material especially for the examination of palatalized trills which are extremely rare sounds, see 2.3.

The role of markedness of plain and palatalized trills is furthermore examined in the phonology of Czech and Polish. The process of r-palatalisation is analyzed in both languages in the framework of Optimality Theory (OT) (Prince and Smolensky 1993, McCarthy and Prince 1993). It is shown that universal phonetically grounded constraints ranked differently in Polish and Czech lead to the selection of the optimal candidates in both languages.

This analysis is more economic than the one by rules (see Rubach 1984) since it requires neither an intermediate stage, i.e. an abstract segment /ř'/, nor additional Spell-out rules. Constraints alone make a distinction between the treatment of an underlying /ž/ and an underlying /r/, two potential inputs for palatalisation.

The article is organized as follows. In section 2, I evaluate plain and palatalized trills against various markedness relations and I show that aerodynamic evidence as well as perceptual saliency are important aspects of the concept of markedness. In section 3, I propose an analysis of r-palatalisation in Polish and Czech in terms of constraints. My conclusions are outlined in section 4.

2 The (un)markedness of trills

2.1 Rhotics as a natural class

Before discussing markedness of trills, let us consider the natural class of rhotics of which trills constitute a part. There is no consensus in the literature on how rhotics can be defined as a natural class. The difficulty is caused among others by the wide range of places of articulations (labial, dental, alveolar, postalveolar and uvular) in connection with different manners of articulation starting from trills via taps, flaps and through to approximants. This is shown in (1) where rhotics are systematized according to their place and manner of articulation. Below, some examples of languages are listed in which the single sounds are attested as phonemes. For the sake of completeness a labial trill (B) is also included which does not however belong to rhotics.

(1) Rhotics+trills²

	labial	dental/alveolar	postalveolar (retroflex)	uvular
Trill	ʙ	r (ř)	ɽ	ʀ
Flap/taps		ɾ	ɽ	
Approximants		ɹ	ɻ	
Fricative				ʁ

ʙ	Kele, Titan (Oceanic branch of Eastern Malayo-Polynesian)
r	Polish, Russian, Finnish, Icelandic, Kele, Toda, Titan
ř	Czech
ɽ	Toda
ʀ	German
ɾ	Hausa (West Chadic subgroup), Ngizim (Chadic branch, Bade subgroup)
ɽ	Hausa, Ngizim
ɹ	American English, Edo (Niger-Kordofanian language, Kwa group)
ɻ	Hausa
ʁ	French

There were some attempts to define rhotics with an acoustic feature. It has been suggested by Ladefoged (1975) and Lindau (1978) that rhotics display a common acoustic feature, i.e. a lowered third formant (F3). This conclusion cannot be maintained because as pointed out by Lindau (1985:165) and Ladefoged and Maddieson (1996:244) there are rhotics that do not display such a characteristic, i.e., Swedish and German uvular r-sounds show a third spectral peak over 2500 Hz and the Arrente retroflex approximant also has a high third formant. However, as recognized by Lindau (1985) rhotics behave phonologically as a natural class and therefore have been specified by the feature [rhotic]. This view is shared by Hall (1997a:107) who provides several arguments in favor of rhotics functioning as a natural class.

A different view is presented by Wiese (2001: 350) who offers a prosodic definition of r-sounds, according to which rhotics constitute a point on the sonority scale between laterals and vowels. However, Wiese points out that there are languages in which /r/s take a different position than the one predicted by the sonority scale, e.g., Polish /rt/, /rd/, /rv/. According to Wiese, such cases have to be treated separately. The present paper attempts to show that the position of /r/ in such clusters can be explained by taking into consideration the perceptual properties of /r/, presented in 2.5.

² Maddieson and Ladefoged (1996) also list a lateral flap /l/.

2.2 *Trills and palatalized trills in light of the markedness concept*

The concept of markedness comprises several aspects that refer to different dimensions of linguistic analysis, cf. Trubetzkoy (1931), Jakobson (1941), Chomsky and Halle (1968), Stampe (1972), Bailey (1973), Vennemann (1983, 1988, 1989). One of the most important characteristics of a marked segment is that it occurs less frequently than its unmarked counterpart in the inventories of the languages of the world. Further, it is acquired later in the process of language acquisition than an unmarked segment. This is often connected with more articulatory effort required for the production of a marked than an unmarked sound. From a diachronic perspective a marked segment is often less stable than its unmarked pendant in a sense that it undergoes various kinds of changes. Furthermore, as far as its phonotactics are concerned, a marked segment is more narrowly distributed than its unmarked counterpart. Finally, the presence of a marked sound in a given language implies the occurrence of its unmarked counterpart.

Apart from the properties mentioned above, the present paper also takes into account aerodynamic, acoustic and perceptual evidence, which, besides articulatory characteristics, is indispensable in accounting for the (un)markedness of certain natural classes as, e.g., trills. These sounds are reported to occur fairly frequently in the languages of the world, cf. 2.3., but they are articulatorily and aerodynamically extremely complex, cf. 2.4. This potential conflict can be resolved only if one takes into consideration perceptual characteristics of trills, cf. 2.5.

As far as the comparison between plain and palatalized trills is concerned, i.e. the main subject of the present study, the aerodynamic, acoustic and perceptual evidence highlights the markedness of palatalized trills providing a complete characteristic of these sounds.

The perceptual aspect of the linguistic markedness follows from two well-known underlying principles of linguistic developments proposed by Passy (1891), and presented in (2).

- (2) Passy's principles (1891:227)
 - a. Principle of economy:
'languages tend to get rid of anything that is superfluous'.
 - b. Principle of emphasis:
'languages tend to stress or exaggerate anything that is necessary'.

Both principles can be understood in articulatory and perceptive terms in the way proposed, e.g., by Boersma (1998) who translates the principle of economy into a speaker-oriented principle, see (3a), and the principle of emphasis into a listener-oriented principle, see (3b).

(3) Boersma's principles (1998:2)

- a. A speaker-oriented principle: minimization of articulatory effort
- b. A listener-oriented principle: minimization of perceptual confusion

According to the speaker-oriented principle, the articulatory effort put into the process of production of a sound should be minimal, and according to a listener-oriented principle, the sound should be so distinctive perceptually that it is easily distinguishable from other sounds.³

For the markedness concept, it means that a marked sound requires more articulatory effort than an unmarked one. As far as the perceptual aspect is concerned, I assume that a marked sound is perceptually less distinct than its unmarked counterpart; for an overview of possible perceptual approaches see Hamann (2003).

In addition to perceptual characteristics, aerodynamic evidence is taken into consideration which is indispensable in illustrating the complexity of trills as well as the differences between trills and their palatalized counterparts.

In the following, I evaluate plain and palatalized trills with respect to the parameters presented in (4).

- (4) (a) frequency of occurrence
- (b) articulatory and aerodynamic characteristics
- (c) perceptual features
- (d) emergence in the process of language acquisition
- (e) stability from a diachronic point of view
- (f) phonotactic distribution
- (g) implications

Taking into consideration the parameters in (4) it is expected that if palatalized trills are marked counterparts of plain ones, they should be, in contrast to plain trills: (a) less frequent in the phonemic inventories of the languages of the world, (b) articulatory and aerodynamically more complex, (c) perceptually less salient, (d) acquired later in the process of language acquisition, (e) less stable in diachronic processes, (f) more narrowly distributed in the phonotactics, (g) they imply the presence of plain trills in phonemic inventories.

³ Perceptual markedness has been neglected in the phonological literature since SPE (1968). Recent phonological development has shown the indispensability of perceptual characteristics in accounting for various phenomena, cf., e.g., Flemming (1995), Padgett (1997), NíChiosáin & Padgett (2001), Hume & Johnson (2001). The perceptual aspects of trills have been mentioned by e.g. Passey (1891), Solé (1998).

In the next sections it will be shown that there are strong arguments in favor of the marked character of palatalized trills following from the examination of parameters listed above.

2.3 *Frequency of occurrence*

Plain trills occur fairly frequently in the phonemic inventories of the languages of the world. The UCLA Phonological Segment Inventory Database (UPSID) in which phonemic inventories of 451 languages are listed, displays 155 languages having a plain trill in their inventories, i.e., 9 languages with a dental trill, 95 with an alveolar trill, 51 with either a dental or alveolar trill.

Palatalized trills occur extremely rarely in the languages of the world. UPSID includes 5 languages having [r^j] as a phoneme: Russian (dental), Lithuanian, Bulgarian (alveolar), Saami, and Nenets (dental or alveolar). The list can be extended by some Slavic languages not included in UPSID such as Upper and Lower Sorbian, Polabian and Ukrainian as well as Muinane, Daur and Toda provided by Hall (2000:8).

Additionally, as observed by Hall (2000:9), palatalized rhotics⁴ are less frequent in phonemic systems than palatalized non-rhotics, as, e.g., palatalized nasals or palatalized obstruents, cf. also Bhat (1978), Hock (1986:134ff) and Walsh-Dickey (1997:129ff.).

In summary, the palatalized trills are extremely rarely found in the phonemic inventories of the languages of the world. At the same time their plain counterparts [r]s occur fairly frequently as phonemes.

2.4 *Articulatory and aerodynamic characteristics*

In this section articulation and aerodynamics, as well as their interaction in the production of trills and their palatalized counterparts, is provided.

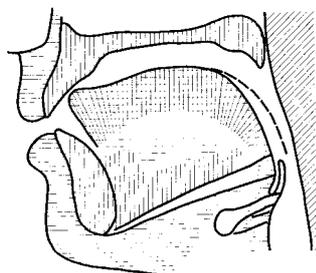
As far as articulation is concerned, there is a consensus in the literature concerning the main articulator of the coronal trills. It is namely the tongue tip that vibrates, cf. Wierzchowska (1971), Bolla (1981), Ladefoged&Maddieson (1996), Recasens (1991), Recasens & Pallarès (1999), Hall (2000), Solè (1998, 1999, 2002a,b).

The situation is less clear when palatalized trills are analyzed. Bolla (1982) shows that in the production of Russian [r^j] the tongue tip still remains the main articulator while the predorsum is raised which is a typical gesture for

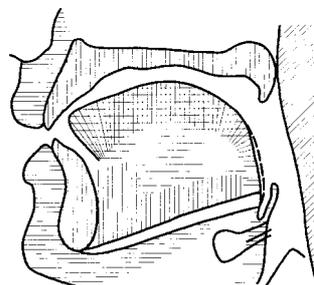
⁴ This generalization pertains to palatalized trills as well as other palatalized rhotics such as [r^j, R^j, x^j], see Hall (2000:8).

palatalized sounds. The x-ray tracings in (5) illustrate the articulation of Russian [r] and [r^j], see Bolla (1981: 76, 77).

(5) a. Russian [r]



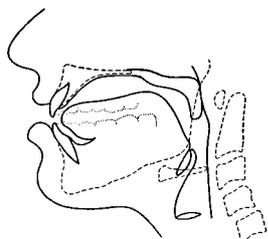
b. Russian [r^j]



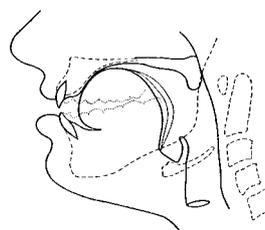
These x-ray tracing in (5b) clearly shows that the tongue tip remains the main articulator in the articulation of Russian [r^j]. This conclusion has been confirmed by Skalozub (1963) who has shown on the basis of palatograms that Russian [r^j] is articulated with the tongue tip at the dental place of articulation. The articulation of [r^j] with the tongue tip has also been postulated for all languages with palatalized rhotics by Hall (2000).

This is in contrast to Polish palatalized trill [r^j] which has been shown to be pronounced with the tongue blade, see e.g. Wierzchowska (1971). The x-ray tracings provided in (6) show Polish [r] in comparison to [r^j]; see Wierzchowska (1971:168, 191).⁵

(6) a. Polish [r]



b. Polish [r^j]



The x-ray tracings provided in (6) show that Polish [r^j] is pronounced with the tongue blade whereas [r] is articulated with the tongue tip. Wierzchowska (1971) is not explicit about this point. In addition, it is also not clear what phase of [r^j] articulation is presented by the x-ray tracing in (6).

Apparently there are at least two different manners of articulation of [r^j]: with the tongue tip and with the tongue blade. Since in the former case the predorsum is raised, the difference between the two manners are minimal. But a crucial fact for the present paper is that in both cases the articulation requires

⁵ Note that in Polish [r^j] occurs allophonically when [r] is followed by [i] or [j]. Such sequences are limited to words of foreign origin.

more articulatory effort than in the case of [r]. This is due to the lighter mass and reduced stiffness of the tongue tip which can vibrate more freely than in cases when the tongue tip is accompanied by a raised predorsum or when a tongue blade vibrates, see, e.g., Solè (1998).

Apart from the difficulty of maintaining vibration in [r^h] by the main articulator, it has been observed that in this sound two antagonistic articulatory gestures have to be combined, namely, the trill conveys tongue predorsum lowering and backing, whereas a high vocoid requires tongue predorsum raising and fronting, see Recasens & Pallarès (1999:144).⁶ These two antagonistic gestures tend to be avoided which results in different outputs, see 2.7. Recasens (1991), describing the lingual configuration in the production of [r], also stresses precise (but not antagonistic) requirements on tongue body positioning. He shows for Catalan that there is a clear difference between a production of an apicoalveolar tap [r] and an apicoalveolar trill [r] in symmetrical VCV sequences in a sense that the latter trill displays lesser V-to-C effects than the tap. Recasens (1999:92) draws the conclusion that the tongue body during the production of [r] is more constrained than in the production of [r].

As far as the aerodynamics of trills are concerned, it has been widely observed that the mechanics of trilling resemble the vibration of vocal folds during voicing, cf. Ladefoged and Maddieson (1996). When the mobile tongue tip reaches the hard palate, the air accumulates behind the constriction and the increasing air pressure separates the tongue tip from the palate. Due to the Bernoulli effect the tongue tip is sucked towards the palate and the cycle automatically repeats.

Ladefoged and Maddieson (1996:217) point out that no muscular action controls each single vibration, but ‘a sufficiently narrow aperture must be created and an adequate airflow through the aperture must occur.’ McGowan (1992) provides a model of a tongue-tip trill production which creates a favourable pressure-flow relation at the tongue tip for sustained vibration. He shows that a trill is the result of interaction between air and the solid tongue tip. What is very important for initiating and sustaining the trill, is the permanent control of the tongue tip, glottal aperture, and the lungs.

Solè (1998, 1999, 2000a,b) who conducted several experiments on the aerodynamics of trills stresses that (voiced) trills require very precise aerodynamic conditions in order to maintain trilling and voicing. On the one hand the oropharyngeal pressure (P_0) has to be high enough in order to produce and maintain trilling and on the other hand it has to be low enough in order not to impair the transglottal flow required for voicing, see Solè (1998:407). Solè

⁶ Recasens & Pallarès (1999) also show that in comparison to [r], the trill is more resistant to the effects from [i].

estimated the allowable range of aerodynamic variation for trills from aerodynamic data. The results show that (P_0) may vary between a narrow range of 5.4 to 4 cm H₂O in order to sustain voicing and trilling, see Solè (2000b:674). If P_0 variations do not fall into this range, the trilling may devoice or cease.

Due to their articulatory and aerodynamic complexity trills are often reported to reduce the number of contacts or lose their trilling and be realized in a different manner. For example, Łobacz (2000) who investigates the production of Polish [r] from an acoustic point of view (with respect to its segmental, spectral and durational properties) reports its very wide variability. She distinguishes four main realization types of [r] which include single taps, single flaps, multiple taps/flaps, and trills. A similar situation is found in Czech and Serbian. Šimáčková (2001) shows that the Czech alveolar trill is mostly realized as a single contact tap. In some cases the trill appeared as an approximated contact or even a vowel. Out of 150 words containing [r], only 2 were realized with a (two-contact) trill. In Serbian, on the other hand, 55 of the 165 words had a two-to-three contact trill. Solé (2002b) states that Catalan trills usually involve four contacts and less often five or six. In Recasens' & Pallarès' study (1999:151), the number of cycles for Catalan trills varies between two and five. Bolla (1981:99) observes three to four cycles for the Russian trills. In sum, the number of cycles in the production of trills varies from two to six which also depends on the syllable position, context, speech rate and style.

As far as the aerodynamics of palatalized trills are concerned the oropharyngeal pressure is even higher than in the case of [r]. This is due to the fronting and raising of the tongue dorsum in the case of a high vocoid. Alternatively, Solè (2000b:663) proposes that the higher P_0 of trills in the [i] context than in that of the [a] context is to be attributed not only to the high tongue position but also to tongue tension which is higher in the [i] context. In consequence, a higher pressure is needed to cause the tongue-tip to vibrate. If the pressure is higher than the narrow range of possible airflow variation, it is exposed to further modifications, which in consequence might lead to the cessation of voicing or trilling.

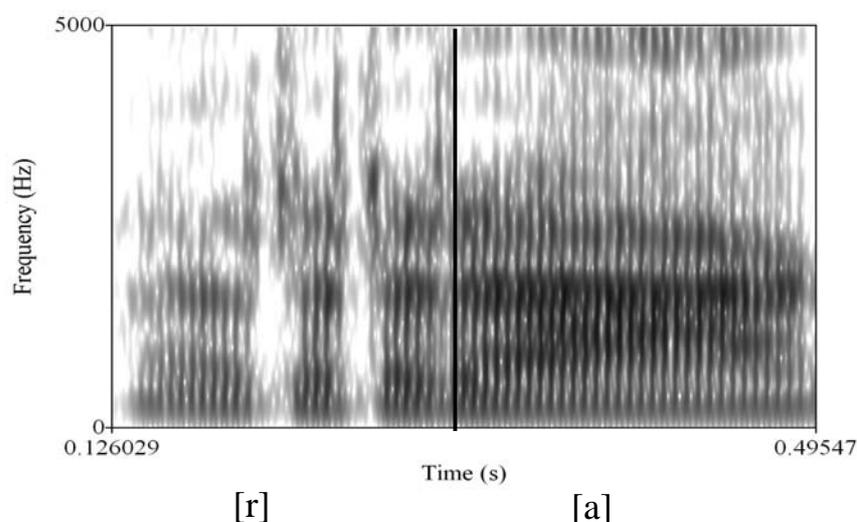
In sum, trills articulated with the tongue tip require very precise aerodynamic conditions in order to maintain trilling and voicing. As soon as a trill appears in the context of the following [i], and furthermore becomes palatalized, its articulation requires more articulatory effort due to the raised and fronted predorsum. Such an articulatory setting demands higher pressure in order to trigger vibration which can disturb the conditions required for voicing and trilling. This indicates that palatalised trills are more marked than their plain counterparts.

2.5 Acoustic and perceptual features

As it has been mentioned in the previous section, a trill consists of cycles of vibration produced by the tongue tip. The number of cycles has essential influence on the acoustic and perceptual characteristics of these sounds.

In (7) a spectrogram of a sequence /ra/ excerpted from a Polish word *rak* ‘crab’ is presented. The trill [r] is separated from a following vowel [a] by a bold line.

(7) Spectrogram of Polish [ra]



The trill in (7) consists of three periods (cycles), each of which is composed by a closed phase and an open phase. In the closed phase the articulators are in contact whereas in the open phase they are slightly apart (Ladefoged & Maddieson 1996:218). The light areas of /r/ visible in the spectrogram in (7) correspond to the initial (closed) phase of a cycle because the formant energy is absent or weak, while dark areas show characteristic formant regions indicating concentrations of energy corresponding to the second (open) phase. The duration of cycles and each phase does not always have to be constant. Generally, the first cycle is longer than the following ones. This is due to the oropharyngeal pressure which is higher at the beginning in order to initiate the trilling than to sustain it, see Solè (1998:406).

In the case presented in (7) the first cycle lasts 80 ms, while the second and the third ones are 50ms long. As far as the latter cycles are concerned, the closed phase of the second and the third cycle amounts to about 20ms. The duration of the closed phase of the first cycle is evidently difficult to measure due to the beginning of the signal, i.e. a word-initial position of [r]. Ladefoged

and Maddieson (1996:218) report on trills where a single cycle also lasts about 50 ms.

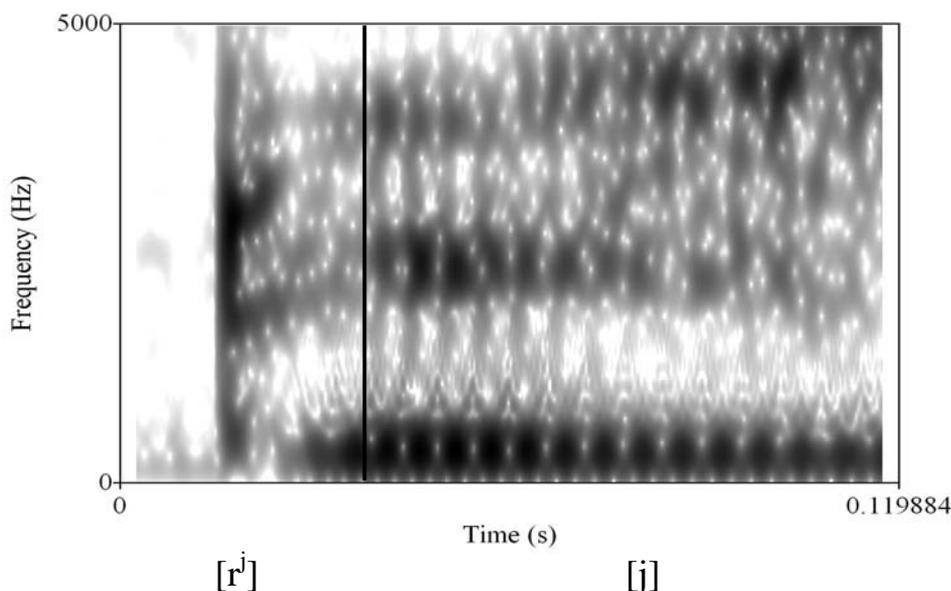
How are the acoustic relations mirrored in the perception of trills? If we assume that one cycle of trilling lasts approximately 50 ms long, then its frequency of vibration is 20 Hz (20 cycles in a second), cf. also Ladefoged and Maddieson (1996:218). Lindau (1985) who investigated trills of Edo, Degema, Ghotuo, Kalabari, Bumo, Spanish and Swedish produced by a total of 25 speakers reports the mean frequency of vibration to be 25 Hz.

What all the analyses cited above have in common is that the frequency of vibration is at least 20 Hz which is a crucial characteristic for perception, namely, the frequency of vibration is a conceivable frequency for a human ear, cf. e.g. Zwicker&Feldtkeller (1967). Importantly, the vibration frequency of 20 Hz in order to be perceived has to have an amplitude of about 80 dB. My preliminary recordings show that this is exactly the case of Polish trills (from 77,8 to 80 dB). Based on this preliminary evidence I assume that the vibration of trilling is perceptually distinguishable. This point obviously requires experimental underpinnings.

As far as the palatalized trills are concerned, their acoustic characteristics considerably differ from that of plain ones. The main difference pertains to the number of cycles. In contrast to plain trills, the palatalized ones display a single cycle of vibration (see Wierzchowska 1971). The cycle is followed by a brief period of frication which overlaps with formants of /j/.

The spectrogram in (8) shows a sequence of [r^j] excerpted from a Polish foreign-word risotto ‘risotto.’ The bold line separates [r^j] from the following [j].

(8) Spectrogram of Polish [r^j]



The spectrogram of [r^j] in (8) shows that there is a single cycle of the trill, i.e., the trill loses its underlying characteristics, i.e. trilling. Instead the first cycle is followed by a brief period of frication (ca. 15ms) which is too short to be perceivable.⁷ This indicates that the palatalised trill is less salient perceptually than the plain one due to the absence of the trilling and the shortness of frication. These relations change as soon as the frication becomes longer and spectrally more similar to the fricative. In such a case the palatalized trill converts to a fricative trill, as in Czech, or to a fricative sibilant, as in Polish, see 2.7.

2.6 *Emergence of trills in the process of language acquisition*

Starting with undoubted cross-linguistic evidence, trills are acquired late in the process of language acquisition in comparison with other sounds, cf. Vihman (1996). In addition, trills are not present in the babbling stage, cf. Stark (1980).

McGowan (1992) states that Spanish trills (besides sibilants) are the final gestures to be mastered by native speakers. In Polish, as reported by Łobacz (1996) and Zarębina (1980), [r] is required by Polish children at the age of four or even later. This sound is often replaced by [l] or [j]. Similarly, in Russian trills belong amongst the last sounds acquired by children.

As far as the acquisition of palatalized sounds is concerned, they are reported to be acquired even later than plain trills which is typical for secondarily palatalized sounds. This has been shown for Polish by Łobacz (1996) and confirmed for Russian by Gagarina (p.c.).

In sum, in the process of language acquisition trills and palatalized trills are acquired late in comparison to other sounds. In contrast to trills, the palatalized ones are acquired even later than their plain counterparts.

2.7 *Stability in diachronic processes*

Trills are widely reported to undergo various changes especially if they are followed by high vocoids.

Several possible outputs of /r/ followed by a high vocoid have been reported, as, e.g., [r^j] (Russian, Lower Lausatian), [rj] (Slovene), [ž] (Polish), or [s] (Jita);⁸ For example, in Jita, a Bantu language, the stem-final /r/ changes to [s] before a glide [j] <y> in the process of causative doubling, as shown in (9),

⁷ Solé (2002) states that the audible friction in the range of 50ms is required for voiced fricatives (and 30 ms for voiceless ones).

⁸ Cf. Banner-Inouye (1995) for other realization of /r/.

cf. Downing (2001), Hyman (1994), Labroussi (1999), Schadeberg (1995), Zoll (1995).

(9) Input	Causative Doubling	
gur-y-a	gus-y-a	‘to sell’
gur-y-ir-a	gus-i:s-y-a	‘to sell to someone’
gur-y-ir-an-a	gus-i:s-y-a:n-y-a	‘to sell to each other’

Slavic languages provide a list of possible developments of [r^j] which emerged from a Proto-Slavic [r] followed by [j] or front vowels. The process is shown in (10), see Rospond (1971:115, 116). Note that the symbols used in (10) are taken from the original source. I assume that <ě> corresponds to the short vowel [e], the front jer <ь> to a short front vowel which is similar to [ɪ], and the nasal <ę> resembles [ẽ].

(10) r → r^j / _ j, i, e, ě, ь, ę

In the thirteenth century, the secondary palatalised [r^j] became spirantized and converted to a fricative trill [ř]. In Polish, this sound did not change its place of articulation until the eighteenth century, cf. Stieber (1966:69). Then the fricative secondary articulation became stronger than the primary articulation. Finally, the palatoalveolar fricative [ʒ] was pronounced which changed into a retroflex [z].⁹

The table in (11) lists the outputs of [r^j] in the present Slavic languages, see Kavitskaya (1997:751). It shows that apart from [r^j], [ř] and [z], the secondary palatalisation was entirely lost, as, e.g., in Belorussian, Croatian, Slovak, Serbian or Macedonian. Finally, the [r^j] changed into a sequence of [rj] as in Slovene.

(11)	Language	Output of /r ^j /
	Russian	+
	Belorussian	–
	Ukrainian	±
	Polish	[z]
	Czech	[ř]

⁹ The motivation for the change /ʒ/ → [z] is discussed in Žygis (2003) in detail.

Slovak	–
Upper Sorbian	±
Lower Sorbian	+
Slovene	[rj]
Croatian	–
Serbian	–
Macedonian	–
Bulgarian	±

- + still present in all environments
 – former palatalisation entirely lost
 ± former palatalisation partially lost

In sum, trills are reported to be unstable especially in one context, i.e., when they are followed by high vocoids. This context often leads to the emergence of the palatalized trills which in turn are highly apt to undergo changes independently of the context in which they occur.

2.8 *Phonotactic distribution*

According to the markedness concept marked segments have a more narrow distribution than their unmarked counterparts in a given language. Hence, we expect a more limited distribution for palatalized trills than for plain ones in the thirteen (Slavic) languages examined below.

As far as plain trills are concerned, general cross-linguistic evidence shows that trills themselves display a limited distribution in comparison to other sounds because they tend to occur close to the syllable nucleus, cf. Lindau (1985), Ladefoged and Maddieson (1996), Wiese (2001). However, the distribution of palatalized trills is even more restricted. This will be illustrated on the basis of phonotactic restrictions of /r/ and /r^j/ occurring in Slavic languages.

The table in **Fehler! Verweisquelle konnte nicht gefunden werden.**) lists the distribution of plain and palatalized trills according to the syllable position in which they occur. The plain trills are not systematized with respect to the onset coda and a consonant cluster because they are attested in all these positions.

(12)

Language	Plain r	Palatalized trills		
		Onset	Coda	Before a consonant
Czech	√	–	–	–
Polish	√	–	–	–
Slovak	√	–	–	–
Upper Sorbian	√	√	–	–
Lower Sorbian	√	√	√	√
Bulgarian	√	√	–	–
Croatian	√	–	–	–
Macedonian	√	–	–	–
Serbian	√	–	–	–
Slovene	√	–	–	–
Belorussian	√	–	–	–
Russian	√	√	√	√
Ukrainian	√	√	–	–

From the table in **Fehler! Verweisquelle konnte nicht gefunden werden.**), it follows that plain trills are present in every Slavic language in onset, coda and before consonants. Palatalized trills occur only in Bulgarian, Russian, Upper and Lower Sorbian, and Ukrainian. However, their presence is limited to certain positions in a syllable. While in Bulgarian, Upper and Lower Sorbian, Russian, and Ukrainian palatalized trills are found in the onset, their presence in the coda is only attested in Lower Sorbian and Russian. Palatalized trills can also be found in a preconsonantal position, i.e., in consonant clusters. Among Slavic languages, there are only two languages, Russian and Lower Sorbian, in which palatalized trills occur in consonant clusters.

This distribution provides convincing evidence that the palatalized trills are marked in comparison to plain ones in Slavic languages. It also implies markedness positions of palatalized trills: the unmarked place for the realization of a palatalized trill is the syllable onset. A more marked position is a coda and the most marked is a preconsonantal position. The hierarchy of markedness position is presented in (13).¹⁰ The symbol ‘<’ means ‘less marked than.’

- (13) Positions of markedness of palatalized trills:
onset < coda (word-final) < consonant clusters

¹⁰ The hierarchy in (13) also holds true for a diachronic development of palatalisation of labials discussed in Rochoń (1999). It has been shown that the secondary palatalisation of labials disappeared in consonant clusters earlier than in word-final positions but is still attested in the onset.

The hierarchy in (13) is also motivated phonetically. Palatalisation as a secondary articulation is realized in an optimal way when it is followed by a vocalic element, see e.g. Ladefoged and Maddieson (1996). Such an optimal condition is offered by the onset position. A direct perceptual consequence of the presence of a vocalic element is that palatalisation is perceptually much more salient than in a preconsonantal position. Indeed, as the data in (13) indicate the least appropriate position for the realization of palatalisation occurs when a palatalized trill is followed by a consonant, i.e. in a consonant cluster.

A word-final position seems to be more optimal for the realization of the palatalized trills than the preconsonantal position. The distribution of /r^j/ in Lower Sorbian and Russian shows that the presence of a palatalized trill in a consonant cluster indicates its presence in word-final position but not vice versa. However, from a phonetic point of view it is not exactly clear why this is the case. One of the reasons could be that a secondarily palatalized consonant occurring word-finally might be followed by an initial vocalic element of the next word in connected speech. In the case of consonant clusters no such option is available. Another -more probable reason- could be that in the coda position [r] contrasts with [r^j] and therefore the neutralization of the latter sound is blocked.

Finally, it is worth noting that trills also occur as syllable nuclei. For example, in Czech the trill functions as a syllable nucleus, e.g. *mrk* ‘a wink’, *mr.kat* ‘to wink’, *smrk* ‘spruce tree’, see Šimáčková (2003). Similarly, the syllabic /r/ is found in Slovak, Croatian, Serbian and Macedonian (Dalewska-Greń 2002:88). However, the comparison of trills with their palatalized counterparts within these languages is not possible because [r^j]s are not present in their inventories. I am also not aware of languages having /r^j/ in a nucleus position.

In sum, the phonotactic distribution of [r] and [r^j] in the Slavic languages shows that the latter segment is more narrowly distributed in comparison to the former and therefore can be considered as its marked counterpart.

2.9 Implications

If palatalized trills are marked sounds, they are expected to imply the presence of the plain trills in the phonemic inventories.

In the languages I examined there are no exceptions concerning this point. Languages having [r^j] in their phonemic inventories include also the plain trill [r], cf. languages listed in UPSID as Bulgarian, Nenets, Lithuanian, Russian, Saami as well as Daur (Wu 1996), Muinane (Walton&Walton 1967), Polabian (Polański 2002), Upper and Lower Sorbian (Schuster-Šewc 1996), Toda (Spajić, Ladefoged & Bhaskararo 1996) and Ukrainian (Zilyns’kyj 1979).

Conversely, the plain trill does not imply the presence of the palatalized counterpart which is confirmed by many languages in inventories of which an /r/ and not /r^j/ is attested. In UPSID 150 languages of this type are included.

2.10 Summary

The evaluation of plain and palatalized trills against several dimensions of linguistic description leads to the conclusion that palatalized trills are marked pendants of plain trills. The [r]s are less frequent than [r^j]s. The latter are also articulatorily and especially aerodynamically more complex as well as less salient from a perceptual point of view than [r]s. They are also acquired later in the process of language acquisition and are less stable from a diachronic point of view than plain trills. From the point of view of the phonotactics, [r^j]s are more narrowly distributed than [r]s and finally their presence implies the presence of [r]s in a given language.

A direct theoretic contribution to the concept of markedness is that in order to obtain a complete picture of (un)markedness of trills, aerodynamic, acoustic and perceptual characteristics of the sounds have to be taken into consideration.

The perceptual features of trills also resolve a potential problem arising from the comparison of trills with other natural classes, as e.g. stops or fricatives. It is far from obvious why trills despite their articulatory complexity occur fairly frequently in the languages of the world. This question can be answered only if the perceptual saliency of trills are considered.

3 A phonological account of r-palatalisation in Polish and Czech

In the remaining part of the article I will show how the phonetic characteristics of trills and especially their palatalised counterparts are mirrored in the phonology of two Slavic languages, namely, in Polish and Czech. The two languages differ in the output of palatalisation of /r/: in Polish /r/ alternates with [z], while in Czech the output of r-palatalisation is [ř]. What both languages have in common is that palatalisation processes in which /r/ alternates are not predictable but instead are triggered by a limited number of suffixes, as will be shown below.

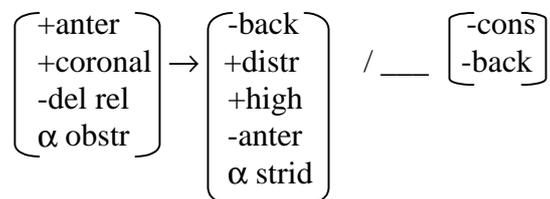
3.1 Data ¹¹

In Polish /r/ undergoes Coronal Palatalisation, a rule which converts //s z t d n ł

¹¹ A part of this analysis is presented for Polish in my dissertation (Rochóń 2000).

r//¹² into alveolopalatals /ś ź t^ś d^ź ń l' ř'/.¹³ The process is shown in (14), see Rubach (1984).

(14) Coronal Palatalisation

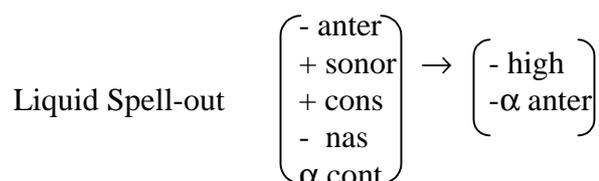


Coronal Palatalisation presented in (14) does not generate the surface outputs found in Polish, see (15a), but only intermediate and nonsurface attested steps, see (15b).

- (15) a) ko//r//a ko[ž]e ‘bark’ nom.sg./loc.sg.
 ko//ʎ//o ko[l]e ‘wheel’ nom.sg./loc.sg.
 b) ko//r//a ko ř'/e ‘bark’ nom.sg./loc.sg.
 ko//ʎ//o ko/l'/e ‘wheel’ nom.sg./loc.sg.

Since neither ř'/ nor l'/ are attested outputs of //r// and //ʎ// palatalisation in Polish, Rubach (1984) proposes the so-called Liquid Spell-out rule, see (16), that converts l'/ to [l] and [ř'] to [ř], see Rubach (1984:199).

(16) Liquid Spell-out rule



The rule in (16) would be sufficient in accounting for the data (see examples

¹² Rubach (1984) uses double slashes for marking the underlying representation and single ones for a derivational representation. He also uses traditional Slavic symbols as, e.g., [ž] for the retroflex [z] and ['] for the secondary palatalization. These conventions will be adopted for presenting Rubach's analysis.

¹³ Another rule accounting for the palatalisation of anterior consonants (the so-called I-Anterior Palatalisation) was proposed earlier by Gussmann (1978:86). This rule converts the coronals //s z t d n l r// into their palatalized counterparts /s' z' t' d' n' l' r'/, which subsequently have to undergo various Spell-out rules such as Obstruent Spell-out /s' z' t' d'/ → [ś ź t^ś d^ź], Nasal Spell-out /n/ → [ń], r-Spell-out /r'/ → [ř'], and Lateral Spell-out /l'/ → [l].

given below). However, [ř] does not occur in the Polish inventory and it not an attested output of r-palatalisation. Therefore in order to get an optimal candidate Rubach (1984) proposes the so-called r-spell-out rule which turns /ř/ to [r] before a consonant and to [ž] otherwise. This is shown in (17), see Rubach (1984:200).

(17) r-Spell-out rule

$$\text{r-spell-out} \left(\begin{array}{l} +\text{sonor} \\ +\text{cons} \\ -\text{anter} \\ -\text{high} \end{array} \right) \rightarrow \begin{array}{l} [+anter] / _ C \quad (a) \\ [+obstr] \quad (b) \end{array}$$

The first part of the rule in (17) accounts for alternations, e.g. gó//r//a-gó[ž]e ‘mountain’, sg./pl., see also examples in (18a). The second part of this rule has been postulated in order to account for data in which //r// is realized as [r] before palatalizing suffixes starting with consonants. The intermediate stage /ř/ undergoes depalatalisation. Examples in (18) show stems with palatalizing suffixes such as -c+a [tsa], -ł+a [wa], -n+y [ni].

- (18) dwo//r//+c+a dwo[rt̪s̪]a ‘station’ nom.sg./gen.sg.
o//r//+ł+a o[rw]a ‘eagle’ nom.sg./gen.sg.
o//r//+n+y o[rn]y ‘to plough’/‘arable’ adj. nom.sg.

According to the rules presented in (14), (16), and (17), the underlying /r/ in (18) undergoes some derivational steps, as shown in (19).

(19) //r// → /ř/ → /ř/ → [r]

It has to be stressed that there are also cases in which [ž] occurs on the surface before the palatalizing suffixes shown above. Examples presented in (20) show that /ž/ is present underlyingly.

- (20) nos+o+ro//ž//+c+a nosoro[žt̪^s]a ‘rhino’ gen.sg.
tę//ž//+c+a tę[žt̪^s]a ‘tetanus’ gen.sg.
wy//ž//+ł+a wy[žw]a ‘pointer’ gen.sg.
oreę//ž//+n+y oreę[žn]y ‘armed’ adj.nom.sg.

The examples in (15), (18) and (20) show why Coronal Palatalisation converts /r/ into /ř/ followed by an additional Spell-out rule, i.e., /r/ → /ř/ → [r]. Rubach (1984:72) motivates this step in the following way:

“If Coronal Palatalisation changed /r/ to [ř] directly, there would be no way of recovering /r/ in the contexts where it undergoes depalatalisation. Thus, for instance, *wierny* ‘faithful’ and *podróżny* ‘journey’ (Adj.) would have the same structure after Palatalisation had applied: /ž+in/”

As far as Czech is concerned there are also processes attested in which /r/ is palatalized. In contrast to Polish, the output of these processes is the fricative trill [ř]. But similarly to Polish, the palatalisation output is not predictable and it is triggered by a limited number of suffixes. In (21) alternations /r/ - [ř] in some morphological contexts are shown: formation of adjectives (a), plural formation in masculine animate declension (b), formation of the 1 ps. sg. from infinitives (c), formation of a possessive adjective (d), formation of the locative singular case in masc., fem. and neut. (e), see Scheer (2001:16).

(21) a)	vyd/r/+í	vyd[ř]í	‘otter’ adj.sg.
b)	dob/r/+í	dob[ř]í	‘good’ 1 ps.pl.masc.
c)	o/r/+u	o[ř]u	‘to plough’ 1ps.sg.pres
d)	Vě/r/+in	Vě[ř]in	‘Věra’ adj.sg.
e)	papí/r/e	papí[ř]e	‘paper’ loc.sg.

Another /r/-[ř] alternation context is the vocative case of masculine nouns. In (22) there are examples in which /r/ in the nominative of masculine nouns alternates with [ř] in the vocative, see Scheer (1998:56). It has to be stressed that this alternation occurs only in cases where the trill is preceded by an obstruent, cf. (22a) and not by vowels, cf. (22b).

(22)	Nominative	Vocative	
a)	kmot/r/	kmot[ř]e	‘godfather’
	svet/r/	svet[ř]e	‘pullover’
b)	docto/r/	docto[r]e	‘doctor’

3.2 An Analysis

Since in Czech and Polish palatalisation of /r/ is not predictable I propose that there is a palatalizing feature called [PAL]¹⁴ that is a property only of those

¹⁴ [PAL] is to be considered as a shorthand for a palatalizing feature, cf. also the discussion on the palatalizing feature by Hall (1997a:82ff).

suffixes that induce palatalisation.¹⁵ In (23a) palatalizing Polish suffixes are shown with their representation, while in (23b) representation of Czech palatalizing suffixes is displayed.

(23) a) Polish

_ ^[PAL] ist+i	adjectival
_ ^[PAL] isk+o	augmentative
_ ^[PAL] ic+a	feminine
_ ^[PAL] in+a	pejorative
_ ^[PAL] i	nom.pl.masc.

b) Czech

_ ^[PAL] u	1 ps. sg. pres.
_ ^[PAL] í	nom. pl. masc. animate declension
_ ^[PAL] in	possesive, adj.
_ ^[PAL] í	nom. adj.
_ ^[PAL] e ¹⁶	loc.sg.

In the following I present an analysis of Polish and Czech data by means of constraints in the framework of Optimality Theory, see McCarthy and Prince (1993), Prince and Smolensky (1993).

In (24) a faithfulness constraint MAX-IO is shown which assures that no segment will be deleted from the underlying representation, see McCarthy and Prince (1995: 264).

(24) MAX-IO

Every segment in the input has a correspondent in the output.

In order to assure that not only segments but also subsegments such as [PAL] will not be deleted, a constraint called MAX (Subseg) has to be involved in the analysis. The general formulation of MAX (Subseg) is given in (25a) (see Zoll 1996:59) and its specification in (25a).

¹⁵ It is important to stress that a floating feature might also project its own root node. For example, in Yowlumne a floating feature either moves to find a suitable docking site in the base or projects its own root node, cf. Zoll (1996:167ff). There are also cases in Polish where a floating feature might emerge as a segment. Traditionally such segments were represented as underlying abstract vowels called jers, which either disappear or are vocalized. Recent studies show that it is possible to represent jers in terms of features, cf. e.g. Gussmann (1992), Szpyra (1992), Zoll (1993), Zoll (1996).

¹⁶ I assume that stems which do not undergo palatalisation are additionally specified in the underlying representation, cf. (22). In the present paper I will not analyze such cases.

- (25) a. MAX (Subseg): Every subsegment in the input has a correspondent in the output.
 b. MAX [PAL]: Every [PAL] in the input has a correspondent in the output.

The phonetic content of [PAL] will not be specified because its application leads to different outputs in different languages. I assume that there is a range of possible outputs of r-palatalisation (ř, ř^j, r^j, r̥^j, ž, ž^j), and MAX [PAL] is satisfied if the featural content of /r/ converts to one of the possible outputs, cf. also the discussion below.

Constraints do not only prohibit deletion of [PAL] but they also have to assure that this subsegment is realized at the appropriate segment, i.e., on the stem-final consonant. Ellison (1995) and Zoll (1996) propose a constraint called No-Intervening, presented in (26), which prohibits intervening material between a category <ρ> (segment, subsegment) and an edge <E> that can refer to a morphological and prosodic category. The constraint is presented in (26).

- (26) No-Intervening (ρ; E; D) There is no material intervening between ρ and an edge E in domain D, cf. Ellison (1995:2), Zoll (1996:108).

In Polish and Czech, the constraint ‘No-Intervening’ demands [PAL] to be realized at the right edge of a stem, i.e. the stem-final consonant. Its specific formulation is given in (27).

- (27) No-Intervening ([PAL]; Right; Stem)

The interaction of No-I([PAL]; R,S) and MAX [PAL] is given in (28).

(28)

or+ ^[PAL] u	No-I([PAL];R,S) ; MAX[PAL]	
oř ^[PAL] u		
oru		*!
o ^[PAL] ru	*!	
o ^[PAL] ř ^[PAL] u	*!	

The first candidate in (28) is selected as the most optimal one since the palatalizing subsegment [PAL] docks on the final-stem consonant. The second candidate violates MAX [PAL] since it deletes the underlying [PAL]. However, it does not incur any violation of No-I([PAL]; R,S) by satisfying it vacuously. Although the third and the fourth candidates satisfy MAX [PAL] by not deleting the palatalizing subsegment [PAL], they are excluded from consideration because they violate No-I([PAL];R,S) by letting [PAL] skip the stem-final

consonant. Note the final candidate displays two palatalizing subsegments which was possible by insertion of [PAL].

In a similar way the constraints interact in Polish, which is illustrated by an example in (29).

(29)

kor+ ^[PAL] e	No-I([PAL];R,S) ; MAX[PAL]	
☞ koř ^[PAL] e		
kore		*!
k ^[PAL] ore	*!*	
k ^[PAL] or ^[PAL] u	*!*	

The process of palatalisation in Polish is similar to that of Czech in a sense that a candidate who realizes a subsegment [PAL] on its stem-final consonant is selected as optimal. This is illustrated by the first candidate in (29). Although the mechanism is similar there is a crucial difference between Polish and Czech, namely, the r-palatalisation outputs. The tableaux in (28) and (29) show the optimal candidates in terms of their representation but not in terms of their realization ([PAL] is still present in the surface representation).

If we assume that the palatalizing feature [PAL] contains the same features in Polish and Czech we will obtain the same palatalisation results, e.g., either [ř] or [ž]. In order to avoid such an undesired situation it is assumed that the realization of [PAL] leads to a range of possible outputs of r-palatalization shown in (30). The selection of optimal outputs is regulated by the markedness constraints differently ranked in individual languages, cf. the analysis below. Note the for the sake of convenience I dispense with the voiceless counterparts of these segments. However, there are languages in which voiceless sibilants are output of the palatalisation, cf. examples in Jita shown in (9).

- (30) ř, ř^j, r^j, r̥^j, ž, ž^j

All the segments are listed in terms of features in (31).¹⁷ Segments which are palatalized in terms of secondary articulation, i.e., ř^j, r^j, r̥^j, ž^j are specified as [-back, +high]. Two other segments ř and ž specified as [0 back, 0 high] are neither palatalized nor velarized.

¹⁷ [-back, +high]= palatalized, [+back]=velarized, [0 back, 0 high]= neither palatalized nor velarized.

(31)

	ř	ř ^j	ɹ ^j	r ^j	ž	ž ^j
[cor]	+	+	+	+	+	+
[ant]	+	+	-	+	-	-
[dist]	+	+	-	-	-	-
[back]		-	-	-		-
[high]		+	+	+		+
[son]	-	-	+	+	-	-
[strident]	+	+	-	-	+	+
[rhotic]	+	+	+	+	-	-

In (32) the segments are formulated as markedness constraints. They are limited to three constraints relevant for the analysis of r-palatalisation in Polish and Czech.

- (32) * [ř]: [ř] is avoided.¹⁸
 * [ž^j]: [ž^j] is avoided.
 * [ž]: [ž] is avoided.

The constraints in (32) express the avoidance of the given sounds due to their articulatory complexity: [ř] is a laminal trill in which the tongue blade vibrates, whereas [ž] and [ž^j] are retroflex sounds. The palatalisation of retroflex sounds has been argued to be avoided, see Hall (2000), or even impossible, see Hamann (2003).

The interaction of the constraints in (32) together with No-I([PAL];R, S) and MAX [PAL] is illustrated in (33) for Polish.

(33)

ko/r/+ ^[PAL] e	* [ř]	* [ž]	No-I([PAL];R)	MAX [PAL]	* [ž]
ko[ř]e	*!				
ko[ž ^j]e		*!			
☞ ko[ž]e					*
ko[r]e				*!	

The candidate *ko[ž]e* is selected as optimal, because the feature [PAL] is realized on the stem-final consonant and it does not violate the high-ranking constraints.

¹⁸ (a) * [-dist, -back] (* [ž'])
 (b) * [+dist, +rhotic] (* [ř])
 (c) * [-dist, -ant] (* [ž])

In Czech the constraint ranking looks differently. In (34) the selection of the candidate *dobří* is shown in (34).

(34)

dob/r/+ ^[PAL] í	*[žʲ]	*[ž]	No-I([PAL];R)	MAX[PAL]	*[ř]
☞ dob[ř]í					*
dob[žʲ]í	*!				
dob[ž]í		*!			
dob[r]í				*!	

The first candidate in (34) is selected as optimal. This is due to the low-ranking markedness constraint prohibiting [ř].

If we take into consideration Polish r-palatalisation in cases when /r/ is followed by consonant-initial suffixes, the ranking proposed thus far for Polish selects a candidate that is unattested in Polish. This is illustrated by the tableau in (35).

(35)

wie/r/+PALn+y	*[ř]	*[žʲ]	No-I([PAL];R)	MAX[PAL]	*[ž]
wie[ř]ny	*!				
wie[žʲ]ny		*!			
☞ wie[ž]ny					*
wie[r]ny				*!	

The /r/ in the underlying representation *wie/r/+PALn+y* undergoes palatalisation and the candidate *wie[žʲ]ny* is selected as optimal. However, this candidate proves false since it is not attested in Polish. The attested candidate *wie[r]ny* is not selected because it violates MAX [PAL] by rejecting palatalisation.

A constraint that is responsible for the selection of the optimal candidate is shown in (36).

(36) Faith (rC) /r/ is not changed when followed by a consonant.

The constraint Faith (rC) prevents any changes of the feature context of /r/ if it is followed by another consonant. It does not allowed /r/ to be palatalized either by a secondary articulation or by changing the place of articulation. A phonetic grounding of this constraint is that a preconsonantal position is the least appropriate position for palatalisation of trills.

In the ranking for Polish Faith (rC) is ranked very high since there is no alternation of /r/ if it is followed by a consonant. Therefore only the candidate that is faithful to the underlying representation proves as optimal. Consider the tableau in (37).

(37)

wie/r/+PALn+y	Faith (rC)	*[ř]	*[žʲ]	No-I([PAL];R)	MAX[PAL]	*[ž]
wie[ř]ny	*!	*!				
wie[žʲ]ny	*!		*!			
wie[ž]ny	*!					*
☞ wie[r]ny					*	

The last question concerns palatalisation of retroflexes in Polish when they are followed by palatalising consonant-initial suffixes. In (38) it is shown that a markedness constraint *[žʲ] prohibiting palatalized retroflexes is responsible for the selection of the optimal candidate.

(38)

podró/ž/+PALn+y	*[žʲ]	No-I([PAL];R)	MAX[PAL]	*[ž]
☞ podró[ž]ny			*	
podró[žʲ]ny	*!			

Alternatively, one could invoke faithfulness constraints that are responsible for the stability of flat retroflexes in Polish. The constraints are shown in (39).¹⁹

(39) Ident (+back, -anterior, -distributed): Faith (š ž t^š d^ž)

The application of Faith [ž] is shown in (40). The attested candidate *podró[ž]ny* is selected as optimal.

(40)

podró/ž/+PALn+y	Faith [ž]	No-I([PAL];R)	MAX[PAL]	*[ž]
☞ podró[ž]ny			*	*
podró[žʲ]ny	*!			
podró[rʲ]ny	*!			
podró[r]ny	*!			

To sum up, it has also been shown that the same universal phonetically grounded constraints ranked differently in Polish and Czech select attested candidates in both languages.

The analysis proposed for Polish also shows that accounting for palatalisation by means of constraints seems to be more economic than by rules since constraints require neither an intermediate stage, i.e. an abstract segment /ř'/ nor additional Spell-out rules. Constraints alone make a distinction between the treatment of an underlying /ž/ and an underlying /r/, a potential input for palatalisation.

¹⁹ Faithfulness constraints that are high-ranking in Polish include:
 Ident (+high, -anterior, -back, +distributed): Faith (š ž t^š d^ž ń)
 Ident (+back +anterior, -distributed, +del.release): Faith (t^š d^ž)

4 Summary

The evaluation of trills [r] and their palatalised counterparts [r^j] from the point of view of markedness has shown that despite the articulatory and aerodynamic complexity of [r], it is an unmarked counterpart of [r^j]s. This follows from the evaluation of several aspects of linguistic markedness including aerodynamic, acoustic and perceptual evidence, mostly neglected when discussing (phonological) markedness relations.

The phonetic features of [r^j] have been incorporated into universal phonetically grounded constraints employed for a phonological OT-analysis of r-palatalisation in Polish and Czech.

For Polish, it has been shown that in contrast to a derivational account of palatalisation as proposed by Rubach (1984) some abstract unmotivated intermediate stages can be avoided with the help of faithfulness constraints, which due to their highly-ranked status in Polish prohibit the palatalisation of retroflexes and disallow the palatalisation of /r/ if it is followed by another consonant.

The analysis proposed for r-palatalisation in Polish is more economic than by rules since it requires neither an intermediate stage, i.e., an abstract segment /r̥/ nor additional Spell-out rules. Constraints alone make a distinction between the treatment of an underlying /ž/ and an underlying /r/, two potential inputs for palatalisation.

5 References

- Bailey, C.-J. N. (1973). *Variation and Linguistic Theory*. Arlington: Center for Applied Linguistics.
- Banner-Inouye, S. (1995). *Trills, Taps and Stops in Contrast and Variation*. PhD thesis. Los Angeles, University of California.
- Bhat, D.N.S. (1973). Retroflexion: an areal feature. *Working Papers on Language Universals*. 27-67.
- Boersma, P. (1998). *Functional Phonology. Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.
- Bolla, K. (1981). *A Conspectus of Russian Speech Sounds*. Budapest: Akademiai Kiado.
- Chomsky, N & M. Halle (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Dalewska-Greń, H (2002). *Języki słowiańskie*. Warszawa: Wydawnictwo Naukowe PWN.
- Delattre, P. (1971). Pharyngeal Features in the consonants of Arabic, German, Spanish, French, and American English. *Phonetica* 23: 129-155.
- Downing, L. J. (2001). Jita Causative Doubling and Paradigm Uniformity. Paper presented at Conference on Paradigm Uniformity, ZAS, March 9-10.
- Ellison, T. M. (1995). *Phonological Derivation in Optimality Theory*. Ms, University of Edinburgh.

- Flemming, E.S. (1995). Auditory representations in phonology. Ph.D. dissertation, UCLA. [Published (2002). New York: Garland]
- Gussmann, E. (1992). Back to front: Non-linear palatalisation and vowels in Polish. In: Fisiak, J. and S. Puppel (eds.) (1992). *Phonological Investigations*. Amsterdam: Benjamins. 5-66.
- Hall, T.A. (1997a). *The Phonology of Coronals*. Amsterdam: Benjamins.
- Hall, T.A. (1997b). The Historical Development of Retroflex Consonants in Indo-Aryan. *Lingua* 101: 203-221.
- Hall, T.A. (2000). Typological generalizations concerning secondary palatalisation. *Lingua* 110:1-25.
- Hamann, S. (2003). *The Phonetics and Phonology of Retroflexes*. Utrecht: LOT.
- Hardcastle, W. J. and N. Hewlett (1999). *Coarticulation. Theory, Data and Techniques*. Cambridge: Cambridge University Press.
- Hume, E. & K. Johnson (eds.) (2001). *The Role of Speech Perception Phenomena in Phonology*. San Diego, CA: Academic Press.
- Hyman, L.M. (1994). Cyclic phonology and morphology in Cibemba. In: Cole, J. & Ch. Kisseberth (eds) *Perspective in Phonology*. CSLI, 81-112.
- Kavitskaya, D. (1997). Aerodynamic constraints on the production of palatalised trills: the case of Slavic trilled [r]. In: *Proceedings of Eurospeech 1997*, vol. 4 . 751-754.
- Labrousse, C. (1999). Vowel systems and spirantization in S.W. Tanzania. In J.-M. Hombert & L. M. Hyman (eds). *Bantu Historical Linguistics: Theoretical and Empirical Perspectives*. CSLI, 335-377.
- Ladefoged, P. (1975). *A course in phonetics*. (1st edition). New York: Harcourt Brace Javanovich.
- Ladefoged, P. & I. Maddieson (1996). *The Sounds of the World's Languages*. Cambridge, Mass.: Blackwell.
- Lindau, M. (1978). Vowel features. *Language* 54: 541-563.
- Lindau, M. (1985). The story of /r/. In: Fromkin, V.A. (ed.) *Phonetic Linguistics*. Orlando. Academic Press. 157-168.
- Łobacz, P. (1996). *Polska fonologia dziecięca*. Warszawa: Energeia.
- Łobacz, P. (2000). The Polish Rhotic. A Preliminary Study in Acoustic Variability and Invariance. *Speech and Language Technology* 4: 85-101.
- Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.
- McCarthy, J. J. & A. S. Prince (1993a). Prosodic Morphology I. Constraint Interaction and Satisfaction. Ms. University of Massachusetts and Rutgers University.
- McCarthy, J. J. & A. S. Prince (1995). Faithfulness and Reduplicative Identity. In: J. N. Beckman, L. W. Dickey & S. Urbanczyk (eds.). *University of Massachusetts Occasional Papers* 18: 249-372.
- McGowan, R.S. (1992). Tongue-tip trills and vocal-tract wall compliance. *Journal of The Acoustic Society of America* 91: 2903-2910.
- NíChiosáin, M. & J. Padgett (2001). Markedness, segment realization, and locality. In: Lombardi, Linda (ed.) *Segmental Phonology in Optimality Theory*. Cambridge: Cambridge University Press.
- Padgett, J. (1997). Perceptual distance of contrast: vowel height and nasality. In Walker, R., Katayama, M. & D. Karvonen (eds). *Phonology at Santa Cruz* 5, 63-78.

- Passy, P. (1891). *Etude sur les changements phonétiques et leurs caractères généraux*. Paris: Librairie Firmin – Didot.
- Polański, K. (2002). Polabian. In: Comrie, B. & G.G. Corbett (eds.). *The Slavonic Languages*. London: Routledge. 795-824.
- Recasens, D. (1991). On the production characteristics of apicovelar taps and trills. *Journal of Phonetics* 19: 267-280.
- Recasens, D. (1999). Lingual coarticulation. In: Hardcastle, W. and Hewlett, N. (eds.). *Coarticulation. Theory, Data and Techniques*. Cambridge: Cambridge University Press. 80-104.
- Recasens, D. (2001). The Effect of Syllable Position in Consonant Reduction. Ms. Paper presented at Conference on Phonetics-Phonology Interface, Berlin, October 2001.
- Recasens, D. & M. D. Pallarés (1999). A study of /r/ and /r/ in the light of “DAC” coarticulation model. *Journal of Phonetics*, 27: 143-169.
- Rochoń, M. (1999). Application of Universal Constraints in Polish: Palatalized Labials. *Linguistica Silesiana*, 31-59.
- Rochoń, M. (2000). *Optimality in Complexity: The Case of Polish Consonant Clusters*. Berlin: Akademie Verlag.
- Rospond, S. (1971). *Gramatyka historyczna języka polskiego*. Warszawa: PAN.
- Romportl, M. (1967). R und das tschechische Konsonantensystem. In: *Acta Universitatis Carolinae – Philologica* 6. *Phonetica Pragensia*, 7-27.
- Rubach, J. (1984). *Cyclic and Lexical Phonology: The Structure of Polish*. Dordrecht: Foris.
- Schadeberg, T. (1995). Spirantization and the 7-to-5 vowel merger in Bantu. In: Dominicy, M. & D. Demolin (eds). *Sound Change*. Amsterdam: John Benjamins. 73-84.
- Scheer, T. (1998). A unified model of Proper Government. *The Linguistic Review*, 15: 41-67.
- Scheer, T. (2001). Čestina ve Wařsawě. Instytut Anglistyki Uniwersytet Warszawski. Ms.
- Šimáčková, S. (2001). Trilled r in Czech and Serbian. Ms. Palacky University Olomouc.
- Šimáčková, S. (2003). Variability in the realization of Czech trilled r. In: Kosta, P., Blaszczyk, J., Frasek, J., Geist L. & M. Zygis (eds.). *Investigations into Formal Slavic Linguistics*. Berlin: Peter Lang Verlag. 119-135.
- Skalozub, L.G. (1963). *Palatogrammy i Rentgenogrammy Soglasnych Fonem Russkogo Literaturnogo Jazyka*. Kiev: Izdatelstvo Kievskogo Univesiteta.
- Solé, M. (1998). Phonological Universals: Trilling, Voicing, and Frication. In: *Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society*, 403-416.
- Solé, M. (1999). Production requirements of apical trills and assimilatory behavior. In: Ohala, J. Y Hasegawa, M. Ohala, D. Granville & A. C. Bailey (eds.)
- Solé, M. (2002a). Assimilatory processes and aerodynamic factors. In: Gussenhoven, C. & N. Werner (eds). *Papers in Laboratory Phonology*: Berlin: de Gruyter. 351-386.
- Solé, M. (2002b). Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics* 30, 655-688.
- Schuster-Šewc, H. (1996) *Grammar of the Upper Sorbian Language*. München: Lincom Europa.
- Spajić, S., Ladefoged, P. & P. Bhaskararo (1996). The trills of Toda. *JIPA* 26: 1-21.
- Stark, R.A. (1980). Stages of speech development in the first year of life. In: Yenikomshian, G. , Kavanagh, J.F. & C.A. Ferguson (eds.) *Child Phonology 1: Production*. New York: Academic Press.

- Stampe, D. (1972). *How I Spent my Summer Vacation*. Ph.D dissertation, Ohio State University.
- Stieber (1966). *Historyczna i współczesna fonologia języka polskiego*. Warszawa: PAN.
- Szpyra, J. (1992). Ghost Segments in Nonlinear Phonology: Polish Yers. *Language*, 68 (2): 277-312.
- Trubetzkoy, N. S. (1931). Die phonologischen Systeme. *Travaux du Cercle Linguistique de Prague*, 4: 96-116.
- Van de Velde, H. & van Hout, R. (eds.) (2001). *'r-atics. Sociolinguistic, phonetic and phonological characteristics of /r/*. Bruxelles: Etudes and travaux.
- Vennemann, T. (1983). Causality in language change: Theories of linguistic preference as a basis for linguistic explanations. *Folia Linguistica Historica*, 6: 5-26.
- Vennemann, T. (1988). *Preference Laws for Syllable Structure and the Explanation of Sound Change. With special Reference to German, Germanic, Italian, and Latin*. Berlin: de Gruyter.
- Vennemann, T. (1989). Language change as language improvement. In: Orioles, V. (ed.). *Modelli esplicativi della diacronica linguistica*. Pisa: Gardini, 11-35.
- Vihman, M.M. (1996). *Phonological development*. Cambridge: Blackwell.
- Walsh-Dickey, L. (1997). *The phonology of liquids*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Walton, J. & J. Walton (1967). Phonemes of Muinane. In: Waterhouse, V. (ed.) *Phonemic systems of Columbian languages*, 37-47.
- Wierzchowska, B. (1971). *Wymowa polska*. Warszawa: Państwowe Zakłady Wydawnictw Szkolnych.
- Wiese, R. (2001). The phonology of /r/. In: Hall, T.A. (eds.) *Distinctive Feature Theory*. Berlin: Mouton de Gruyter. 335-368.
- Wu, C. (1996). *Daur*. München: Lincom Europa.
- Wurzel, W.U. (1994). *Grammatisch initiiertes Wandel*. Bochum: Brockmeyer.
- Wurzel, W.U. (1998). On Markedness. In: *Theoretical Linguistics*, 24: 153-71.
- Zarębina, M. (1980). *Język polski w rozwoju jednostki. Analiza tekstów dzieci do wieku szkolnego. Rozwój semantyczny języka dziecka*. Kraków: Wydawnictwo Naukowe Wyższej Szkoły Pedagogicznej.
- Zilys'kyj, I. (1979). *A phonetic transcription of the Ukrainian Language*. Cambridge, Mass.: Harvard University Press.
- Zoll, Ch. (1993). Ghost segments and Optimality. *The Proceedings of the Twelfth West Coast Conference on Formal Linguistics*, 183-199.
- Zoll, Ch. (1995). Consonant mutation in Bantu. *Linguistic Inquiry*, 26: 536-545.
- Zoll, Ch. (1996). *Parsing below the Segment in a Constraint-Based Framework*. Ph.D. dissertation, University of California. ROA-143-0996.
- Zwicker, E. & R. Feldtkeller (1967). *Das Ohr als Nachrichtenempfänger*. Stuttgart.
- Żygis, M. (2003). The Role of Perception in Slavic Sibilant Systems. In: Kosta, P., Blaszcak, J., Frasek, J., Geist, L. und M. Żygis (eds.). *Investigations into Formal Slavic Linguistics*. Berlin: Peter Lang Verlag. 137-154.

Prosody and Information Structure in Chichewa*

Laura J. Downing

ZAS Berlin

Al Mtenje

University of Malawi

Bernd Pompino-Marschall

Humboldt-Universität Berlin

This paper presents preliminary results of a phonetic and phonological study of the Ntcheu dialect of Chichewa spoken by Al Mtenje (one of the co-authors). This study confirms Kanerva's (1990) work on Nkhotakota Chichewa showing that phonological re-phrasing is the primary cue to information structure in this language. It expands on Kanerva's work in several ways. First, we show that focus phrasing has intonational correlates, namely, the manipulation of downdrift and pause. Further, we show that there is a correlation between pitch prominence and discourse prominence at the left and right periphery which conditions dislocation to these positions. Finally, we show that focus and syntax are not the only factors which condition phonological phrasing in Chichewa.

1 Introduction

Kanerva's (1990) pioneering study of focus in Chichewa concentrates on the prosodic expression of in situ focus. The effect of in situ focus is illustrated in (1). Under neutral phrasing the entire VP is a single phonological phrase, as shown in (1b)). Narrowing the scope of focus within the Verb Phrase, as in (1c) – (1e), results in an increase in the number of phonological phrases. (The phonological phrases are indicated with parentheses; the phonological motivation for the phrasing is presented in the next section). No prosodic

* We would like to thank the audiences of the Bantu Grammar: Description and Theory IV workshop at the University of Leiden for useful feedback on this paper. Al Mtenje and Laura Downing would also like to thank the British Academy for funding this workshop and their participation in it through an International Network Grant.

(intonational) correlates for these phrase boundaries are reported in previous work.

(1) Focus and phrasing in Nkhotakota Chichewa (Kanerva, 1990: 98, fig. (101))

- (a) anaménya nyumbá ndí mwáála
he hit house with rock ‘He hit the house with a rock.’
- (b) *What did he do?* (VP focus)
(anaményá nyumbá ndí mwáála)
- (c) *What did he hit the house with?* (Oblique PP focus)
(anaményá nyumbá ndí mwáála)
- (d) *What did he hit with the rock?* (Object NP focus)
(anaménya nyuúmba) (ndí mwáála)
- (e) *What did he do to the house with the rock?* (V focus)
(anaméénnya) (nyuúmba) (ndí mwáála)

Ex situ focus is not discussed in Kanerva (1990). This is surprising, as work like Bresnan & Mchombo (1987) shows that word order is rather free in Chichewa. In other Southern Bantu languages, like Makua (Stucky 1985), word order variation correlates with distinct information structures. However, the influence of word order variation on information structure or prosody has not received detailed attention in previous work on Chichewa.

In this paper, we present preliminary results of a study of the interaction of both in situ and ex situ focus (or discourse prominence) with prosody in Al Mtenje’s dialect of Chichewa (Ntcheu Chichewa). The paper is organized as follows. First, we give a background sketch, from previous studies, on phonological and syntactic phrasing in Chichewa. Then we present our findings, which expand on earlier work in the following ways. First, we show that the phonological phrasing induced by focus does, in fact, have intonational correlates. We then show that word order is not as free as reported in earlier work, as there are information structure asymmetries at the left vs. right periphery. Finally, we show that focus and syntax are not the only factors which condition phonological phrasing in Chichewa. The final section presents conclusions and issues for further research.

2 Background on syntactic and phonological phrasing in Chichewa

Kanerva (1990) argues that two main factors condition phonological phrasing in Chichewa: syntax and focus. Sentences in Chichewa have three main constituents – an optional Subject NP, an obligatory VP and an optional Topic NP – which Bresnan & Mchombo (1987), Kanerva (1990) and Mchombo

(2004) suggest can all be freely ordered. Each of these three constituents, when they co-occur, is parsed into its own phonological phrase (Bresnan & Kanerva 1989; Kanerva 1990). Important cues to phonological phrase boundaries are that there are no High tones on phrase-final vowels in the Nkhotakota dialect of Chichewa analyzed in Kanerva (1990), and phrase penult vowels are lengthened.¹ The data in (2) illustrates syntactically-motivated phonological phrasing. As shown in (2b) and (2c), dislocated (object) NPs – Topics – are in a distinct syntactic and phonological phrase, and can occur in either order with respect to the VP:

(2) (a) (Subj) (VP) – Kanerva (1990: 103, fig (114b))

(mwaána)	(anapézá	galú	kúdáambo)
1.child	1-found	dog	at swamp

The child found the dog at the swamp.

(b) (Subj) (VP) (Top) – (Kanerva 1990: 107, fig (123b))

(mwaána)	(a-na-’m – pézá	kúdáambo)	(gaálu)
1.child	1. found it	at swamp	dog

The child found it at the swamp, the dog.

(c) (Top) (VP) (Subj) – (Kanerva 1990: 102, fig (110c))

(a-leenje)	(zi-ná-wá-luuma)	(njúuchi)
2.hunter	10-TA-2-bite	10.bees

The hunters, they bit them, the bees [did].

Focus also plays a role in conditioning phonological phrasing in Chichewa. In situ focus on an element of the VP is possible in Chichewa, and is realized, according to Kanerva (1990), only by a change in the phonological phrasing of the VP. The effect of focus on phonological phrasing is what Hyman (1999) and Truckenbrodt (1995) call ‘boundary narrowing’. A phonological phrase boundary must follow the focused element, and each remaining constituent of the VP is parsed into a separate phonological phrase. As we saw in (1), this results in more, smaller phonological phrases under narrow focus. Another example of focus-induced boundary narrowing is provided in (3):

¹ See Kanerva (1990) for detailed discussion of the range of phonological processes motivating phonological phrases in this dialect of Chichewa.

(3) Nkhotakota Chichewa (Kanerva 1990: 98, fig (102))

- (a) a-na-góná m-nyumbá yá Mávúuto.
 They slept in-house of Mavuto.
 ‘They slept in Mavuto’s house.’
- (b) *Where did they sleep?* (Object NP focus)
 (a-na-góná m-nyumbá yá Mávúuto)
- (c) *What did they do in Mavuto’s house?* (V focus)
 (a-na-góona) (m-nyumbá yá Mávúuto)

Kanerva (1990: chapter 5) shows that these phonological phrases are grouped into Intonational Phrases (IPs): the next highest constituent in the Prosodic Hierarchy (Nespor & Vogel 1986; Selkirk 1986). IPs in Chichewa have the following prosodic correlates: intonational boundary tones, IP-final lengthening, and tonal catathesis (downdrift). These properties are illustrated in (4) and discussed in the lengthy descriptive quote which follows:

(4) IPs in Nkhotakota Chichewa (Kanerva 1990:140)

(a-na-pátsá	mwaána)	(njiínga)↘	(ósatí	mfúumu) ↘
S/he gave	CHILD	bicycle	not	chief
S/he gave the child a bicycle, not the chief.				

“Both IPs in (4) end in low falling contours and show IP-final lengthening; in particular, the IP-penultimate syllable in *njiínga* is noticeably longer than the penultimate syllable in *mwaána*, which is lengthened only at the FP [phonological phrase] level [...]. Tonal catathesis occurs twice in the first IP, not only within the first [phonological phrase], but also between it and the second [phonological phrase]. No catathesis, however, occurs between the IPs; in fact, the High tones of the second IP are all higher pitched than those in *njiínga* and even *mwaána*.”

Notice that there is no discussion here of boundary tones or register adjustments at phonological phrase boundaries, only at IP boundaries.

Like all good pioneering studies, Kanerva (1990) and the other work cited above both establish the groundwork for future research on phonological phrasing in Chichewa and raise new questions. Is there really no intonational correlate of phonological phrasing, especially phrasing conditioned by focus? Work like Selkirk (2004) proposes focus always has a culminative prosodic correlate, and Yip’s (2002) tone survey shows focus has prosodic correlates in other tone languages. These studies motivate taking a second look at Chichewa focus phrase intonation. Is the order of Subject, VP and Topic really ‘free’? That is, is there no difference in context of use or prosody associated with the

different orders? Work like Stucky (1985) and Morimoto (2000) suggests word order variation is strongly influenced by discourse factors in Bantu languages, making it important to take a second look at Chichewa. The remainder of the paper addresses these questions, reporting new research on Ntcheu Chichewa, Al Mtenje's dialect.²

3 Manipulation of downdrift/downstep and information structure

3.1 Downdrift, the suspension of downdrift and focus

One of our most striking findings is that downdrift (catathesis) is manipulated at phonological phrase boundaries, providing an additional salient prosodic cue to focus. To return to the oft-cited examples in (1), we find differences both in phrasing and in pitch register in expressing in situ focus in Ntcheu Chichewa, as shown in (5). The context for each pronunciation is given in square brackets:

(5) Focus and phrasing in Ntcheu Chichewa

(a) anáménya nyumbá ndí mwáálá
s/he hit house with rock
'S/he hit the house with a rock.'

(b) (A-ná-ménya nyumbá ndí mwáálá).
[neutral declarative]

(c) (A-ná-ménya NYUÚMBÁ) ! (ndí mwáálá).
[Answers the question: *A-ná-méenya chiyáani ndi mwáálá?*
'What did he hit with the stone?']

(d) (A-ná-ménya nyuúmbá) (NDÍ MWÁÁLÁ). [no downstep]
[Without downstep, emphasizes 'stone', but 'house' – as it is postverbal and not coreferenced by an OM – is also new information]

(e) (A-NÁ-MÉENYA) ! (nyuúmbá) ! (ndí mwáálá).
[focus on hitting]

We find a number of differences in (5) compared to the description summarized in (1). First, there are dialectal tone differences. For example, as can be seen by

² Other recent analyses of focus and phrasing in Chichewa – eg., Truckenbrodt (1995, 1999), Seidl (2001), Gussenhoven (2004), Samek-Lodovici (in press) – are reanalyses of the data presented in Kanerva (1990). As far as we know, this is the first new study of Chichewa focus prosody.

In the data in the remainder of the paper, '||' indicates pause; '!' indicates downstep. Parentheses continue to set off phonological phrases.

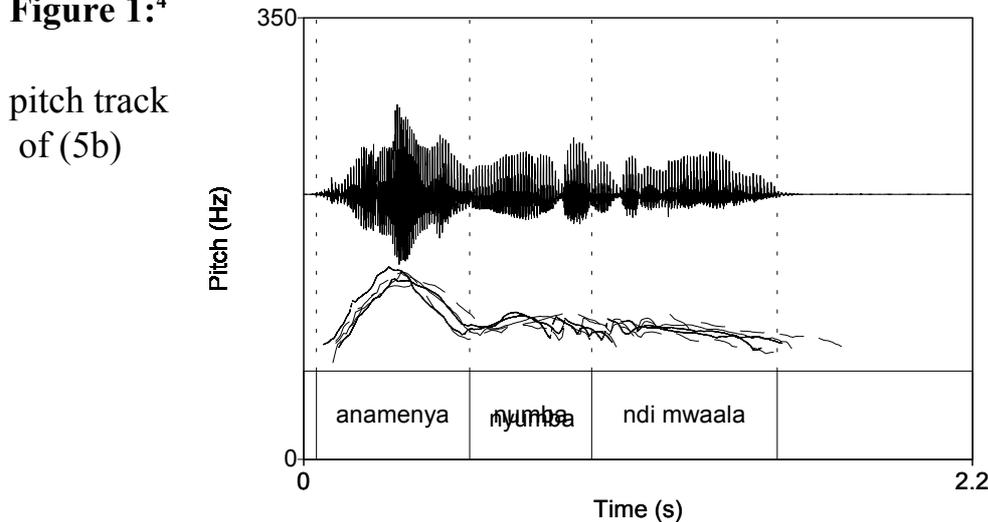
comparing (1c) with (5c), retraction of a High tone from a phrase-final vowel onto the lengthened penult leaves a High tone on the phrase-final syllable in this dialect (see Mtenje (1987), Hyman & Mtenje (1999) for more discussion). We also find a phrasing difference: there is an obligatory phrase break before a VP-final focused element, as can be seen by comparing (5d) with (1c).

The most striking differences, though, lie in the realization of downdrift in (5b) vs. (5c) vs. (5d). Kanerva's (1990) discussion, quoted in (4), above, gives the impression that all of these phrases should have similar downdrift patterns in Nkhotakota Chichewa. Our phonetic study shows that focus interferes with the realization of downdrift, noticeably raising the expected pitch of a High tone in the focused phrase. As shown by the figures in (6) and the pitch tracks in Figure 1, we still find downdrift across the phrase, no matter where focus is found. However, the pattern of downdrift is disrupted by focus:

(6) Pitch of High tones in (5b, c, d):³

	<u>H1</u>	<u>H2</u>	<u>H3</u>	
(5b)	140	115	102	[downdrift in neutral phrasing]
(5c)	155	134	101	[H3 is relatively much lower than H2]
(5d)	134	113	112	[no downdrift if H3 is in focus]

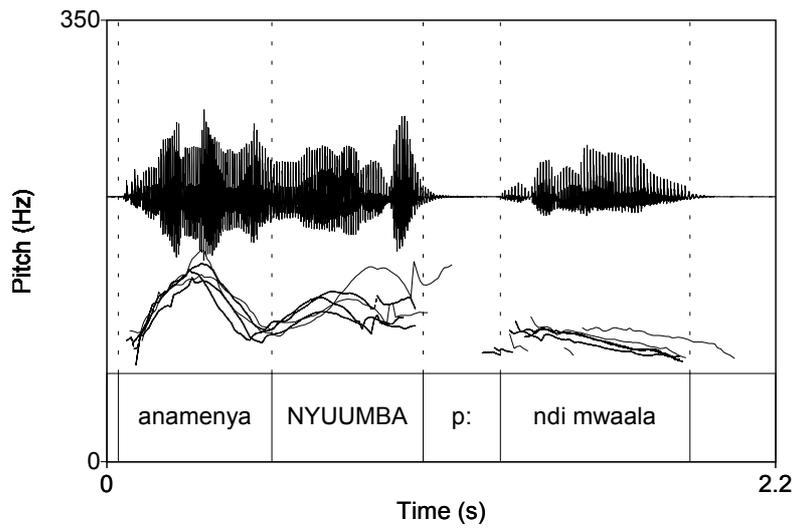
Figure 1:⁴



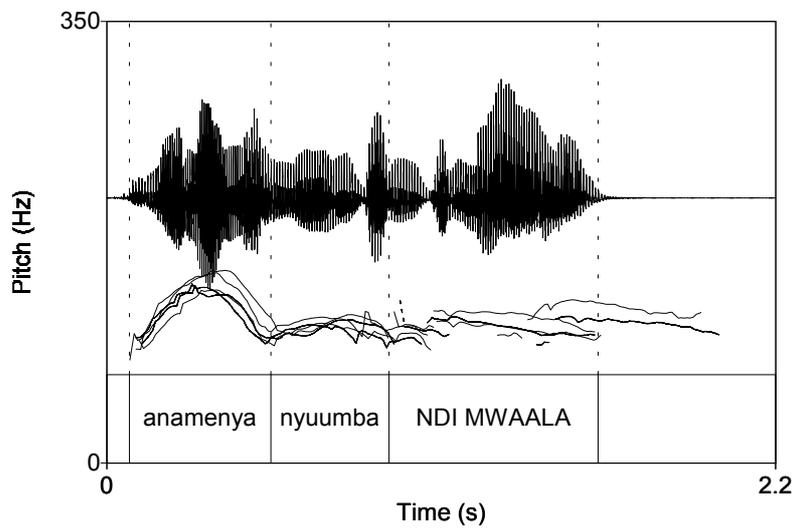
³ Representative f_0 values [inHz] in one repetition out of five.

⁴ The figures – for purpose of demonstration only – always display the audiosignal of one repetition with its segmentation and labelling together with the overlaid pitch tracks of all five repetitions aligned at a constant time point before acoustical signal onset. N.B.: since no time warping was applied there are discrepancies with respect to f_0 peak locations and duration due to differences in speech rate and pausing between the repetitions of the same phrase.

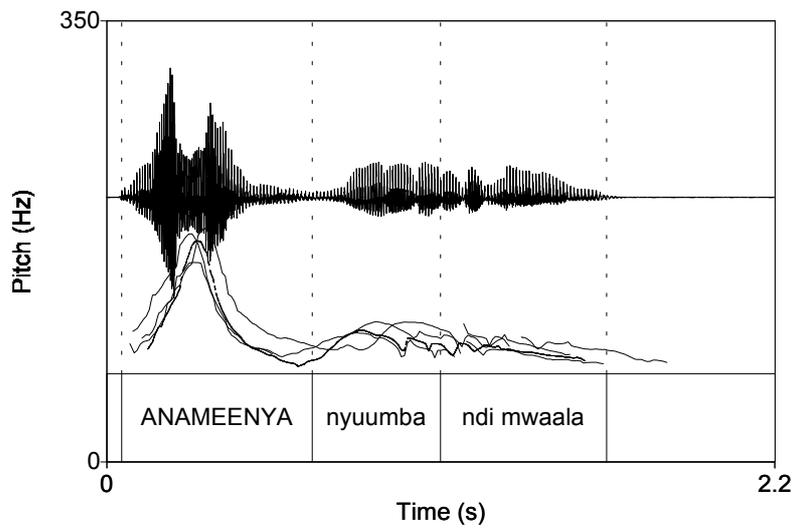
pitch track
of (5c)



pitch track
of (5d)



pitch track
of (5e)

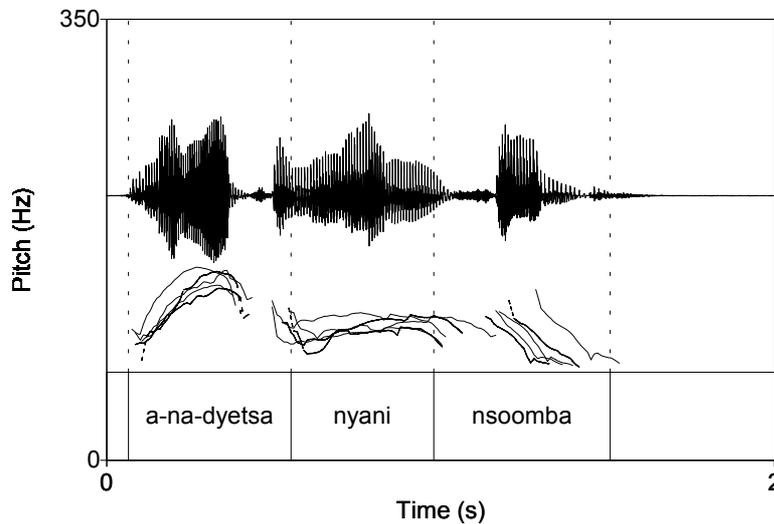


The same patterns of manipulation of downdrift triggered by focus are found consistently, as shown by the pitch tracks in Figures 2 and 3 for the data sets in (7) and (8), respectively:

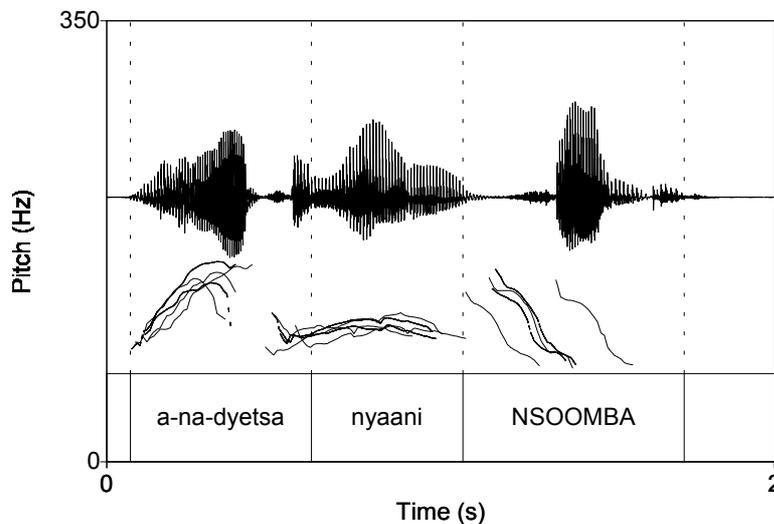
- (7) (a) (A-ná-dyétsa nyaní nsóomba).
 they fed baboon fish
 [neutral declarative]
- (b) (A-ná-dyétsa nyaání) (NSÓOMBA). [no downstep]
 [Focus on ‘fish’]
- (c) (A-ná-dyétsa NYAÁNI) ! (nsóomba).
 [With downstep, focus on ‘baboon’. Answers question:
A-ná-dyétsa ndaání nsóomba? ‘Who did they feed fish to?’]

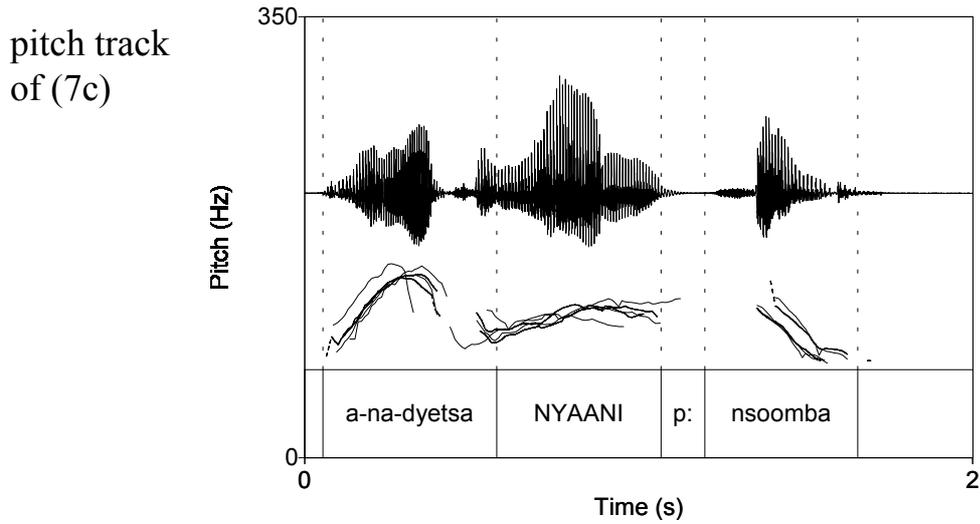
Figure 2:

pitch track
of (7a)



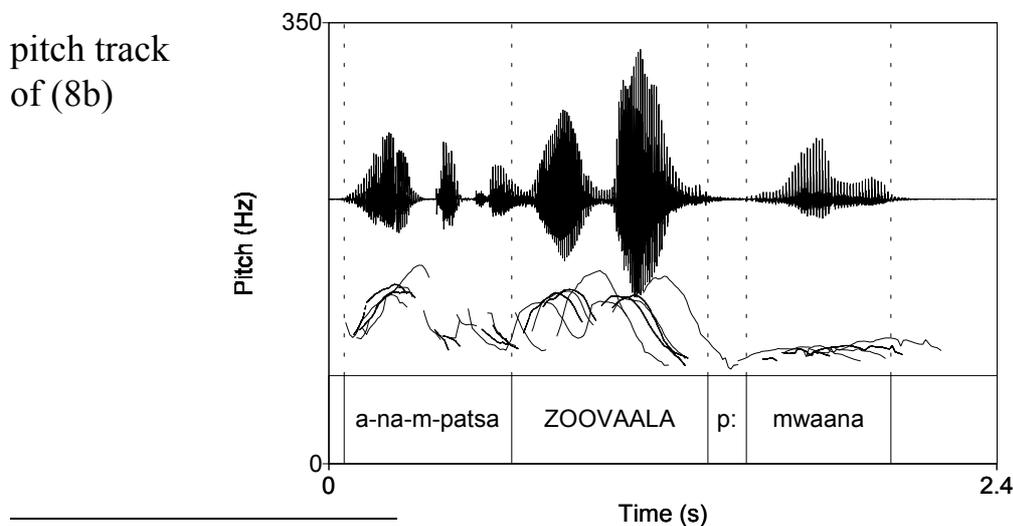
pitch track
of (7b)





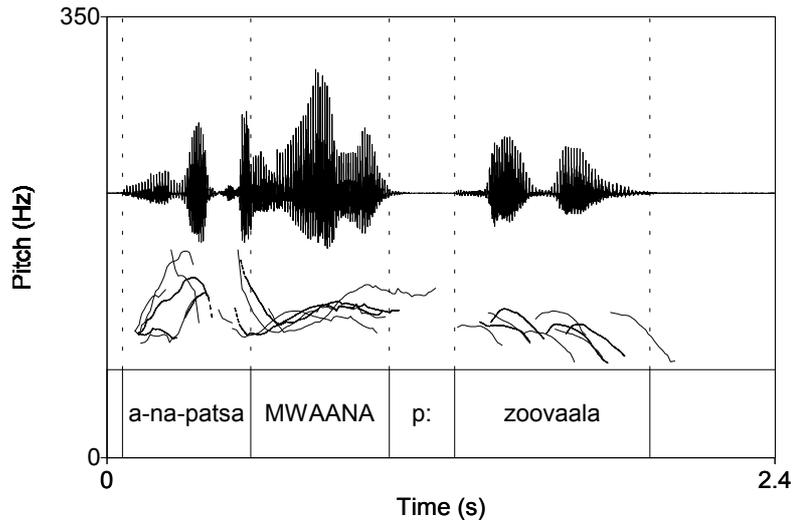
- (8) (a) (Mfúumu) ! (i-ná-pátsa mwaná zóóváala).
 chief gave child clothes
 [neutral declarative]
- (b) (A-ná-'m-patsa ZÓÓVÁALA) ! (mwaáná).
 [Answers question: *A-ná-'m-patsa chiyáani mwaáná?*
 'What did they give to the child?' 'Clothes, and not something else.']
- (c) (A-ná-pátsa MWAÁNÁ) ! (zóóváala).
 [Answers question: *A-ná-pátsa ndááni zóóváala?*
 'Who did they give clothes to?']
- (d) (A-ná-pátsa mwaáná) (ZÓÓVÁALA). [no downstep]
 [If no downstep, then focus on 'clothes'; 'child' is also new information
 as it is positioned right after the verb]

Figure 3: ⁵

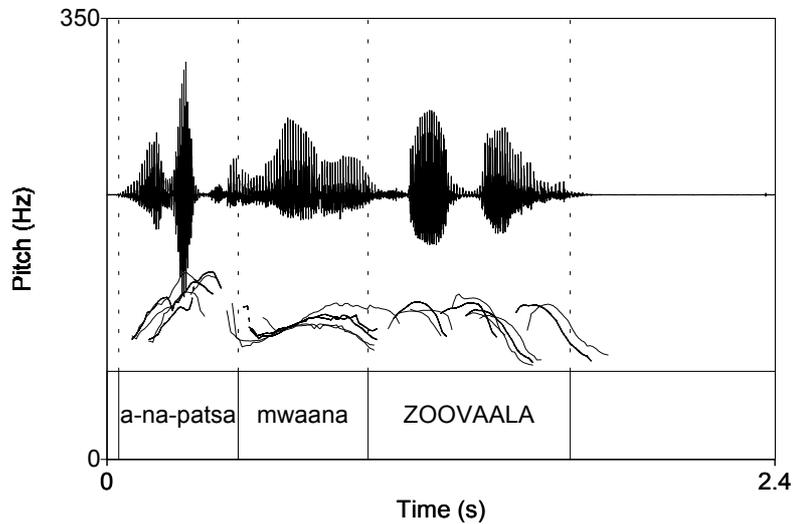


⁵ For discussion of the different timing of the pitch tracks cf. footnote 4.

pitch track
of (8c)



pitch track
of (8d)



The data in (9) present a further set of examples which illustrate the same downdrift patterns. (To save space, pitch tracks for this data set are not provided.)

- (9) (a) (A-ná-pézá galú kudáambo).
 s/he found dog in swamp)
 [neutral declarative]
- (b) (A-ná-pézá GAÁLÚ) ! (kudáambo).
 [Answers question: *A-ná-pézá chiyáani kudáambo?*
 ‘What did they find in the swamp?’]
- (c) (A-ná-pézá gaálú) (KUDÁAMBO). [no downstep]
 [Focus on ‘swamp’ if no downstep; ‘dog’ is also new information]

To sum up this section, focus does have an intonational correlate in Ntcheu Chichewa: the expected degree of downdrift is violated. If the focused element is not the last word of the phrase, its High tones are raised enough to create a steep fall to the next High tone, giving what Downing (2004a,b) calls an ‘anti-accent’ effect. That is, the focused element is not made prominent by giving it the culminative pitch height of the utterance – this is what is typically meant by ‘accent’. Rather, it is made prominent by raising the pitch enough to make the following elements relatively much lower in pitch – this is what is meant by ‘anti-accent’. If the focused word is final in the phrase, the expected lowering of its pitch due to downdrift is suspended. This violation of expected downdrift patterns in tonal sequences triggered by focus makes manipulation of pitch register at phonological phrase boundaries an additional prosodic cue to focus.

3.2 *Downdrift and positions of contrast*

A second striking finding is that pitch height asymmetries correlate with information structure asymmetries for dislocated (Topic) NPs at the left- vs. right periphery. As shown in (2), Subject and Topic NPs can occur either before or after the VP in Chichewa. The previous work on Nkhotakota Chichewa, cited above, suggests that this ordering is free. Our work shows that the relative order of sentential constituents has discourse consequences. Sentence-initial Topic/Subject is understood as contrastive and discourse prominent. Sentence-final Topic/Subject cannot be contrastive; it is downplayed, discourse-old information. When two NPs precede (or follow) the VP, the order of discourse salience matches their ‘left to right’ order of occurrence. This is illustrated by the data sets in (10) and (11):

- (10) (a) (Mikáango) || ! (mbúzi iizi) ! (i-ná-zí-saaka).
lions goats these they hunted them.
The lions hunted these goats.
[focus on ‘lions’ as topic; more prominence on ‘these goats’ than if it follows verb]
- (b) (I-NÁ-ZÍ-SAAKA) || ! (mikáango) || ! (mbúzi iizi).
[Emphasizes the hunting]
- (c) (Mbúzi iizi) || ! (mikáango) || ! (i-ná-zí-saaka).
[Emphasis on ‘goats’ as contrastive topic; ‘lions’ highlighted compared to ‘hunting’: each downstepped item less discourse prominent]

- (d) (Mbúzi iizi) || ! (i-ná-zí-saaka) || ! (mikáango).
[Emphasis on ‘goats’ as contrastive topic; ‘hunting’ highlighted compared to ‘lions’]
- (11) (a) (ALEÉNDÓ) || ! (a-ná-’m-dyetsa nsóomba).
visitors fed him fish
[Answers question: *Ndaání á-ná-m-dyétsá nsóomba?* ‘Who fed him fish?’]
- (b) (A-ná-’m-dyetsa nsóomba) ! (aleéndó).
[Answers question: *A-ná-’m-taání nyaání alééndó?* ‘What did the visitors do to the baboon?’]
- (c) (A-ná-’m-dyetsa nsóomba) (aleéndó). [no downstep]
[Emphasis is on ‘visitors’ as choice in a known list of possibilities: *A-ná-m-dyétsá nsóomba ndaání?* ‘WHO fed him fish?’]

The contrast in the discourse function and pitch for the word for ‘visitors’ in (11b) and (11c) is especially striking. In (11b), when the word is right-dislocated and downstepped, it is clearly old information, an afterthought. When downdrift is suspended for this word, as in (11c), then it can be a contrastive reiteration of old information.

To sum up this section, word order variation is not free in Ntcheu Chichewa but rather reflects discourse salience and contrast. It is not an accident that positions of relative discourse salience are realized sentence-initially. Prosodically, recall that downdrift across an utterance has as a result that sentence-initial elements are pronounced with the highest pitch and sentence-final with the lowest pitch. What we find, then, is that relative pitch prominence correlates with relative discourse prominence. Due to downdrift, sentence-initial elements are prosodically more salient, realized at a relatively higher pitch.

4 The role of pause as a cue to structure

Another common cue to phonological phrase boundaries, in our study, is pause. As we show in this section, the placement of pauses is, in fact, often an important cue to the information structure or grammatical structure of the sentence.

4.1 Information structure and pause

Often pauses occur in positions one might expect, for example, to set off dislocated Topic NPs. The data in (12) – (14) illustrates the correlation between pause and left-dislocation of NPs:

- (12) (a) (A-ná-pézá galú kudáambo).
s/he found dog at swamp)
[neutral declarative]
- (b) (Kudáambo) || ! (a-ná-pézá gaálú).
[Answers question: *Ndí kúuti á-ná-pezá gáálú?*
It is where that s/he found the dog?]
- (13) (a) (A-ná-dyétsedwa nsómbá ndí mávúuto).
They were fed fish by Mavuto
[Answers question: *A-ná-táání anthu awa?*
‘What happened to these people?’]
- (b) (Nsóomba) || ! (a-ná-dyétsedwa ndí mávúuto).
[Big pause required, plus right context: Two things were given as food, meat and fish. As for the fish, it was given by Mavuto. Meat given by someone else.]
- (14) (a) (Mikáango) ! (i-ná-sáka mbúzi iizi).
lions hunted goats these
[neutral phrasing]
- (b) (I-NÁ-ZÍ-SAAKA) || ! (mikáango) || ! (mbúzi iizi).
[Emphasizes the hunting]

Focus can also be accompanied by pause. As we saw in the previous section, when a phonological phrase boundary follows a focused NP, a pause (as well as a downstep) generally marks the phrase boundary. This is clearly shown in the pitch tracks for (5c) and also (7c) and (8c). In those examples, pause is a secondary cue to focus.

Pause is sometimes the primary cue to focus. For example, the important difference between an in situ focused subject and in situ topic subject is the greater degree of pause following a focused subject.⁶

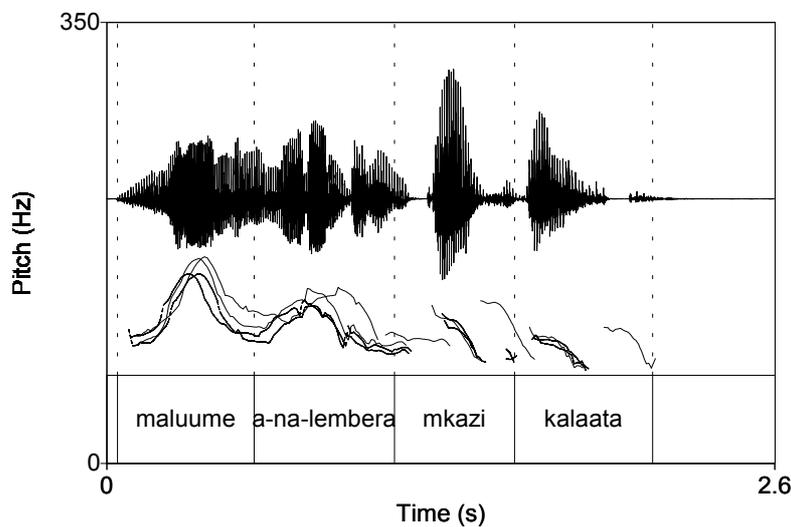
⁶ Kanerva's (1990: 103) study of Nkhotakota Chichewa shows no difference in pronunciation or phrasing when the subject is a topic vs. focused. Perhaps this is a dialect difference.

- (15) (a) (Mikáango) ! (i-ná-sáka mbúzi iizi).
 lions hunted goats these
 Lions hunted these goats.
 [neutral phrasing]
- (b) (MIKÁANGO) ! || (i-ná-sáka mbúzi iizi).
 [focus on ‘lions’: contrasts them with other possible agent and
 answers the question: *Ndaáni á-ná-saká mbúzi iizi?*
 ‘Who hunted these goats?’]
- (16) (a) (Malúume) ! (a-ná-lémbera mkázi kálaata).
 uncle s/he wrote woman letter
 [simple declarative]
- (b) (MALÚUME) || ! (a-ná-lémbera mkázi kálaata).
 [Focus on ‘uncle’: *Ndaáni á-ná-lembéla mkázi kálaata?*
 ‘Who wrote the letter to the woman?’]

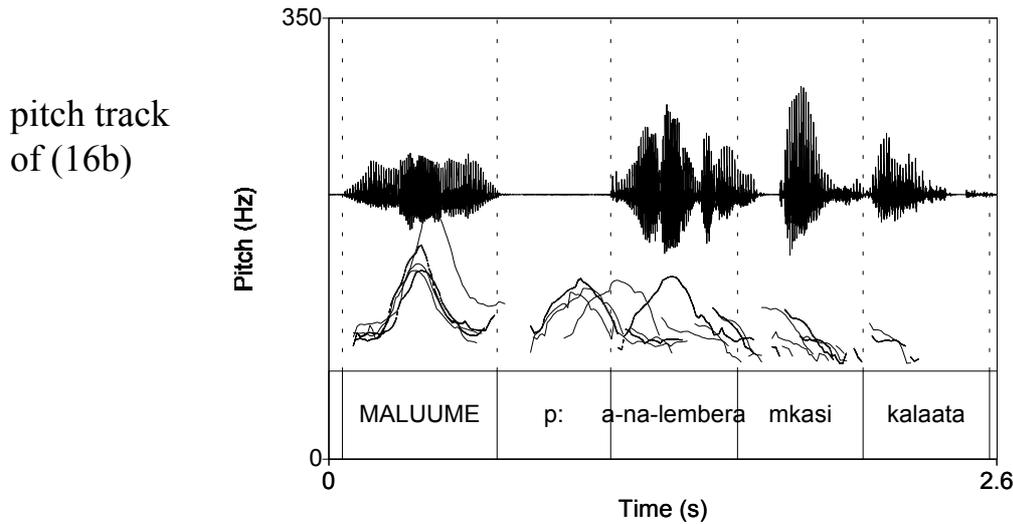
The pitch tracks in Figure 4 clearly show how pause correlates with a focused, new information subject (16b),⁷ while lack of pause correlates with the neutral, topic role of subject (16a).

Figure 4:

pitch track
of (16a)



⁷ For discussion of the different timing of the pitch tracks cf. footnote 4.



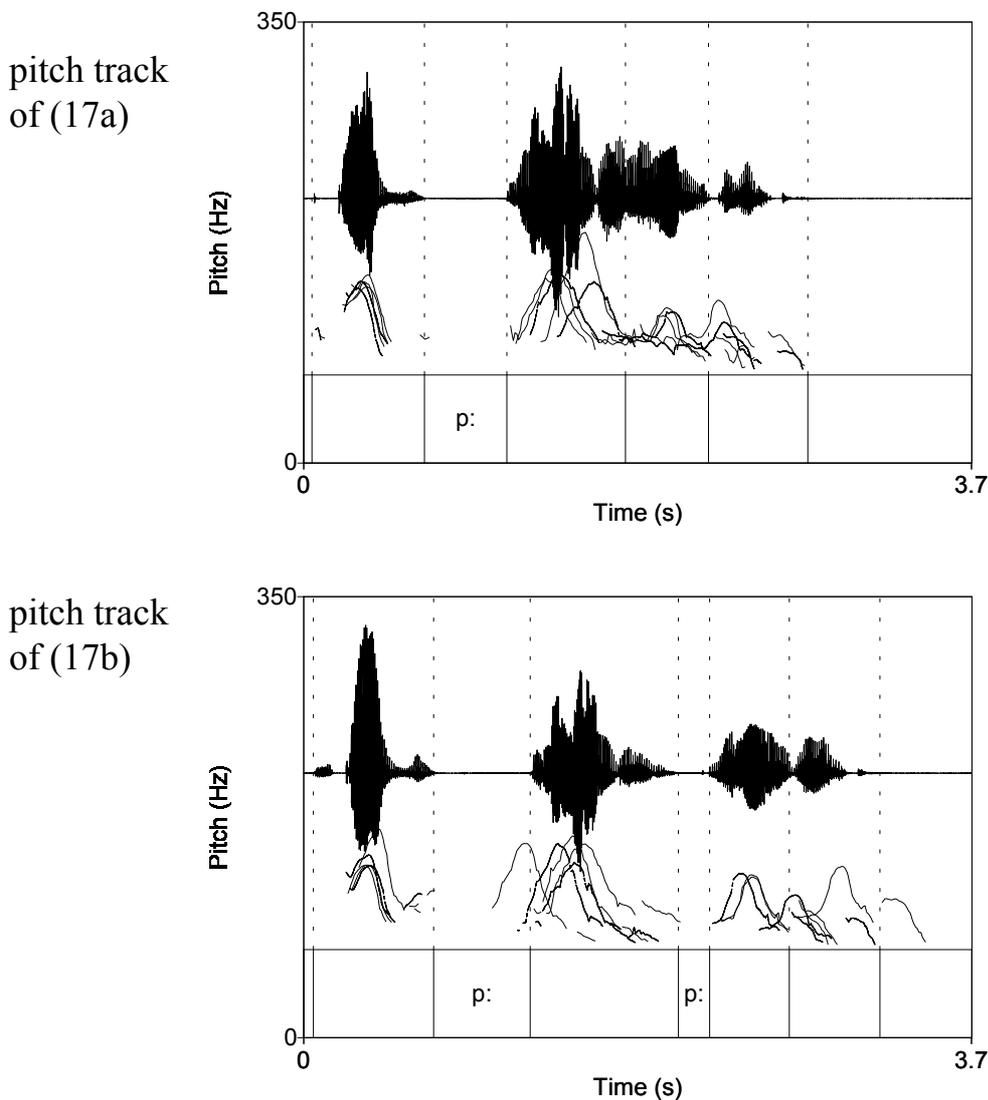
4.2 Grammatical role and pause

In some cases, differences in phonological phrasing and pause placement can change the grammatical role attributed to NPs. In the few examples we have, this use of pause seems to be most likely when the verbal morphology does not disambiguate the grammatical role. For example, in (17), ‘woman’ and ‘uncle’ belong to the same noun class. As a result, it is ambiguous from the verb agreement morphology which NP is the subject and which is the object. The contrastive placement of pauses in (17a) compared to (17b) disambiguates the grammatical role of these two NPs:

- (17) (a) (mkáazi) || (a-ná-mú-lembera malúume) ! (kálaata).
 woman s/he wrote him uncle letter
 [‘woman’ is understood as the agent and ‘uncle as the recipient]
- (b) (mkáazi) || ! (a-ná-mú-lembeera) || ! (malúume) ! (kálaata).
 [‘woman’ is understood as the patient and ‘uncle’ is the agent]

The pitch tracks in Figure 5 show the contrastive pause placements in (17a) vs. (17b) clearly.

Figure 5:



Phonological phrasing can play a similar role in disambiguating the grammatical role of NPs. The sentence in (18), like the one in (17), contains ambiguous verb agreement morphology. In this case, the object prefix (-m-) could refer either to ‘child’ or to a third human NP not mentioned in the sentence. If ‘child’ is phrased with verb, as in (18a), it is unambiguously understood as the recipient. If it is phrased separately, as in (18b), it is ambiguously either agent or recipient.

- (18) (a) (mwaná ! a-ná-’m-patsa zóóváala)
 child s/he –him-gave clothes
- (b) (mwaáná) ! (a-ná-’m-patsa zóóváala)

4.3 *Mixed signals*

Surprisingly, in some ‘neutral phrasings’, no phonological phrase boundary or pause follows a sentence-initial NP, whether subject or object, but noticeable downstep still follows as if there were a phrase boundary. One example has already been seen in (18a), others are given below:⁸

(19) (a) (Aleéndó) ! (a-ná-’m-dyetsa nsóomba).
Visitors fed him fish
[neutral declarative]

(b) (Alendó ! a-ná-’m-dyetsa nsóomba).
[neutral declarative]

(20) (a) (Mpháatso) || ! (a-ná-péreka kwá mfúumu).
gifts they gave to chief

[As for gifts, they were given to the chief, and other items had something else happen; must pause to make it clear that ‘gift’ is the object if no OM is used.]

(b) (Mpháatso) ! (a-ná-yí-pereka kwá mfúumu).
[‘gift’ is simple topic, as if subject of a passive sentence]

(c) (Mphátsó ! a-ná-yí-pereka kwá mfúumu).
[neutral declarative as in (20b)]

What is interesting in this data is that larger phonological phrasings continue to correlate with neutral information structure, while boundary narrowing adds some discourse prominence to the word set off by phrasing. It is also interesting to note that the phonological phrasings in (18a), (19b) and (20c) are not predicted by previous work on Chichewa, as the pre-verbal NPs should, in all these cases, be in separate phonological phrases to match the syntactic phrasing. This data emphasizes that focus and syntax are not the only factors conditioning phonological phrasing in Ntcheu Chichewa.

5 **Conclusion and issues for further research**

To sum up, our preliminary phonetic study of Ntcheu Chichewa provides a new perspective on both the prosody of focus and the function of word order

⁸ Some participants in the Leiden Workshop were not sure the downstep is as noticeable in examples like (19b) as the authors perceived. A phonetic study of data sets like these is necessary to decide this question.

variation. One of our most significant findings is that downdrift is manipulated at phonological phrase boundaries as an additional cue to focus. Most work on focus has investigated the role of sentence accent as the primary cue to focus (Gussenhoven 1984, 1996, 1999, 2004; Selkirk 1984, 1995, 2004). Indeed, Truckenbrodt's (1995) reanalysis of Kanerva (1990) suggests that focused elements – as the heads of their utterances, by definition – must have sentential prominence. As Samek-Lodovici (in press) shows, theoretical problems are caused by Chichewa focus phrasing if focused elements do not, in fact, have sentential prominence. While the manipulations of downdrift resemble culminative focus accent in raising the pitch of focused constituents, we would argue that it is not equivalent to accent. No culminative prominence is given to the focused element, rather a relative prominence: the expected degree of downdrift is violated. This study, then, confirms work like Ladd (1996: 195-196) which argues that accent is not a universal cue to focus. It also confirms work like Hayes & Lahiri (1991), Ladd (1996: ch. 5) and Downing (2003) arguing that focus involves paradigmatic prominence, rather than culminative prominence. Words or phrases are recognized as being in narrow focus because they do not have the unmarked prosody that correlates with the 'neutral' pronunciation of the same phrase.

A further important role for downdrift is that it explains why word order is not actually free in Chichewa. We find an asymmetry in discourse prominence of sentence-initial topic – which is potentially contrastive – vs. downstepped sentence-final topic – which is in discourse 'oblivion.' There is a matching asymmetry in relative pitch height at the left vs. right periphery: sentence-initial high tones are highest in pitch. Due to downdrift, sentence-final high tones are lowest in pitch. What we find, then, is a correlation between pitch and discourse prominence asymmetries at the peripheries of the sentence.⁹

Our study also identifies additional roles for phonological phrasing and pause than those reported in Kanerva (1990). Syntactic structure and focus are not the only factors which condition phrasing. As shown in the preceding section, a sentence-initial NP is not always a separate phonological phrase, even though it is presumably always a separate syntactic constituent. Our study also shows that phrasing is a cue to grammatical role as well as focus and syntactic structure. Finally, we have shown that pause is a cue to subject focus, and that distinct pauses are required as a parsing aid with certain word orders when the verbal morphology does not unambiguously define the grammatical role. Pause

⁹ See Downing (2004a) for discussion of a similar correlation between pitch prominence and discourse prominence in Chitumbuka. See Hock (1998) for discussion of final reduction as a factor motivating pre-final focus position in non-Bantu languages.

does not just occur at IP boundaries, as previous work on the distribution of pauses, like Nespor & Vogel (1986) suggests.

We would like to close by identifying areas for future research. Further phonetic studies are clearly needed to test if similar prosody is found in other languages. For example, in Xhosa (Jokweni 1995) and Haya (Byarushengo et al. 1976) focus also conditions boundary narrowing. Do they also have similar manipulation of downdrift at these phonological phrase boundaries? Herero (Möhlig et al 2002) is reported to have register raising under focus. Is raising in Herero like Chichewa, where raising means suspension of downdrift? Or is it like stress-accent, where the focused element has the highest pitch of the utterance? We hope this study lays the foundation for pursuing these questions.

References

- Bresnan, J. and S. Mchombo (1987) Topic, pronoun and agreement in Chichewa. *Language* 63: 741-782.
- Bresnan, J. and J. Kanerva (1989) Locative inversion in Chichewa: a case study of factorization in grammar. *Linguistic Inquiry* 20: 1-50.
- Byarushengo, E. R., L. M. Hyman and S. Tenenbaum (1976) Tone, accent, and assertion in Haya. In Larry M. Hyman, ed. *Studies in Bantu Tonology (SCOPIL 3)*: 185-205.
- Downing, L. J. (2003) Stress, tone and focus in Chichewa and Xhosa. In Rose-Juliet Anyanwu (ed.). *Stress and Tone – the African Experience. Frankfurter Afrikanistische Blätter* 15: 59-81.
- Downing, L. J. (2004a) Focus and phrasing in Chitumbuka. Paper presented at CALL 34, University of Leiden, 23-25 August 2004.
- Downing, L. J. (2004b) The prosody of focus in Bantu languages and the primacy of phrasing. Paper presented at the TIE Conference, Santorini (Greece), 9-11 September 2004.
- Gussenhoven, C. (1984) *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris.
- Gussenhoven, C. (1996) Sentence accents and argument structure. In I.M. Roca, ed. *Thematic Structure: Its Role in Grammar*. Berlin: Foris, 79-106.
- Gussenhoven, C. (1999) On the limits of Focus Projection in English. In Peter Bosch & Rob van der Sandt, eds. *Focus: Linguistic, Cognitive, and Computational Perspectives*. Cambridge: Cambridge University Press, 43-55.
- Gussenhoven, C. (2004) *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.
- Hayes, B. and A. Lahiri (1991) Bengali Intonational Phonology. *NLLT* 9, 47-96.
- Hock, H. H. (1998) Finality, prosody and change. In *Proceedings of LP'98*, O. Fujimura, B.D. Joseph and B. Palek (eds.). Prague: The Karolinum Press: 15-30.
- Hyman, L. M. (1999) The Interaction between Focus and Tone in Bantu. In: Georges Rebuschi and Laurie Tuller (eds.). *The Grammar of Focus*. Amsterdam: John Benjamins: 151-177.

- Hyman, L. M. and A. Mtenje (1999) Prosodic Morphology and tone: the case of Chichewa. In René Kager, Harry van der Hulst and Wim Zonneveld, (eds.). *The Prosody-Morphology Interface*. Cambridge: Cambridge University Press: 90-133.
- Jokwani, M. W. (1995) *Aspects of IsiXhosa Phrasal Phonology*. Ph.D. dissertation, University of Illinois at Urbana-Champaign.
- Kanerva, J. (1990) *Focus and Phrasing in Chichewa Phonology*. New York: Garland.
- Ladd, D. R. (1996) *Intonational Phonology*. Cambridge: Cambridge University Press.
- Mchombo, S. (2004) *The Syntax of Chichewa*. Cambridge: Cambridge University Press.
- Möhlig, W. J.G., L. Marten and J. U. Kavari (2002) *A Grammatical Sketch of Herero (Otjiherero)*. Köln: Rüdiger Köppe Verlag.
- Morimoto, Y. (2000) *Discourse Configurability in Bantu Morphosyntax*. Ph.D. dissertation, Stanford.
- Mtenje, A. (1987) Tone shift principles in the Chichewa verb: a case for a tone lexicon. *Lingua* 72, 169-209.
- Nespor, M. and I. Vogel (1986) *Prosodic Phonology*. Dordrecht: Foris Publications.
- Samek-Lodovici, V. (in press) Prosody-syntax interaction in the expression of focus. *NLLT* 23.
- Seidl, A. (2001) *Minimal Indirect Reference: A Theory of the Syntax-Phonology Interface*. London: Routledge.
- Selkirk, E. O. (1984) *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Mass.: MIT Press.
- Selkirk, E. O. (1986) On derived domains in sentence phonology. *Phonology Yearbook* 3: 371-405.
- Selkirk, E. O. (1995) Sentence prosody: intonation, stress and phrasing. In John A. Goldsmith, (ed.). *The Handbook of Phonological Theory*. Cambridge, Mass.: Blackwell, 550-569.
- Selkirk, E. O. (2004) Bengali intonation revisited. In Chungmin Lee, Matthew Gordon & Daniel Büring (eds.), *Topic and Focus: A Cross-Linguistic Perspective*. Dordrecht: Kluwer, 217-246.
- Stucky, Susan U. 1985. *Order in Makua Syntax*. New York: Garland.
- Truckenbrodt, H. (1995) *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*. Ph.D. dissertation, MIT.
- Yip, M. (2002) *Tone*. Cambridge: Cambridge University Press.

The phonetic motivation for phonological stop assibilation*

T. A. Hall

Indiana University

Silke Hamann

Zentrum für Allgemeine Sprachwissenschaft, Berlin

Marzena Żygis

Zentrum für Allgemeine Sprachwissenschaft, Berlin

This article examines the motivation for phonological stop assibilations, e.g. /t/ is realized as [ts], [s] or [tʃ] before /i/, from the phonetic perspective. Hall & Hamann (2003) posit the following two implications: (a) Assibilation cannot be triggered by /i/ unless it is also triggered by /j/, and (b) Voiced stops cannot undergo assibilations unless voiceless ones do. In the following study we present the results of three acoustic experiments with native speakers of German and Polish which support implications (a) and (b). In our experiments we measured the friction phase after the /t d/ release before the onset of the following high front vocoid for four speakers of German and Polish. We found that the friction phase for /tj/ was significantly longer than that of /ti/, and that the friction phase of /t/ in the assibilation context is significantly longer than that of /d/.

1 Introduction

This article examines stop assibilations – defined here as processes which convert a (coronal) stop to a sibilant affricate or fricative before high vocoids, e.g. /t/ is realized as [ts], [s] or [tʃ] before /i/. In a typological study of phonological assibilations in more than 30 typologically diverse languages, Hall & Hamann (2003) postulate the following two implications:

* We would like to thank Susanne Fuchs, Christine Mooshammer and Bernd Pompino-Marschall for comments on an earlier version of this article. We are also grateful to the input provided by the audience of the 26th DGfS Conference in Mainz and the 9th Conference on Laboratory Phonology in Urbana-Champaign. All disclaimers apply.

- (1) Two implications:
 - a. Assibilation cannot be triggered by /i/ unless it is also triggered by /j/.
 - b. Voiced stops cannot undergo assibilation unless voiceless ones do.

In the present study we present the results of three acoustic experiments with native speakers of German and Polish which support the implications in (1). It will also be shown that additional results of these three experiments point to possible universal generalizations on phonological assibilation not discussed in the literature on this process (e.g. Clements 1999, Kim 2001, Hall & Hamann 2003). The present treatment is therefore important because we provide additional evidence that phonological assibilation can only be adequately explained by appealing to phonetics (see the authors cited above).

The purpose of this paper is to consider implications like the ones in (1) from the phonetic perspective. In particular, since stop assibilation is a process which is phonetically motivated by the turbulent noise which occurs after the release of a stop into a following high vocoid (referred to below as the *friction phase*), we postulate that implication (1a) can be supported if it can be shown that the friction phase in /tj dj/ is significantly longer than in /ti di/. The implication in (1b) would similarly derive support if the friction phase in sequences like /ti/ is longer than in /di/. We would furthermore expect that these predictions could be borne out in any one of a number of languages, since Kim's (2001) universal claims are based on a phonetic study of a single language (Korean) and Hall & Hamann's (2003) observations hold for a number of typologically diverse languages. In the following paragraphs we present the results of an acoustic study of /ti tj di dj/ sequences in German and Polish.

This article is organized as follows. In §2 we define in greater detail the kind of processes we understand to be assibilation and summarize briefly the typological findings in Hall & Hamann (2003) which led them to posit the two implications in (1). §3 summarizes two predictions pertaining to the phonetic realization of sequences like /ti/, /di/, /tj/ and /dj/ which – if confirmed – would support the two implications in (1). In §4 we present the results of three acoustic studies in which the friction phase from the release of /t d/ onto a following high vocoid for several German and Polish speakers is measured. The results of these experiments are significant because they lend support to the two predictions established in §3. In §5 we discuss several unexpected factors influencing the friction phase in sequences like /ti/, namely the relevance of stress and quality of the adjacent vowels. In §6 we discuss the phonetic motivation of the results of the three experiments, namely the reason why /j/ triggers a longer friction phase in a preceding stop than /i/ (section 6.1), the motivation for a longer friction phase in /t/ than in /d/ (section 6.2), the influence of stress (section 6.3) and the influence of the quality of adjacent vowels (section 6.4). §7 considers an

alternative phonetic motivation for implications in (1), namely the similarity in center of gravity of burst friction before /i/ and /j/ for both alveolar stops. §8 concludes.

2 Stop assibilations

In this section we define what we mean by stop assibilation and then present several universal properties for such processes discussed by Hall & Hamann 2003, as well as by earlier authors, namely Foley (1973, 1977), Jäger (1978), Bhat (1978), Ohala (1983), Clements (1999) and Kim (2001).

Stop assibilations (or assibilations for short) are defined here as processes whereby stops become sibilant affricates or sibilant fricatives before high vocoids. Three examples of such rules have been presented in (2).

(2) Three examples of phonological assibilation rules:

- | | | |
|---------------------------------------|------------------------------------|------------------|
| a. $t \rightarrow s / _ i$ | Finnish (Kiparsky 1973) | spirantization |
| b. $t t^h \rightarrow ts ts^h / _ i$ | Korean (Kim 2001) | affrication |
| c. $t \rightarrow tʃ / _ i$ | West Futuna-Aniwa (Dougherty 1983) | posteriorization |

The distinction between the three outputs in (2), i.e. ‘spirantizations’, ‘affrications’ and ‘posteriorizations’, is not important for the present study. The three processes are referred to collectively as assibilations because they all display a similar cluster of properties (see 3 below).

Although processes like the ones in (2) can also affect a velar stop (e.g. in Late Latin /k g/ surfaced as [ts dz] before /j/; Pope 1952) and in some rare languages a labial (e.g. in Lahu labial stops and nasals are affricated before /u/; Mattisoff 1982: 3), we restrict our discussion to assibilations which have a coronal stop as the input segment, in particular the input is dental or alveolar, i.e. [+coronal, +anterior] in terms of distinctive features.

Assibilations like the ones in (2) can either be lexical or postlexical rules. For example, in Korean (see 2b) assibilation is lexical because it is restricted to applying within a derived environment and does not affect tautomorphic /ti/, /t^hi/ sequences. In Quebec French (Cedegren, Archambault & Boulianne 1991, Kim 2001) the assibilation rule is postlexical because it applies across the board, both within and across words. Since the properties we discuss below hold for postlexical and lexical assibilations we do not see the need to distinguish between the two rule domains.

The term ‘assibilation’ is used here in a very narrow sense since we restrict our discussion below to processes like the ones in (2), which share the following three properties (based on the findings of Clements 1999 and Kim 2001):

- (3) Three properties of stop assibilation:
- a. the trigger is some subset of the high front vocoids (i.e. /i j/)
 - b. the output is a sibilant (either an affricate or a fricative)
 - c. the trigger is to the right of the target

Earlier literature offers a phonetic explanation for the properties of stop assibilation in (3a-c). Jäger (1978: 316) states that “if a language has rules that either devoice, aspirate, fricate or affricate consonants before only some vowels but not others, it will be before high rather than low vowels.” Jäger attributes this implication to the fact that the narrow constriction for high vowels creates better conditions for turbulence, which can affect the preceding consonant by adding frication to it. This frication together with the preceding stop can be reinterpreted as a fricated stop, i.e. an affricate. Ohala (1983: 204) also observes that stops tend to be realized with a fricated release when they are followed by close vowels. He ascribes this tendency to the fact that the high velocity of the airflow created upon release is maintained longer when a stop is followed by a close vowel as opposed to an open vowel. Kim (2001) and Clements (1999) state that the creation of sibilants from stops has its phonetic origin in the brief period of turbulence (or ‘friction phase’) which occurs at the release of a stop into a following high vocoid. Thus, Clements (1999) and Kim (2001) observe that the friction phase that occurs in some languages following the release of an alveolar stop into a high front vocoid is significantly longer than the friction phase of the same stop which is released into a non-high and/or non-front vocoid.

In Hall & Hamann’s (2003) study not all logical language types are shown to be attested. In (4) the ten logical language types are presented with the variables /i/ and /j/ as triggers and /t/ and /d/ as assibilating segments. Of these ten types only the five in (4a) were shown to be occurring.

- (4) Ten logical language types:

- a. *Occurring assibilation types:*

<i>Language Type</i>	<i>assibilating segment(s)</i>	<i>trigger(s)</i>
A	/t d/	/i j/
B	/t d/	/j/
C	/t/	/i j/
D	/t/	/j/
E	none	/i j/, /i/, /j/, none

- b. *Nonoccurring assibilation types:*

<i>Language Type</i>	<i>assibilating segment(s)</i>	<i>trigger(s)</i>
F	/t d/	/i/
G	/t/	/i/

H	/d/	/i j/
I	/d/	/j/
J	/d/	/i/

Hall & Hamann (2003) propose that the nonoccurring language types in (4b) fall out from the two implications in (1) which we have repeated in (5) for convenience:

- (5) Two implications:
- Assibilation cannot be triggered by /i/ unless it is also triggered by /j/.
 - Voiced stops cannot undergo assibilations unless voiceless ones do.

Given the two triggers /i/ and /j/ and the two input segments /t/ and /d/, the implications in (5a-b) mean that none of the languages in (4b) is expected to occur.

3 Predictions

Since assibilations have their phonetic origin in the friction phase which arises after the release of a coronal stop before a high vocoid, we argue that the implications in (5) and the typology in (4) can be shown to be grounded in phonetics if the four predictions in (6) are borne out:

- (6) Four predictions:
- The friction phase in /tj/ is of longer duration than in /ti/
 - The friction phase in /dj/ is of longer duration than in /di/
 - The friction phase in /tj/ is of longer duration than in /dj/
 - The friction phase in /ti/ is of longer duration than in /di/

In essence, a longer friction duration after a coronal stop is more likely to be interpreted as affricate than a shorter one. If prediction (6a-b) can be shown to be correct then this finding would lend support to implication (5a). Along the same lines, if prediction (6c-d) can be substantiated experimentally then this would support implication (5b).¹

The typology in (4) suggests an assibilation hierarchy as in (7), in which the wedge ‘<’ means ‘implies’; hence, the assibilation of /ti/ implies the assibilation of /tj/, etc. Note that this typology makes no predictions concerning

¹ Predictions (6c-d) can probably be attributed to the fact that the friction phase of /t/ is generally speaking longer than the friction phase in /d/ (regardless of the quality of the following vocalic element). See section 6.2 for discussion.

the relationship of /ti/ to /dj/; hence, both sequences occupy the same slot in the hierarchy in (7).

(7) An assibilation hierarchy (to be revised):

$$/tj/ < \{ /ti/, /dj/ \} < /di/$$

Should the predictions in (6a-b) be confirmed then this would lend phonetic support to the hierarchy in (7).

4 A phonetic analysis

In this section we present phonetic evidence supporting the predictions in (6) and the hierarchy in (7). We present below the results of three acoustic studies of /ti tj di dj/ sequences, one with German speakers only (section 4.1) and the second and third with German and Polish speakers, respectively (section 4.2). In our studies we measured the duration from the release of /t d/ to the onset of the following vocoid (the start of the fundamental frequency and continuous formants) as illustrated in Figure 1. This phase includes phonetic instances such as burst friction (BF) and aspiration (A) and will be referred to below as the ‘friction phase’.

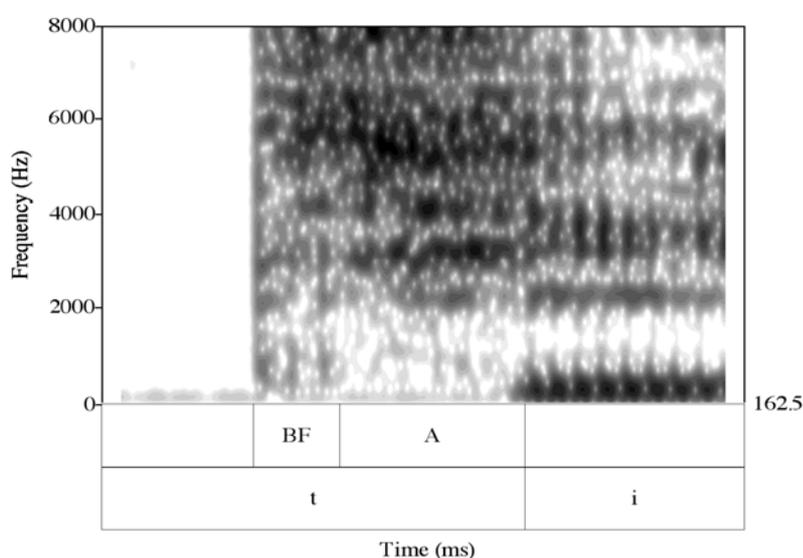


Figure 1: Spectrogram of a sample friction phase. BF = burst friction and A = aspiration.

As can be seen in Figure 1 burst friction differs considerably from aspiration. While burst friction immediately follows the closure phase of /t/, aspiration occurs between the burst friction and the following vocoid. Burst friction is generated at the supraglottal constriction and shows a spectral prominence in the

frequency range from 3500 to 7000 for an alveolar stop which is due to a front cavity that is 1-2 cm long, cf. e.g. Stevens et al. (1999).

Aspiration on the other hand is generated at the glottis and shows a stronger concentration of energy in the higher frequency region but also formant like peaks in lower frequency regions. Furthermore, it displays attenuation, or lack of the first formant, the so-called F1 cutback, cf. Liberman et al. (1958). If the stop is followed by a high vowel, as in Figure 1, aspiration overlaps with friction generated at the constriction of this vowel, cf. Hanson & Stevens (2003). Aspiration is considerably longer durationally than burst friction. What also distinguishes burst friction from aspiration is the amplitude, since the amplitude of the former decreases as the cross-sectional area of its supraglottal constriction increases, cf. Stevens (1998: 362). When the noise generated at the constriction becomes dominated by the noise generated at the glottis, then the amplitude of aspiration increases.

In our experiment we expect the duration of the friction phase for /tj/ to be significantly longer than that of /ti/ (= prediction 6a). Prediction (6b) similarly leads us to hypothesize that the friction phase of /dj/ should be significantly longer than that of /di/. Furthermore, we expect the duration of the friction phase of /di dj/ to be shorter than that of the sequences with /ti tj/ (= predictions 6c and 6d). Two experiments were conducted in order to test these predictions. In the first one the friction phase of /ti/ vs. /tj/ and /di/ vs. /dj/ in contrastive pairs of nonce words was measured, e.g. [ti'a] vs. [tja]. Since all four of the speakers in this experiment often glided the /i/ to [j] in items like [ti'a], we conducted a second experiment with near minimal pairs of nonce words such as [tjak] vs. [tik], in which the gliding of /i/ to [j] was not possible. It will be shown below that the results of Experiment 1 support (6c-d). The results of Experiment 2 support (6a-d).

4.1 Experiment 1

Four speakers of German (two male, CG and JD, and two female, SF and SH) participated in an acoustic study in which the duration of friction phase of the stops /t/ and /d/ was measured. The four speakers were asked to read the twenty four nonce sequences given in Table 1 in a randomized order. All are phonotactically well-formed in German.² Capitals indicate stress.

² Sequences of [di] and [ti] plus following vowel are not that common in German, but occur in foreign words such as *Diagramm* [diagʁam] 'diagram' and *Tiara* [tia:ʁa] 'tiara'. An optional gliding of the unstressed /i/ in these positions is possible. Examples containing [dj] occur in many (frequent) nonnative words, e.g. *Studium* [ˈʃtu:djʊm] 'studies' (sg.). Examples of [tj] sequences are rare but they are attested in *Adjektiv* [ˈatjɛkti:f] 'adjective'.

Table 1: Twenty four sequences involving /t d/ with /i j/ used in Experiment 1.

tiA	atiA	atsiA
tjA	atjA	atsjA
tiU	Atia	Atsia
tjU	Atja	Atsja
diA	adiA	astiA
djA	adjA	astjA
diU	Adia	Astia
djU	Adja	Astja

The nonce words were embedded in the frame sentence “Ich habe _ gesagt” ‘I said _.’ The subjects repeated the test words with the frame sentence five times at normal speed.

Figure 2 shows broad band spectrograms (with a window length of 5 ms and a dynamic range of 50 dB) for the examples /atiA/ and /atjA/ for speaker CG. The duration of the friction phase is indicated in the examples in the first row below each spectrogram (the /a/ at the beginning and at the end of each word are not given in their whole duration).

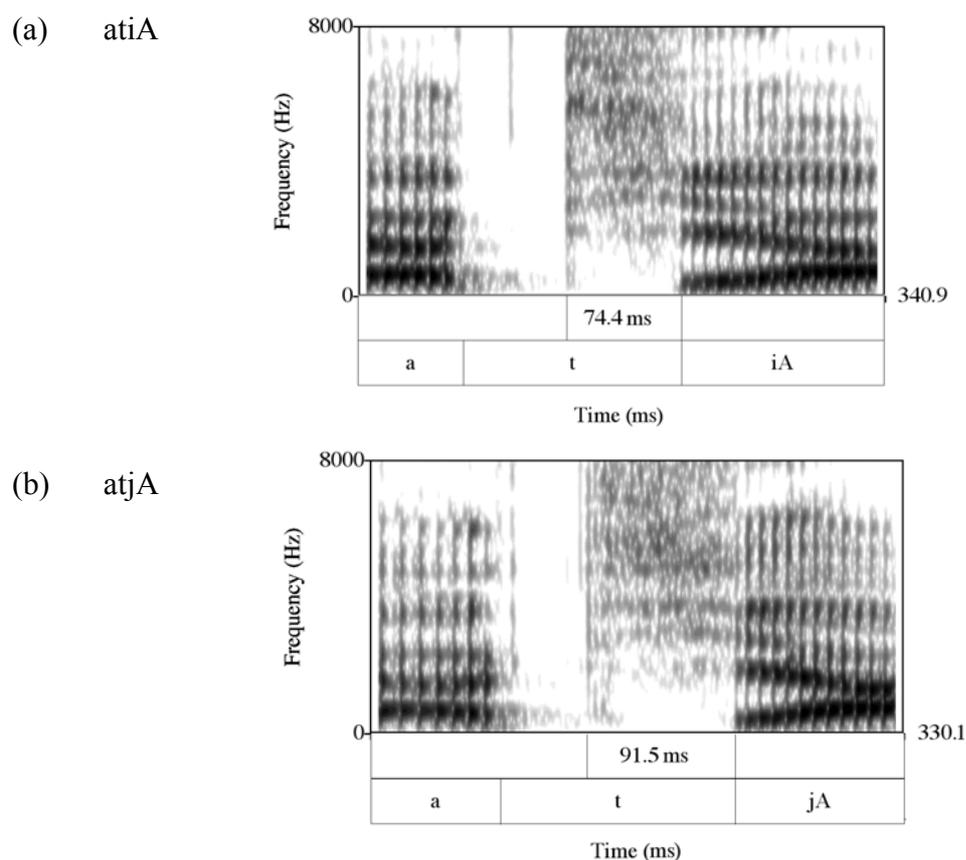


Figure 2: Spectrograms of (a) atiA and (b) atjA (German speaker CG).

See Hall (2004) for discussion of the distribution of [tj] and [tsj] sequences in Modern German.

A comparison of the two spectrograms in Figure 2 shows that the example with a following /i/ in (a) has a shorter friction phase than the one with a glide /j/ in (b).

Examples for a voiced stop followed by /i/ and /j/ are given in Figure 3, again from speaker CG.

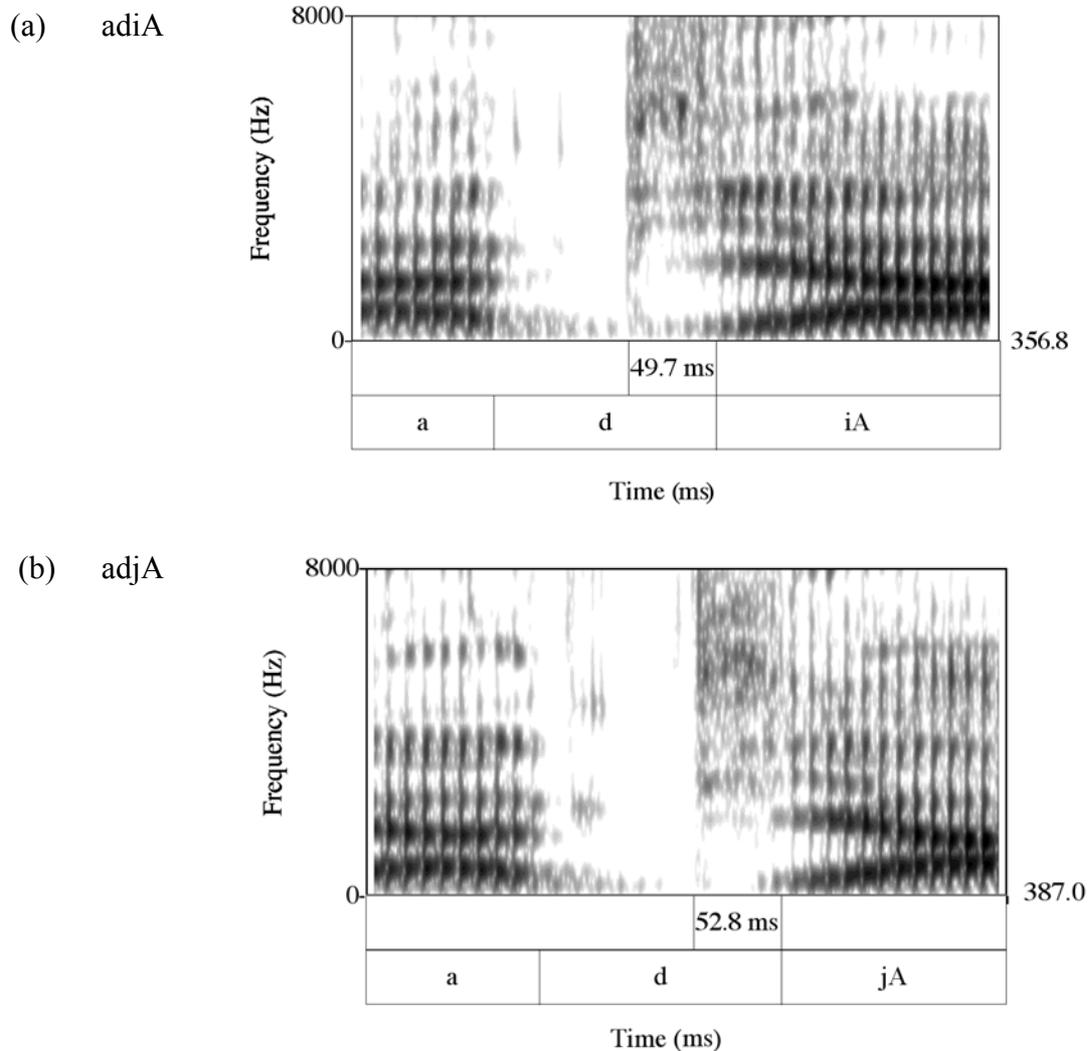


Figure 3: Spectrograms of (a) atiA and (b) atjA (German speaker CG).

For some speakers, there was no visible friction phase in the spectrograms of the voiced stop, e.g. the [d] in *Adja* for speaker SH in Figure 4. In cases like these we took 0 ms as the value for friction duration.

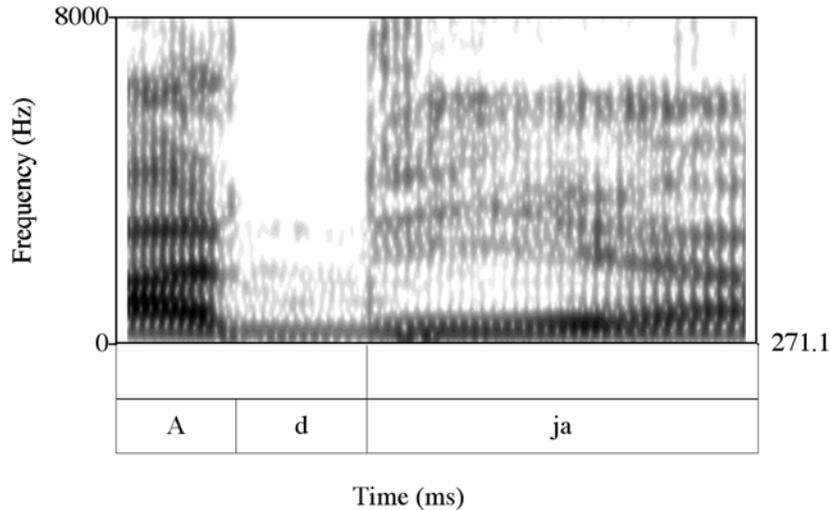
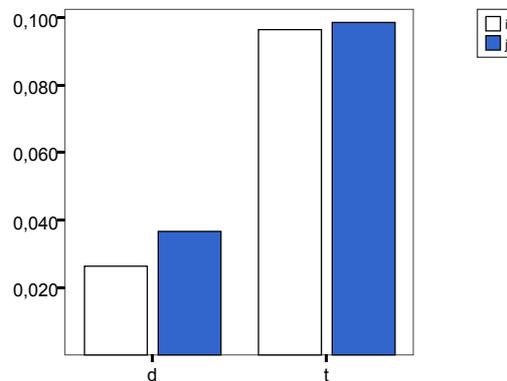


Figure 4: Spectrogram of adjA (German speaker SH).

For all four speakers together, the friction duration after /d/ is shorter than it is after /t/, as illustrated in the following Graph 1. The vertical axis indicates duration of the friction phase in [s]. The different shadings of the bars correspond to /i/ or /j/, as indicated in the legend to the right.



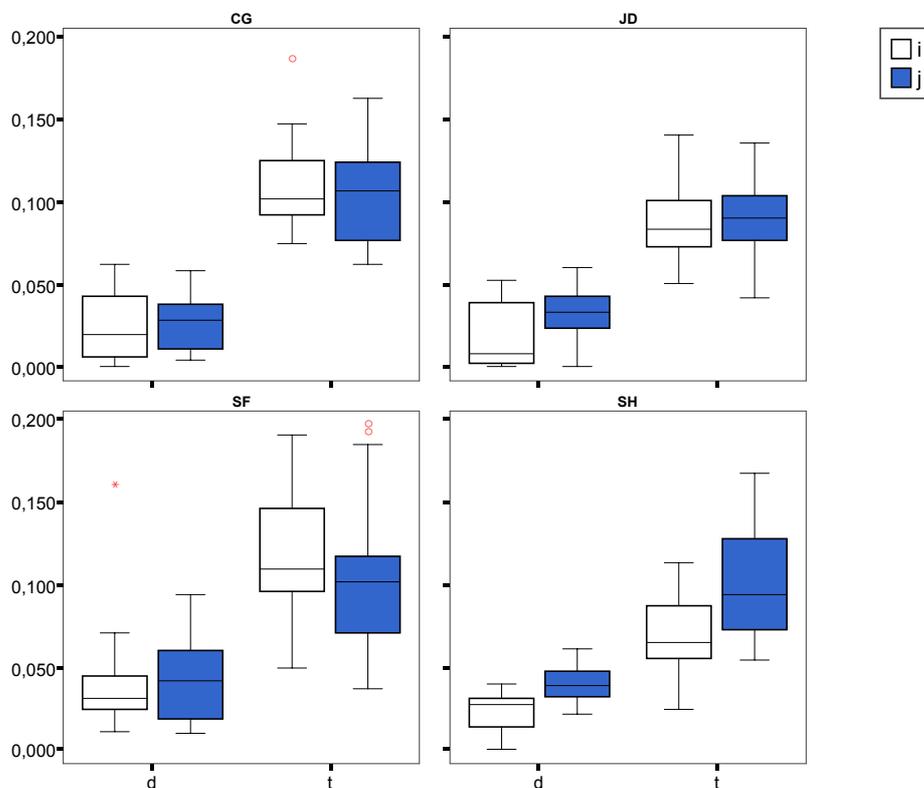
Graph 1: The average duration of friction in s after /d/ or /t/ over five tokens of the test words for all speakers.

The results summarized in Graph 1 are important because they support predictions (6b-d). A two-factorial ANOVA³ with /t, d/ and /i, j/ as independent variables and friction duration as dependent variable shows that the difference between the friction phase duration for /tj/ vs. /dj/ (=6c) and for /ti/ vs. /di/ (=6d) is highly significant ($F(1, 239) = 263.432, p < .001, F(1, 238) = 284.909, p < .001$). A comparison of the friction phase duration for /di/ vs. /dj/ (=6b) on the left of Graph 1 illustrates that the difference is also significant ($F(1, 159) = 8.944, p < .01$). We illustrate below that (6c-d) and to a lesser extent (6b) are also

³ All statistical calculations were made in SPSS 11.5.1.

borne out for each of the four speakers individually. By contrast, prediction (6a) in Graph 1 is not nearly as robust ($F(1, 318) = .417$ $p=.519$). Similar results were obtained for the individual speakers as well (see below).

Graph 2 shows the results of Experiment 1 split between speakers. Note that the order of the single box plots within each window corresponds to the order of bars in Graph 1. Each boxplot shows the median, quartiles, and extreme values within a category. The median value is shown by a horizontal line displayed in each of the boxplots. The vertical line shows 95% of the data and is limited by the highest and the lowest values. The box length is the interquartile range and it covers 50% of the data. It is limited by the first (25%) and third quartile (75%). Graph 2 presents also outliers marked by small circles above the boxes which represent cases with values between 1.5 and 3 box lengths from the upper or lower edge of the box.



Graph 2: The average duration of friction in s after /d/ or /t/ over five tokens of the test words split according to speakers. Boxplots correspond to means of friction duration.

An examination of Graph 2 illustrates that prediction (6c-d) are borne out but that (6a-b) are not. For all four speakers (6c-d) are confirmed because the difference in friction phase duration for /tj/ is significantly longer than for /dj/. Similarly, the friction phase in /ti/ is significantly longer than in /di/; speaker CG $F(1, 59) = 188.788$, $p < .001$, speaker JD $F(1, 59) = 140.243$, $p < .001$, speaker SF $F(1, 59) = 62.166$, $p < .001$, speaker SH $F(1, 58) = 69.382$, $p < .001$.

By contrast, predictions (6a) (and to a lesser extent 6b) are not consistently confirmed by the four individual speakers. Only speaker SH has a significant difference between /tj/ and /ti/ ($F(1,78) = 22.971$ $p < .001$). Speaker JD has a slightly longer friction phase duration for /tj/ than for /ti/, but this difference is not significant ($F(1,79) = .072$, $p = .789$). Speakers CG and SF show a reverse duration of friction phase than predicted by (6a): for both the friction duration for /ti/ is longer than for /tj/. (6b) derives support from speakers SH and JD, whose friction phase duration for /dj/ is significantly longer than /di/ (speaker SH: $F(1,39) = 21.018$, $p < .001$, speaker JD: $F(1,39) = 7.019$, $p < .05$). For speakers CG and SF /dj/ is longer (although not significantly) than /di/ which is the reverse from our prediction in (6b).

The unexpected results concerning predictions (6a-b) can be accounted for by the fact that most nonce words with /i/ plus vowel sequences surfaced as [j] and thus the minimal pairs illustrating /i/-/j/ contrasts were neutralized. The gliding of /i/ to [j] could be observed in a clear majority of the tokens containing [i] plus vowel for all speakers but SH. This point can be attributed to the fact that [i] and [j] are usually analyzed as allophones of /i/ in German which are distributed in such a way that the glide [j] only occurs before a vowel (see, e.g. Wiese 1996). In order to avoid the occurrence of gliding, we conducted a second experiment in which gliding is not possible.

One finding in Experiment 1 we did not predict is that the friction phase duration of /ti/ is significantly longer than that of /dj/ ($p < .001$ for all speakers). In section 5 we return to this point.

4.2 Experiment 2 and 3

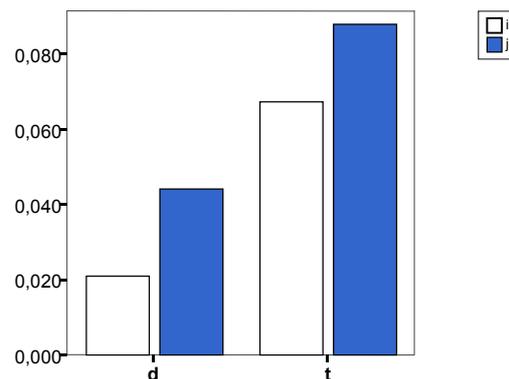
In the second experiment, four native speakers of German (the same subjects as in the first experiment) were asked to read 28 nonce words embedded in a carrier sentence “Ich habe _ gesagt” ‘I said _’. The nonce words in Table 2 were presented in a randomized order and the subjects repeated the sequences ten times at normal speed. Stressed syllables are indicated by capital letters.

Table 2: Twenty four sequences involving /t d/ with /i j/ used in Experiment 1.

Atik	aTIK	Adik	aDIK
Itik	iTIK	Idik	iDIK
Utik	uTIK	Udik	uDIK
Atjak	aTJAK	Adjak	aDJAK
Itjak	iTJAK	Idjak	iDJAK
Utjak	uTJAK	Udjak	uDJAK
Atjik	aTJIK	Adjik	aDJIK
Atjuk	aTJUK	Adjuk	aDJUK

In contrast to Table 1, there are no minimal pairs in Table 2 involving [i] and [j], e.g. *Atia* vs. *Atja*. Instead, we see in Table 2 that there are pairs like *Atik* (in which a consonant and not a vowel follows [i]) vs. *Atjak*. In this way we avoid the possibility of converting /i/ into [j] before a vowel because there is no vowel following the /i/ in items like *Atik*. Our results are summarized in the following graphs.

Graph 3 presents the average duration of friction after /d/ and /t/ when followed by /i/ or /j/ for all German speakers together.

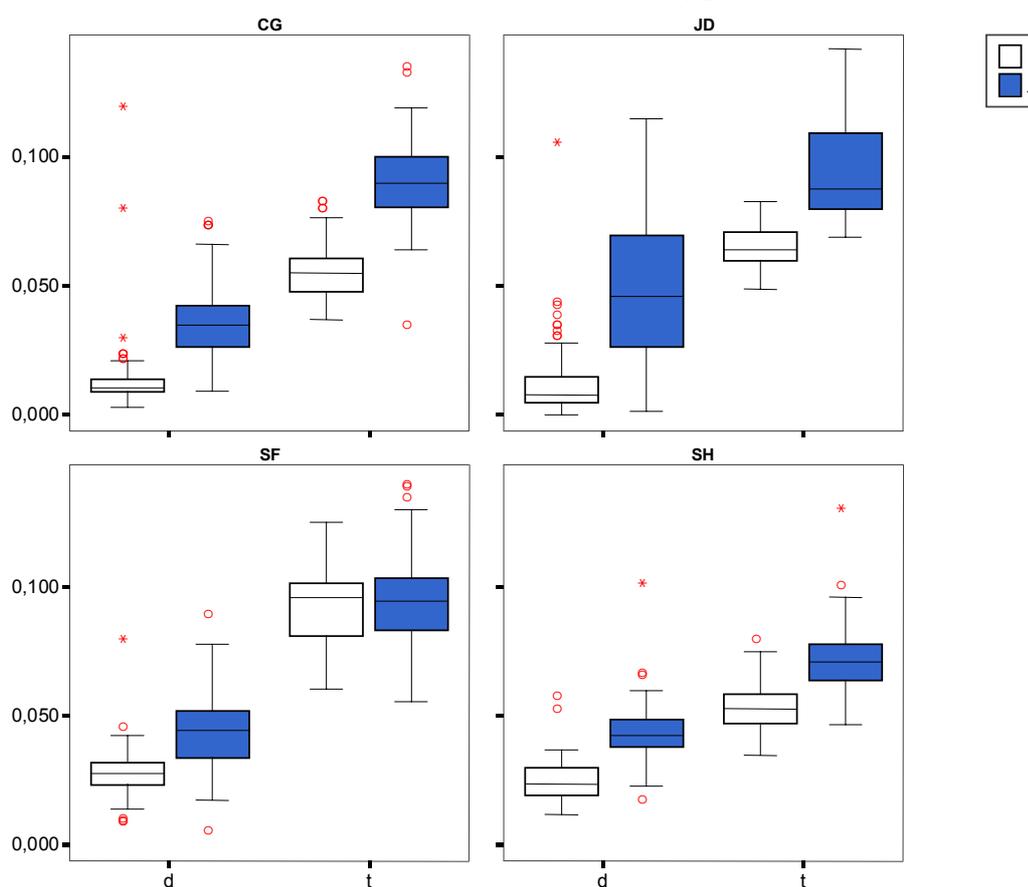


Graph 3: The average duration of friction in s after /d/ or /t/ for all German speakers.

The results of Experiment 2, as summarized in Graph 3, support predictions (6a-d). (6a) is supported because the friction duration for /tj/ is significantly longer than /ti/ ($F(1,634) = 189.122$ $p < .001$), as is (6b) because the friction duration for /dj/ is significantly longer than /di/ ($F(1,643) = 265.508$ $p < .001$). (6c-d) also derive support because the friction phase for /tj/ is significantly longer than for /dj/, ($F(1, 795) = 1112.039$, $p < .001$). For /ti/ vs. /di/ the results are also significant: ($F(1,482) = 887.476$, $p < .001$).

As in Experiment 1, the friction phase for /ti/ in Experiment 2 is significantly longer than /dj/ ($F(3,1278) = 809.933$, $p < .001$), the latter results follow from a post-hoc Scheffé test). See section 5 for discussion.

Graph 4 presents results in the form of box plots, as obtained for individual speakers. The order of the single box plots within each window corresponds to the order of bars in Graph 3. Note that apart from outliers there are also extreme cases marked by asterisks which represent cases with values more than 3 box lengths from the upper or lower edge of the box. Other parameters are the same as in Graph 2.



Graph 4: The average duration of friction in s after /d/ or /t/ split according to speakers. German speakers averaged over vowels and stress condition.

Graph 4 illustrates that all four speakers of German individually support predictions (6c-d): Statistically, the friction phase is significantly longer in /tj/ than in /dj/ (for speaker CG $F(1,198)=687.597$ $p<.001$, speaker JD $F(1,198)=164.580$ $p<.001$, speaker SF $F(1,198)=506.637$ $p<.001$, and speaker SH $F(1,198)=299.221$ $p<.001$) and the friction is significantly longer in /ti/ than in /di/ for speaker CG $F(1,120)=244.719$ $p<.001$, speaker JD $F(1,198)=164.580$ $p<.001$, speaker SF $F(1,119)=838.043$, $p<.001$, and speaker SH $F(1,120)=278.373$ $p<.001$. Prediction (6b), that the friction phase in /dj/ is of longer duration than in /di/, is also significantly supported by all four speakers (for each speaker $p<.001$). Prediction (6a), that the friction phase in /tj/ is of longer duration than in /ti/, is supported by speakers CG, JD and SH but not by speaker SF who has almost equal values for /tj/ and /ti/, the difference is not significant: $F(1,157) = 156$, $p=.693$.

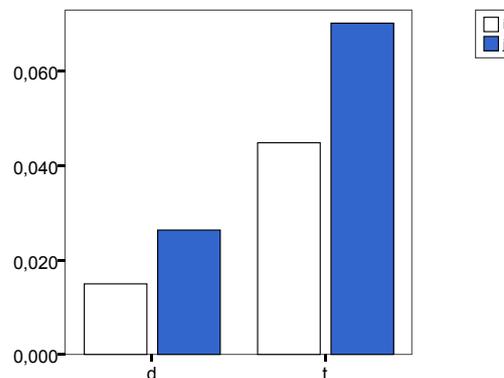
In light of these results the question arises whether the predictions in (6) and the proposed hierarchy in (7) are language specific or whether they can be confirmed by data from other languages. For reasons of comparison we included an investigation of Polish. To be sure, any one of a number of languages could be taken for comparison, but we opted for Polish because this language differs from German in terms of voicing: While German contrasts voiceless with

voiceless aspirated stops, i.e. /p t k/ vs /p^h t^h k^h/ (see, e.g. Jessen & Ringen 2002), Polish contrasts voiced and voiceless stops, i.e. /b d g/ with /p t k/.

This language-specific difference in the realization of the voicing contrast leads us to expect a difference in overall friction length: Since the duration from the burst until the onset of a vowel is shorter for truly voiced stops than for voiceless ones and aspiration adds to this length of friction (see the beginning of section 4), we expect Polish to show generally shorter friction duration than German.

Four native speakers of Polish (two female DZ, MR and two male SL, KZ) were asked to read the nonce words presented in Table 2 in the carrier sentence ‘Powiedziałem ... do ciebie’ ‘I said... to you’ at normal speed.⁴

The average duration of the friction phase duration as obtained by all four speakers together for /di/, /dj/, /ti/ and /tj/ is presented in Graph 5. The friction is in s.



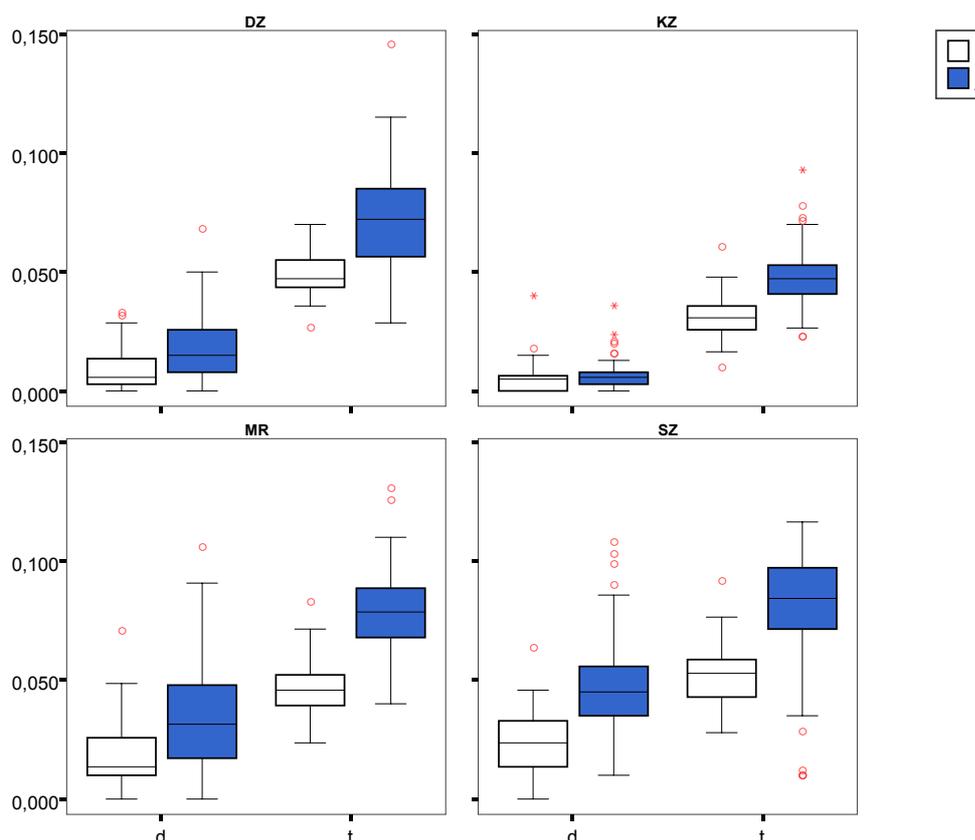
Graph 5: The average duration of friction in s after /d/ or /t/ for all Polish speakers.

The results presented in Graph 5 support the predictions in (6a-d). Again, a two-factorial ANOVA with /d, t/ and /i, j/ as independent variables and friction duration as dependent variables reveals that all differences presented in Graph 5 are highly significant: /tj/ vs. /ti/ ($F(1,624) = 248.425$ $p < .001$), /dj/ vs. /di/ ($F(1,643) = 49.616$ $p < .001$), /tj/ vs. /dj/ ($F(1, 789) = 740.485$ $p < .001$), and /ti/ vs. /di/ ($F(1, 478) = 559.761$ $p < .001$).

As for German, the friction duration for /ti/ in Polish is significantly longer than for /dj/ ($p < .001$). See section 5 below for discussion.

Graph 6 presents the results split between the individual speakers in the form of box plots.

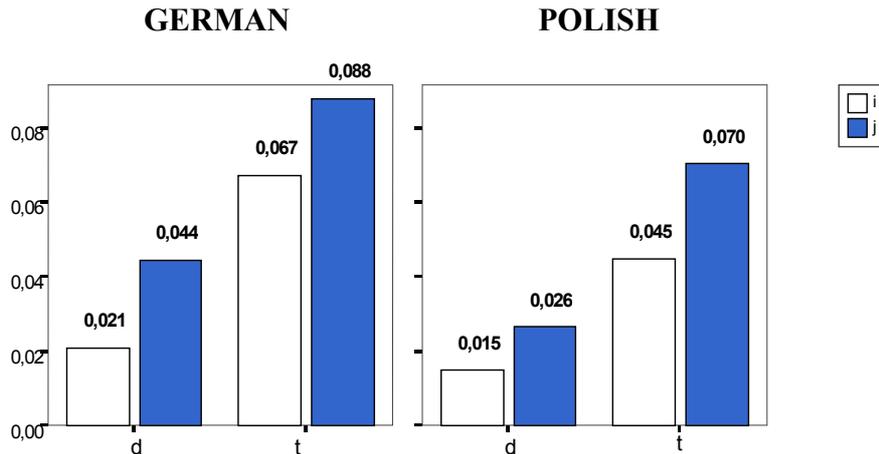
⁴ The sequences [tj dj] plus vowel are well-formed in Polish. They occur in words of foreign origin, e.g. [dj]alekt ‘dialect’, [tj]ara ‘tiara.’



Graph 6: The average duration of friction in s after /d/ or /t/ split according to speakers. Polish speakers.

The results in Graph 6 show that predictions (6a, c-d) are supported by the results from all four speakers (for all three predictions: for (6a) speaker DZ $F(1,157) = 68.830$ $p < .001$, speaker KZ $F(1, 158) = 95.819$, $p < .001$, speaker MR $F(1,148) = 179.022$ $p < .001$, speaker SZ $F(1,158) = 90.152$ $p < .001$, for (6c) speaker DZ $F(1,199) = 542.550$ $p < .001$, speaker KZ $F(1, 199) = 1119.964$, $p < .001$, speaker MR $F(1,189) = 252.337$ $p < .001$, speaker SZ $F(1,199) = 137.535$ $p < .001$ and for (6d) speaker DZ $F(1,118) = 599.272$ $p < .001$, speaker KZ $F(1, 119) = 389.158$, $p < .001$, speaker MR $F(1,119) = 176,979$ $p < .001$, speaker SZ $F(1,119) = 134.688$ $p < .001$). (6b) is supported by three speakers (speaker MR $F(1,160) = 25.943$ $p < .001$, speaker SZ $F(1,160) = 66.688$ $p < .01$). By contrast, speaker KZ did not have a significant difference between the friction phase duration of /di/ vs. /dj/. $F(1,160)=1.644$ $p=.202$

Graph 7 shows a comparison of the results for German and Polish. The numbers on the top of the bars indicate mean friction duration in s as obtained in the given sequences.



Graph 7: Comparison of the average duration of friction in s after /d/ or /t/ for German and Polish speakers.

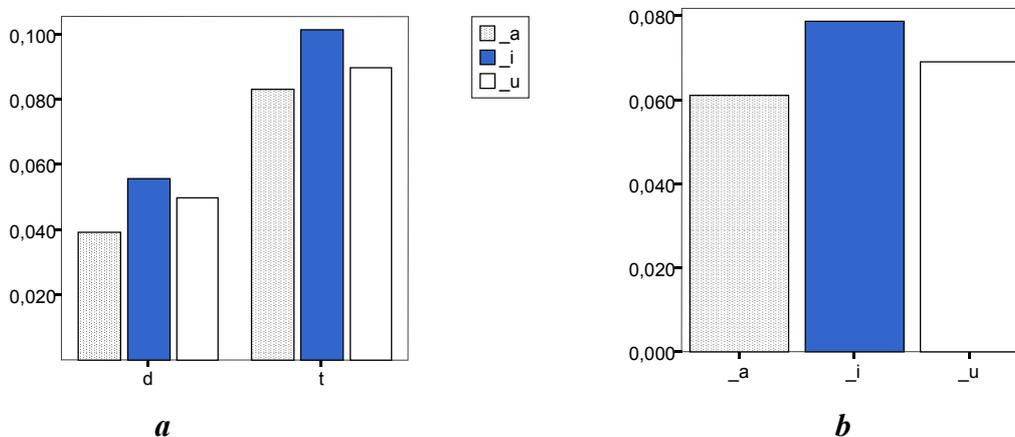
The results of Graph 7 show that the friction phase is on average longer in German than in Polish, which is in accordance with our expectations.

In addition to the voicing parameter (i.e. /t/ vs. /d/) and the parameter for the trigger (i.e. /i/ vs. /j/) there are three other factors which could potentially influence the friction phase duration. We list them in (8):

- (8) a. The quality of the vowel following /tj/ and /dj/
- b. The quality of the vowel preceding /tj/, /ti/, /dj/ and /di/
- c. The stress of the vowels adjacent to /tj/, /ti/, /dj/ and /di/

In the remainder of this section we report on the results of Experiment 2 with respect to the three parameters in (8).

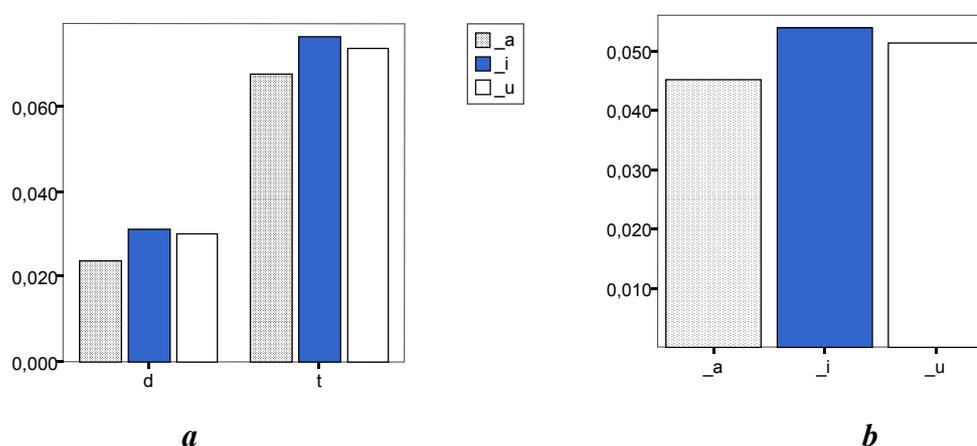
As far as parameter (8a) is concerned, we did not include a vowel following /ti/ and /di/ sequences, as e.g. /atia/ or /atiu/, in order to exclude a possibility of realizing them with a glide, i.e. [atja], [atju]. Graphs 8a and b shows average friction durations dependent on the vowel following /tj/ and /dj/.



Graph 8: The average duration of friction in s dependent on the following vowel for /dj/ and /tj/ separately (a) and for /tj/ and /dj/ together (b). German speakers.

The results provide evidence that the friction duration in German /tj/ and /dj/ sequences is influenced by the following vowel. In particular, the longest friction occurs if the sequences are followed by *_i* and the shortest when they are followed by *_a*. Apart from the difference in mean friction duration between *_u* and *_i* in the /dj/ context (cf. [adjuk] vs. [adjik]), all other results are significant in comparison to each other. Taken together, an average friction for /tj/ and /dj/ leads to statistically significant differences: a following *_i* results in the longest duration followed by a following *_u* and finally by *_a*, cf. Graph 8b. According to a post-hoc Scheffé test all differences are statistically significant (*_i* vs. *_a* $F(2,795) = 23.664$ $p < .001$, *_i* vs. *_u* $p < .05$, *_u* vs. *_a* $p < .01$).

The influence of the vowel after /tj/ and /dj/ in the Polish items is given in Graphs 9a and b.

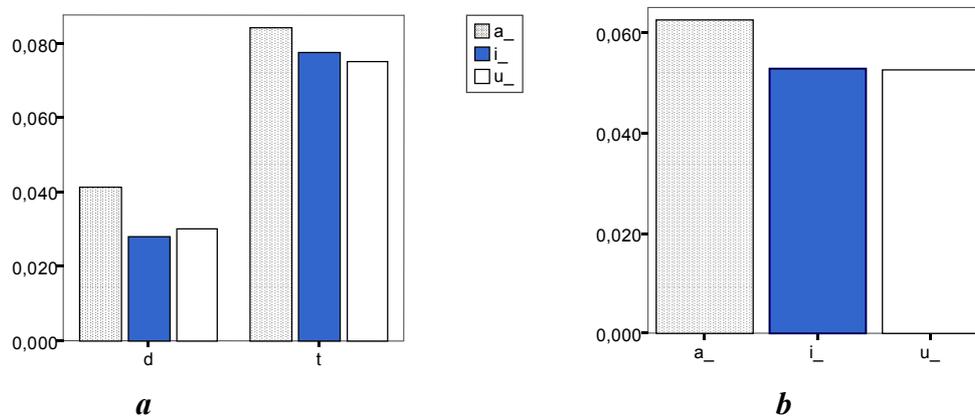


Graph 9: The average duration of friction in s dependent on the following vowel for /dj/ and /tj/ separately (a) and for /tj/ and /dj/ together (b). Polish speakers.

As in German, the *_i* context in Polish triggers the longest duration in /tj/ and /dj/ sequences. The second longest friction occurs in the *_u* context followed by the *_a* context. However, from a statistical point of view the influence of the following vowel is significant in only two contexts: for /tj/ the difference *_a* vs. *_i* is significant ($F(2,385) = 5.630$ $p < .01$) and for /dj/ the same difference is only weakly significant ($F(2,399) = 4.659$, $p < .05$). Taking the results for /dj/ and /tj/ together, Graph 9b shows that the sequences in the *_i* context have the longest friction duration which is only slightly longer than in the *_u* context and statistically not significant. The only statistically significant difference is attested between *_a* vs. *_i* contexts ($F(2, 785) = 5.509$, $p < .05$). The influence of vowel quality on a preceding glide and its consequence on friction duration is discussed in section 6.4 below.

Another factor we varied in this experiment was the preceding vowel, cf. (8b). Here all items of Table 2 were included in the calculations. Graphs 10a and

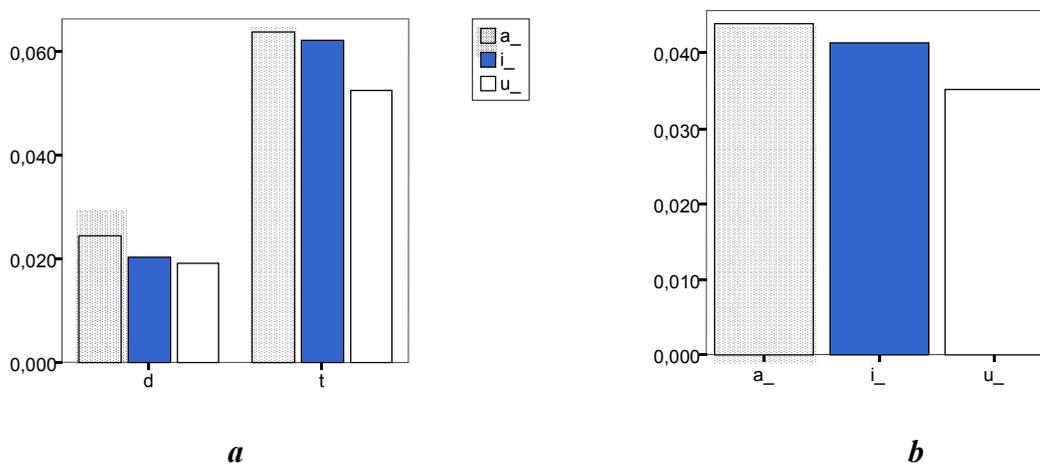
10b show how the vowels a_, i_ and u_ influence the friction length of the stop-vocoid sequences in German.



Graph 10: The average duration of friction in s dependent on the preceding vowel for /dj/ and /tj/ separately (a) and for /tj/ and /dj/ together (b). German speakers.

Results obtained for German show that if /dj/ or /tj/ are preceded by a_ the friction is longest. Other contexts do not show any clear tendencies, cf. Graph 10a. Statistically, the differences in the a_ vs. u_ context is highly significant for /t/ and /d/ (for /t/ $F(2,634) = 12.244$, $p < .001$, for /d/ $F(2,634) = 30,297$ $p < .001$). The difference between the a_ vs. i_ context is significant in /t/ sequences ($p < .01$) and highly significant in /d/ sequences ($p < .001$). The i_ vs. u_ context is significant neither in /t/ nor in /d/ sequences. This context does not influence the friction length if all items are calculated together, cf. Graph 10b. Here, the contexts a_ vs. u_ and a_ vs. i_ are highly significant (for both $F(2,1278) = 16.460$ $p < .001$).

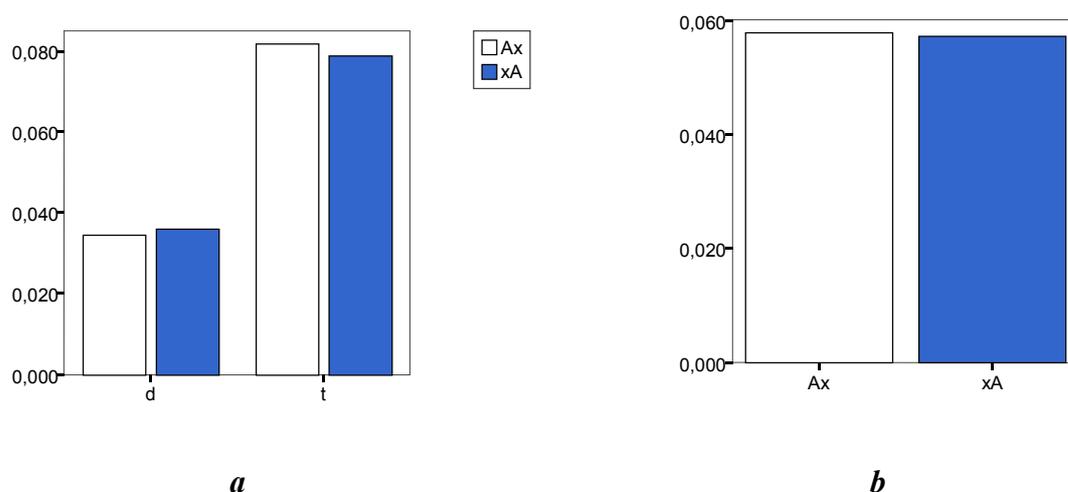
Graphs 11a and 11b present the results of the influence of the preceding vowel obtained for Polish speakers.



Graph 11: The average duration of friction in s dependent on the preceding vowel for /dj/ and /tj/ separately (a) and for /tj/ and /dj/ together (b). Polish speakers.

As Graph 11a shows, the results for the Polish speakers are similar to those of the German speakers in experiment 2: the longest friction in Polish is attested when /t/ and /d/ are preceded by a_. When the stops are preceded by i_ the friction is shorter and when preceded by u_ the friction is the shortest. Not all differences are statistically significant. In /t/ sequences the context a_ vs. i_ is not significant, while u_ vs. i_ is significant ($F(2,624) = 13.583, p < .01$) and a_ vs. u_ is highly significant ($p < .001$). As far as /d/ sequences are concerned only a_ vs. u_ context is (weakly) significant ($F(2,643) = 4.288, p < .05$), while other contexts are not significant. Taking /t/ and /d/ sequences together, cf. Graph 11b, the a_ vs. u_ context is highly significant ($F(2, 1268) = 9.308, p < .001$) and u_ vs. i_ is weakly significant ($p < .05$). The difference in the a_ vs. i_ context is not significant. Possible explanations for the influence of the preceding vowel on friction duration are discussed in section 6.4.

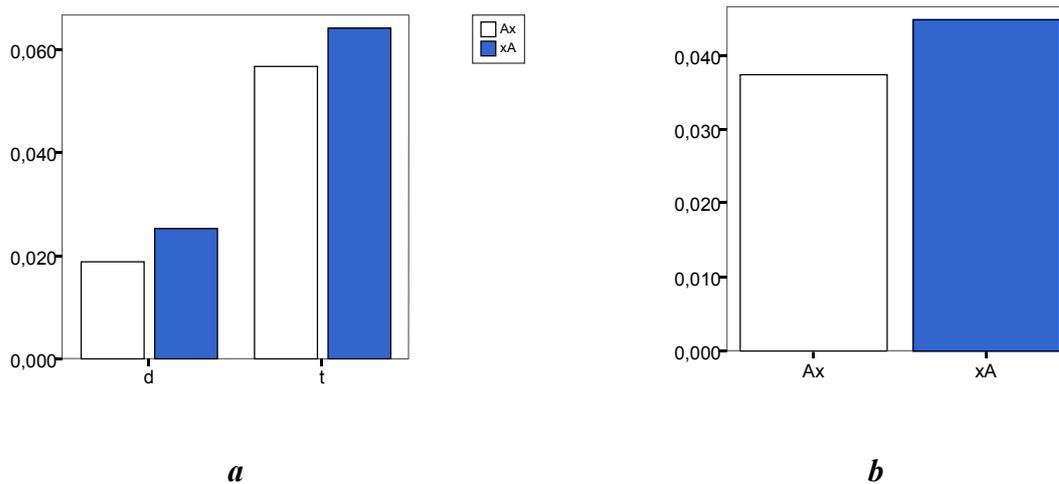
Finally, the influence of stress on the friction duration was investigated, cf. (8c). Graphs 12a and 12b show the friction length in German as influenced by stress. Note that Ax denotes a bisyllabic item with initial stress, and xA a bisyllabic item with final stress.



Graph 12: The average duration of friction in s dependent on stress for /dj/ and /tj/ separately (a) and for /tj/ and /dj/ together (b). German speakers.

The results in Graphs 12a and 12b show that stress does not have a considerable influence on the friction duration in German. If /tj/ and /ti/ occur in a prestressed position, the friction duration is slightly longer than when they occur in a stressed position. The opposite is true for /dj/ and /di/ sequences, cf. Graph 12a. None of the differences is statistically significant. Graph 12b confirms that there is almost no difference in friction duration when the tested items are evaluated together.

Polish speakers performed differently with respect to the influence of stress. The results are depicted in Graphs 13a and 13b.



Graph 13: The average duration of friction in s dependent on stress for /dj/ and /tj/ separately (a) and for /tj/ and /dj/ together (b). Polish speakers.

The differences in friction duration for /d/ and /t/ presented in Graph 13a are statistically highly significant (for /d/ $F(1,643) = 15.254$, $p < .001$ and for /t/ $F(1,624) = 16.590$, $p < .001$) showing that in Polish the friction duration is longer in a stressed than in a prestressed position for both /t/ and /d/ sequences. A similar result is obtained by calculating the friction duration without splitting the data: a longer friction is attested in a stressed than in a prestressed position ($F(1,1268) = 20.928$, $p < .001$), cf. Graph 13b. Section 6.3 will give possible explanations for this language-specific dependence of friction duration on stress.

5 Findings

In addition to supporting the four predictions in (6) and the assibilation hierarchy in (7), the three experiments described in the previous section led to several findings not discussed in the literature on assibilations. In this section we summarize these findings.

One finding supported by our experiments which was not made by Hall & Hamann (2003) or in any earlier study to our knowledge is stated in (9):

- (9) The friction phase in /dj/ is shorter durationally than in /ti/.

The finding in (9) suggests that the assibilation hierarchy presented above in (7) should be modified as in (10):

- (10) An assibilation hierarchy:
 /tj/ < /ti/ < /dj/ < /di/

If the hierarchy in (10) is truly universal then one would expect it to be reflected in a typology like the one summarized above in (4). For example, one should be able to find evidence for an implication of the form “if /d/ assimilates before /j/ then /t/ assimilates before /i/”. This is clearly a question we leave open for further research.

Three additional results of the three experiments conducted are summarized in (11):

(11) Three additional results:

- a. The vowel following /tj/ and /dj/ influences the friction phase in the sense that /i/ results in a longer friction phase than /u/ or /a/;
- b. The vowel preceding /tj/ etc. influences the friction phase in the sense that /a/ results in a longer friction phase than /i/ or /u/;
- c. Stress of the vowels adjacent to /tj/ etc. does not influence the duration of the friction phase in German but it does in Polish in the sense that there is a longer friction phase if the stress follows the /tj dj/ sequence.

Again, results like the ones in (11) can be (dis)confirmed by considering phonological assimilation rules from the typological perspective.

6 Discussion

In this section, possible explanations for the attested differences in friction length are discussed. Section 6.1 focusses on the difference in outcome for /i/ versus /j/, section 6.2 on the difference for /d/ versus /t/, section 6.3 on the influence of stress, and section 6.4 on the influence of the quality of the preceding and following vowel.

6.1 *The distinction between /i/ and /j/*

All three acoustic experiments show that a coronal stop (be it voiceless or voiced) followed by a palatal glide /j/ has a longer total friction phase than a coronal stop followed by the vowel /i/. Three possible explanations for these results can be offered, two articulatory and one aerodynamic.

The first explanation is based on the articulatory difference between the vowel /i/ and the glide /j/: the palatal glide might be articulated with a narrower constriction, which then causes greater impedance to the escaping air and thus a longer friction phase, cf. Klatt (1975) and Ohala (1983). Data on the articulatory difference between palatal glide and high front vowel to attest this point are scarce. In Figure 5 we have provided an x-ray tracing for /i/ and /j/ from one native speaker of Polish (from Wierzchowska 1971).

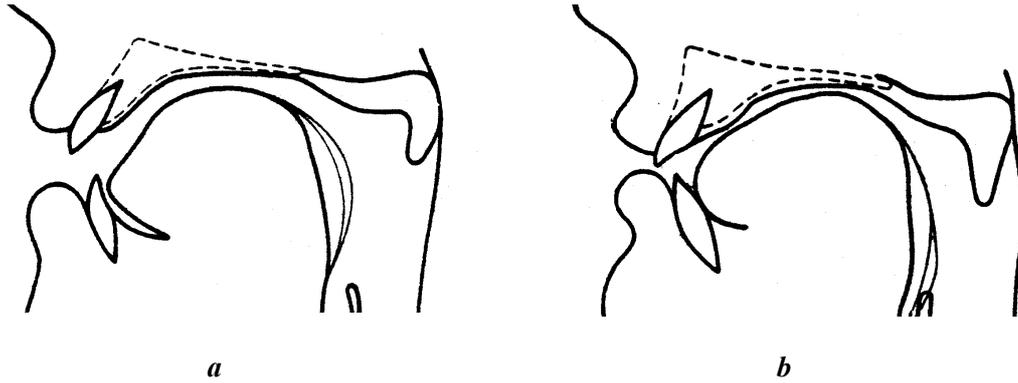


Figure 5: X-ray tracings for (a) /i/ and (b) /j/ in Polish. The thin lines at the back of the tongue indicate the lateral tongue positions.

A comparison of the tongue body in /i/ vs. /j/ reveals an obvious difference between the two: For the latter sound the constriction is longer and narrower than for /i/.

The difference in degree and length of constriction between /i/ and /j/ for the Polish speaker in Figure 5 is also assumed for the languages investigated by Maddieson and Emmorey (1985). They compared the formant frequencies of the palatal glide and the high front vowel in Amharic, Yoruba, and Zuni and found that in all three languages the glide has a lower first formant frequency than the vowel. Their conclusion is that the glide /j/ is produced with a narrower constriction than the vowel /i/.⁵

In other languages the difference in length and degree of constriction is not as obvious as in Polish. For example, Wängler's (1961) x-ray tracings of German /i:/ and /j/ for one speaker show that the two segments are articulated almost identically.

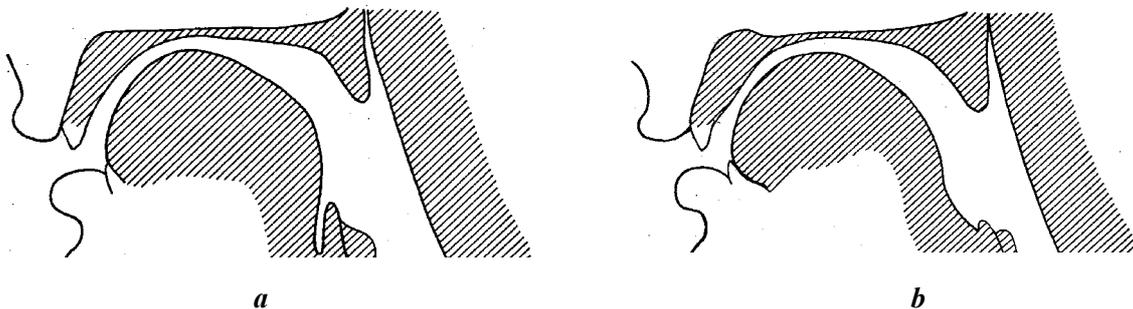


Figure 6: X-ray tracings for (a) /i:/ and (b) /j/ in German.

⁵ The recordings leading to this result only include tokens of /i/ in palatal glide context (in the nonsense word *iji*). However, the glide in this sequence might be articulated with a closer constriction than in other vocalic contexts in order to maintain a perceptual difference between glide and vowel.

A comparison of Figure 6a with 6b reveals that the palatal glide /j/ has a minimally longer constriction, i.e. the tongue front is raised a bit further than for the /i/. At this point one can only speculate that this minor difference between length of constriction between /i/ and /j/ might be responsible for the longer duration of the friction phase in /tj dj/ vs. /ti di/.

Like German, Romanian does not seem to support an articulatory difference between glide and vowel, either. Chitoran (2002, 2003) investigated the phonetic difference between the high front vowel /i/ and the glide /j/ in Romanian, and measured their friction duration after the stop /p/,⁶ and the formant values at the starting point of both segments. The duration of the friction phase was expected to be longer for the glide, indicating a narrower constriction. However, no significant differences could be found. Furthermore, Chitoran expected significant differences in the formant values at the beginning of the two segments. She found – contrary to what one would expect if /j/ has a narrower constriction than /i/ – that the second formant was higher in the vowel for two of her three speakers. Chitoran (2003:3016) interprets the lower F2 values for the glides as a ‘target undershoot’, which means that the glide is articulated with even less constriction than the vowel in the same context.

The results of these cross-linguistic studies might indicate a language-specific difference between the glide /j/ and the vowel /i/: In German, the two seem to be articulated similarly, in Romanian the glide /j/ seems to be articulated with a less narrow constriction than the vowel /i/, and in Amharic, Yoruba, Zuni and Polish the glide seems to have a more narrow constriction. However, it has to be kept in mind that these are mostly the interpretations of acoustic studies and that neither Chitoran, nor Maddieson and Emmorey, nor the present authors conducted articulatory studies on the difference between /i/ vs. /j/ to test these interpretations.⁷

A second explanation for the difference in friction length between glide and vowel could be the influence of the glide on the preceding stop. As B. Hurch (p.c.) pointed out to us, the glide can cause an inherent palatalization of the preceding stop: the stop before the glide is articulated with a raised tongue

⁶ Chitoran’s definition of friction duration differs from the one applied in the present article: it is the duration measured from the end of the [p] release burst to the onset of F1 (Chitoran 2003: 3014).

⁷ The cross-linguistic difference between high front vowel and the corresponding glide might be based on the fact that languages differ in their realization of the vowel [i], as P. Boersma (p.c.) pointed out to us. Thus, a very high [i] as it is the case in German might result in little articulatory difference between vowel and glide, whereas a lower [i] as we assume to be the case in Polish might result in a bigger articulatory difference between vowel and glide. We have at present no comparative data from German, Polish or any other language which might support this hypothesis.

middle, which can, due to a longer constriction, result in more friction than the non-palatalized stop before the high front vowel. The stop before the high front vowel, on the other hand, does not show this inherent palatalization. Like the first articulatory proposal, this explanation still has to be tested.

The third and final explanation proposed here is not based on articulation but on aerodynamics (see also the authors referred to in note 1). It could be argued that the glide is produced with stronger airflow than the vowel, and that the higher airflow results in longer friction noise following the release of the stop. We are currently testing this hypothesis (Hall, Hamann & Zygis in prep.), and preliminary results show that the /j/ following coronal stops has stronger airflow than the /i/ in the same context (results from two German speakers).

6.2 *The distinction between /t/ and /d/*

The voiced stops /di dj/ generally show a shorter friction duration than the corresponding voiceless stops /ti tj/. One reason for this observation is that whereas the open vocal folds for a voiceless stop allow air to pass unimpeded, the vibrating vocal cords of the voiced stop restrict the air flow volume. This results in less air pressure behind the alveolar constriction for voiced stops and thus less friction noise at the release of the voiced stop. Furthermore, the voicing of stops requires a difference between subglottal and supraglottal pressure (in order to let the vocal folds vibrate), which is usually maintained by pharyngeal expansion and larynx lowering (Kent & Moll 1969, Perkell 1969, Bell-Berti 1975). Pharyngeal expansion also results in less air pressure at the constriction and less friction at the stop release (Ohala & Riordan 1979). The higher subglottal pressure and lower intraoral air pressure for voiced stops compared to voiceless stops has been attested by Netsell (1969) for American English.

Another factor to consider is a possible difference in articulation between the voiced and the voiceless stop. If /t/ is a laminal sound and /d/ an apical (as in some West-African languages, see Ladefoged 1964)⁸, then one might expect the former to assibilate, as apicals should be less prone to assibilation than laminals due to their smaller oral closure. Since the two experiments described in section 4 were acoustic and not articulatory, we do not know if the subjects had a distinction between laminal vs. apical /t d/. However, this hypothesis is testable in an articulatory experiment.

⁸ The difference between tongue tip and tongue blade reported by Ladefoged (1964) goes together with a difference in place of articulation, with apicals being articulated in the (post-) alveolar area, and laminals in the dental area.

6.3 *The influence of stress*

Stress was shown to be an influencing factor on friction duration in Polish but not in German, as shown in Graphs 12 and 13: the stressed stops in Polish had significantly longer friction than the unstressed ones. This language-specific difference might be due to the realization of stress in Polish versus German. To account for our data, Polish should have a more enforced articulation of the stop in stressed position than in unstressed position, whereas German should not show such a difference. At present, we do not have comparative data on German and Polish to support this hypothesis.

6.4 *The quality of adjacent vowels*

The vowel preceding /ti/ etc. was another factor demonstrated in our experiments to influence the friction phase. Graphs 10 and 11 showed that a preceding /a/ yields longest friction duration, whereas preceding /u/ and /i/ differ only slightly in their influence on the friction.

Although this point requires a detailed articulatory and aerodynamic investigation, we think that the longest friction in the case of the preceding /a/ may be influenced by its largest oral cavity (the tongue body is flat) and consequently by the largest air volume in comparison to /u/ and /i/ where the tongue body is in a higher position. The largest air volume needs the longest time to pass the supraglottal constriction and thus results in the longest friction. Unfortunately, we are not aware of any study dealing with this issue.

The final parameter to account for is the influence of the vowel following stop plus /j/ sequences. In section 4.2 we saw that the friction duration of a stop is considerably longer if an /i/ follows compared to a following /a/ (see Graphs 8 and 9). The influence of the following /i/ can be ascribed to the fact that languages generally avoid [ji] (and [wu]) sequences for perceptual reasons, as a number of studies on typologically diverse languages have shown (see, for example Kawasaki 1982). If [ji] sequences occur in a language, the glide seems to be articulated with a higher tongue position than in other glide vowel sequences.⁹ This higher tongue position then causes longer friction at the release of the stop into the glide, resulting in a longer friction duration observed in our second experiment. The raising of the tongue position for the glide can be argued to be a gradual process, dependent on the quality of the following vowel. This is attested by the fact that a following /u/ results in longer friction phase

⁹ Evidence for this assumption is given by Laver (1994:298), who remarks that the articulatory starting point for [j] in *ji* sequences of English and the Chentu dialect of Chinese is normally slightly closer and fronter than for [i].

than a following /a/ (this difference was shown to be significant for German but not for Polish): /u/ is also articulated with a high tongue body, requiring an adjacent glide with a high tongue body position. These hypotheses on the influence of the following vowel on stop-glide sequences could in principle be tested articulatorily.

7 Alternative explanation: Center of gravity

In the preceding sections we only considered friction duration as an explanation for the assibilation hierarchy in (10) and the implications in (5) – repeated in (12) and (13) for convenience:

(12) An assibilation hierarchy
/tj/ < /ti/ < /dj/ < /di/

(13) Two implications:

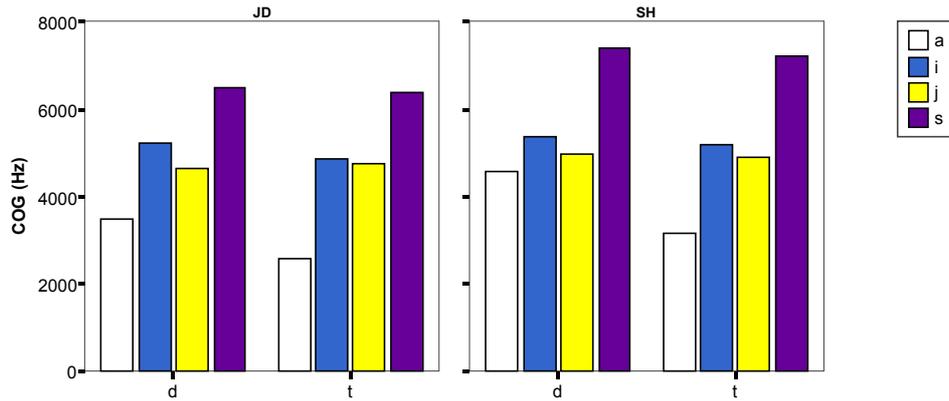
- a. Assibilation cannot be triggered by /i/ unless it is also triggered by /j/.
- b. Voiced stops cannot undergo assibilations unless voiceless ones do.

The assibilation hierarchy in (12) might also be attributable to other acoustic factors, such as the spectral similarities between the friction of the alveolar fricative part of affricates like /ts dz/ and the friction before /i j/ in sequences like /ti tj di dj/. The assibilation hierarchy in (12) would therefore be supported if the friction in /s z/ were spectrally more similar to the friction in /tj/ than the one in /ti dj/, and more similar to the friction in /ti dj/ than in /di/.

In order to test this alternative explanation we conducted an additional pilot study in which we measured the center of gravity (henceforth: COG) (cf. Jassem 1979, Nittrouer, Studdert-Kennedy & McGowan 1989, Forrest et al. 1998, Jongman, Wayland & Wong 2000, Gordon, Barthmaier & Sands 2002). COG is the average of frequencies over the entire frequency domain weighted by the amplitude. In our measurements, the weighting was done by using the power spectrum (p=2) in PRAAT. For the COG calculations we excluded the burst (ca. 10 ms) and measured the friction up to the beginning of continuous formants. To exclude the influence of the fundamental frequency in the voiced items, we band pass filtered all signals with a pass Hann band of 500 – 12000 Hz (and a smoothing of 100 Hz).

COG values were measured for the items in Table 3 spoken by two German speakers (female SH and male JD) and two Polish speakers (female MZ and male SL). The items were repeated five times in the same carrier sentences as used in section 4.2 above.

The results for the German speakers are presented in Graph 14.



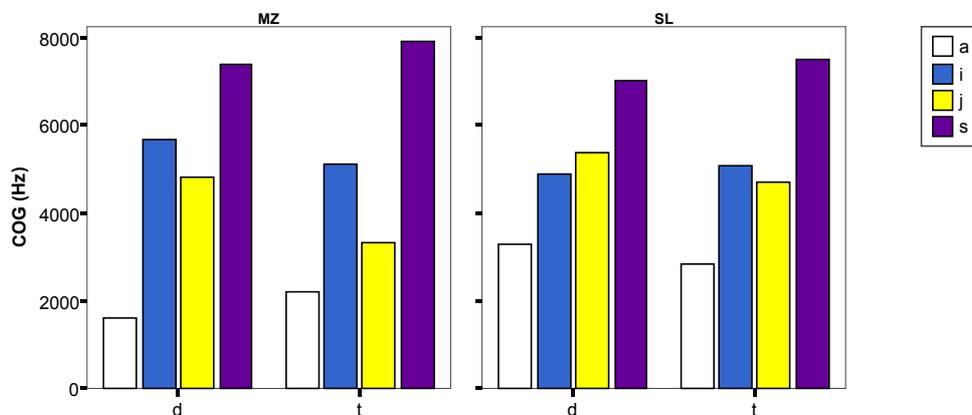
Graph 14: The average COG in Hz for the two German speakers. The context of the following ‘s’ includes the fricative part of the alveolar affricates in *tsik* and *dzik*.

Both German speakers show an ordering of the COG values for items like the ones in (14), where the wedge ‘<’ means ‘has a lower COG value than’; hence, /tak/ has a lower COG value than /tjak/, etc.

- (14) tak < tjak (<) tik < tsik
 dak < djak (<) dik < dzik

The wedges between *tjak* and *tik* and between *djak* and *dik* are in brackets, because a post-hoc Scheffé test shows that for both speakers the difference between /i/ and /j/ for both /t/ and /d/ is not significant. For speaker SH the difference between /da/ and /dj/ is also not significant. All other differences are significant (For JD /da/ vs. /di/ p<.01, /da/ vs. /dj/ p<.05, /da/ vs. /dz/ p<.001, /di/ vs. /dz/ p<.05, /dj/ vs. /dz/ p<.01. For SH /da/ vs. /di/ p<.01, /da/ vs. /dz/ p<.001, /di/ vs. /dz/ p<.001, /dj/ vs. /dz/ p<.001. All remaining differences for /t/ for both speakers are highly significant p<.001).

The results for the Polish speakers are given in Graph 15.

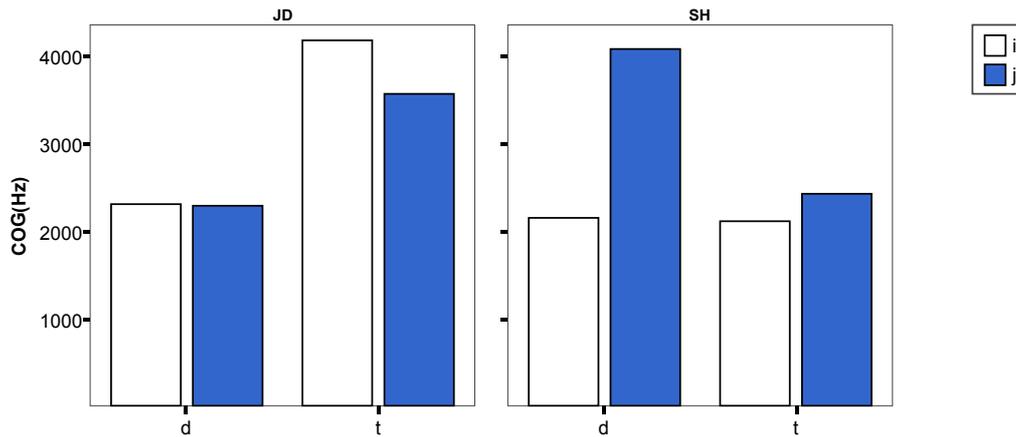


Graph 15: The average COG in Hz for the two Polish speakers. The context ‘s’ includes the fricative part of the alveolar affricates in *tsik* and *dzik*.

The ordering of the items for the Polish speakers according to their COG values is similar to the German speakers in (14), with the only exception being that speaker SL has a reverse ordering for /di/ and /dj/. The difference between /di/ and /dj/ is not significant for either speaker. For speaker SL the difference between /ti/ and /tj/ is also not significant, whereas for speaker MZ this difference is highly significant. All other differences are significant (for MZ /di/ vs. /dz/ $p < .05$, /dj/ vs. /dz/ $p < .01$, all other differences with /d/ are $p < .001$; /ta/ vs. /tj/ is significant $p < .01$, all other differences with /t/ are highly significant. For SL /da/ vs. /di/ $p < .005$, /da/ vs. /dj/ $p < .01$, /di/ vs. /dz/ $p < .01$, /dj/ vs. /dz/ $p < .01$, /da/ vs. /dz/ highly significant; for /t/ all differences apart from /ti/ vs. /tj/ are highly significant).

The COG values for all four speakers thus show that sequences of coronal stop plus /i j/ are more similar to the respective coronal fricatives /s z/ than sequences of coronal stop plus /a/. This finding points to a reason that assibilation is triggered by high front vowels but not by low vowels. However, an explanation for implication (13a) could not be found in the COG values. The reason is that none of the speakers has a significant difference in COG values between /i/ and /j/ items (apart from speaker MZ whose /ti/ items have significantly higher COG values than the /tj/ items) and the ordering between /i/ and /j/ is actually reverse from the expected.

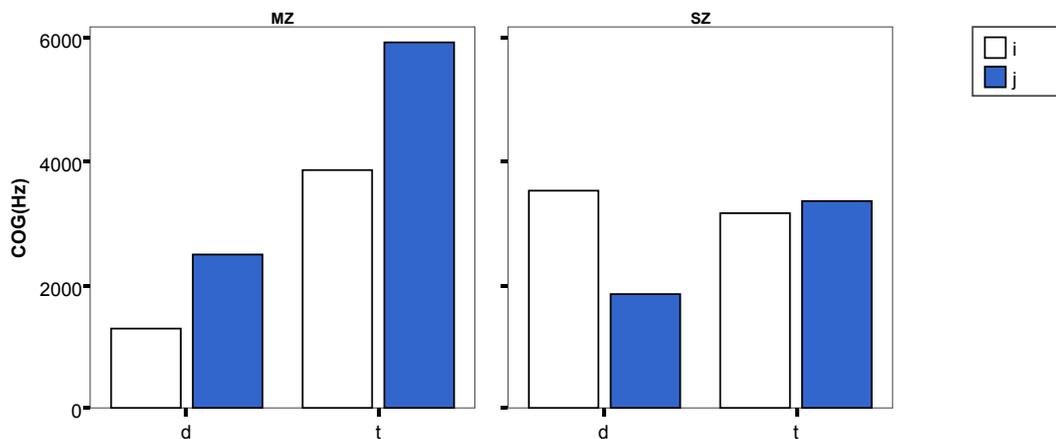
Were the results of these COG measurements to support implication (13b), we expect to find a greater difference in COG values between /di dj/ and /dz/ than between /ti tj/ and /ts/. For these calculations we measured the COG values of the fricatives including the fundamental frequency (i.e., no band pass filtering as in the COG measurements above was conducted). The results for the German speakers are given in Graph 16. COG mean values of /di/ were subtracted from COG mean values of /dz/, which is shown in the first column on the left; the second column shows the difference in the COG mean values of /dz/ and /dj/. In a parallel manner the differences for voiceless stops are shown by the columns on the right.



Graph 16: The average difference in COG in Hz for the two German speakers.

For speaker JD, the value /dz/ - /di/ (first column) is smaller than that for /ts/ - /ti/ (third column), and /dz/ - /dj/ (second column) is smaller than /ts/ - /tj/ (fourth column). This finding is important because it does not support the expected result that differences for the voiced stop are larger than the voiceless stops. A different picture arises for speaker SH, for whom /dz/ - /di/ is almost identical to /ts/ - /ti/, and /dz/ - /dj/ is greater than /ts/ - /tj/.

The differences in COG values between the affricate and /i j/ for the Polish speakers are given in Graph 17, again for the voiced stops on the left and the voiceless stops on the right.



Graph 17: The average difference in COG in Hz for the two Polish speakers.

Speaker MZ shows a larger value for both voiceless alveolar stops than for the voiced alveolar stops, and speaker SZ has a larger value for /dz/ - /di/ than for /ts/ - /ti/, but a smaller value for /dz/ - /dj/ than for /ts/ - /tj/. The results of both German and Polish speakers are so diverse that a statistical analysis is unnecessary. In sum, there is no general tendency that the difference in COG values for voiced stops is greater than the difference between voiceless stops, and therefore no evidence for implication (13b).

Summing up this section, the results of the COG measurements can account for the fact that assibilation takes place before high front vocoids and not before low vowels, but they do not provide evidence for the assibilation hierarchy in (12) or for the implications in (13).

8 Conclusion

In this article we reported on the results of three acoustical experiments in which the duration of the friction phase of /ti di tj dj/ were measured in German and Polish and showed the relevance of these experiments to language typology. The results indicate that the friction phase for these four sequences can be arranged in the order /tj/ < /ti/ < /dj/ < /di/, meaning that the friction phase in /tj/ is longer than in /ti/ etc. The results of these experiments are important because they lend phonetic support to two proposed implications with respect to phonological stop assibilation which were made in an earlier study (Hall & Hamann 2003), namely (a) assibilation cannot be triggered by /i/ unless it is also triggered by /j/, and (b) voiced stops cannot undergo assibilation unless voiceless ones do.

The additional COG measurements showed that this parameter does not provide evidence for the assibilation hierarchy /tj/ < /ti/ < /dj/ < /di/. However, further investigations into the spectral shape of the friction phase in the sequences /tj ti dj di/ and the fricative component of /dz ts/ could account for the observed ordering among alveolar stop plus high vocoid sequences.

References

- Bell-Berti, F. 1975. Control of pharyngeal cavity size for English voiced and voiceless stops. *Journal of the Acoustical Society of America* 57: 456-461.
- Bhat, D. N. S. 1978. A general study of palatalisation. In: J. Greenberg et al. (eds.), *Universals of Human Language*, 47-92. Stanford: Stanford University Press.
- Byrd, D. (1993). 54,000 stops. *UCLA Working Papers in Phonetics* 83, 97-115.
- Cedegren, H. J., J. D. Archambault & G. Boulianne 1991. Stop assibilation in Quebec French: an analysis by articulatory synthesis. *Proceedings of the 12th International Congress of Phonetic Sciences*. Vol. 3. Aix-en-Provence. 186-189.
- Chitoran, I. 2002. A perception-production study of Romanian diphthongs and glide-vowel sequences. *Journal of the International Phonetic Association* 32,2: 203-222.
- Chitoran, I. 2003. Gestural Timing and the Glide Percept in Romanian. *Proceedings of the 15th International Congress of Phonetic Sciences*: 3013-3016.
- Clements, G. N. 1999. Affricates as noncontoured stops. In: O. Fujimura et al. (eds.), *Item, Order in Language and Speech*, 271-299. Prague: Charles University Press.
- Dougherty, J. W. D. 1983. *West Futuna-Aniwa: An Introduction to a Polynesian Outlier Language*. Berkeley: University of California Press.

- Foley, J. 1973. Assibilation as universal phonological rule. *Folia Linguistica* 6 (3-4): 251-262.
- Foley, J. 1977. *Foundations of Theoretical Phonology*. Cambridge: Cambridge University Press.
- Forrest, K., Weismar, G. Milenkovic, P. & R. N. Dougall. 1988. Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America* 84, 115-124.
- Gordon, M., Barthmaier, P. and K. Sands. 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32, 141-171.
- Hall, T.A. & S. Hamann. 2003. Towards a typology of stop assibilation. In: T.A. Hall & S. Hamann (eds.) *Papers on Phonetics and Phonology. ZAS Papers in Linguistics* 32: 111-136.
- Hall, T.A., S. Hamann, & M. Zygis. In prep. The aerodynamics of stop assibilation. Ms. ZAS Berlin.
- Hanson, M.H. & K.N. Stevens. 2003. Models of aspirated stops in English. *Proceedings of the 15th ICPHS*, Barcelona. 783-786.
- Jäger, J. 1978. Speech Aerodynamics and Phonological Universals. Berkeley: *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*. Berkeley Linguistics Society. 311-329.
- Jassem, W. 1979. Classification of fricative spectra using statistical discriminant functions. In: Lindblom B. & S. Öhman (eds.) *Frontiers of Speech Communication Research*. New York: Academic Press.
- Jessen, M. & C. Ringen 2002. Laryngeal features in German. *Phonology* 19: 189-218.
- Jongman, A., Wayland, R. & S. Wong. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America* 108, 1252-1263.
- Kawasaki. 1982. An Acoustical Basis for Universal Constraints on Sound sequences. Ph.D. dissertation. University of California at Berkeley.
- Kent, R. D. & K. L. Moll. 1969. Vocal-tract characteristics of the stop cognates. *Journal of the Acoustical Society of America* 46: 1549-1555.
- Kim, H. 2001. A phonetically based account of phonological stop assibilation. *Phonology* 18, 81-108.
- Kiparsky, P. 1973. Abstractness, opacity and global rules. In: O. Fujimura (ed.), *Three Dimensions of Linguistic Theory*, 57-86. Tokio: Taikusha.
- Klatt, D. 1975. Voice onset time, frication and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research* 18: 686-706.
- Kohler, K. 1995. Einführung in die Phonetik des Deutschen. Berlin: Erich Schmidt Verlag.
- Ladefoged, P. 1964. A Phonetic Study of West African Languages: An Auditory-Instrumental Survey. Cambridge: Cambridge University Press.
- Laver, J. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.
- Lieberman, A.M., Delattre, P. C. and F. S. Cooper (1958). Some cues for the distinction between voiced and voiceless stops in initial position. *Language & Speech* 1, 153-167.
- Lisker, L & A.S. Abramson (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20: 384-422.
- Maddieson, I. & Emmorey, K. 1985. Relationship between semivowels and vowels: cross-linguistic investigations of acoustic difference and coarticulation. *Phonetica* 42: 163-174.

- Matisoff, J. A. 1982. *The Grammar of Lahu*. Berkeley: University of California Press.
- Netsell, R. 1969. Subglottal and Intraoral Air Pressures during the Intervocalic Contrast of /t/ and /d/. *Phonetica* 20: 68-73.
- Nittrouer, S., Studdert-Kennedy, M and R.S. McGowan. 1989. The emergence of phonetic segments: evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research* 32, 120-132.
- Ohala, J. 1983. The origin of sound patterns in vocal tract constraints. In: P. MacNeilage (ed.), *The Production of Speech*, 189-216. New York: Springer.
- Ohala, John J. & C. J. Riordan. 1979. Passive vocal tract enlargement during voiced stops. In: J. J. Wolf & D. H. Klatt (eds.) *Speech Communication Papers*, New York: Acoustical Society of America; 89-92.
- Perkell, J. S. 1969. *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study*. Cambridge: MIT Press.
- Pope, M. K. 1952. *From Latin to Modern French with special consideration of Anglo-Norman. Phonology and Morphology*. Manchester: Manchester University Press.
- Stevens, N. K. 1998. *Acoustic Phonetics*. Cambridge, Mass./London, England: the MIT Press.
- Stevens, K. N., Manuel, S. Y. & M. Matthies. 1999. Revisiting place of articulation measures for stop consonants: implications for models of consonant production. *Proceedings of 14th ICPHS, San Francisco*. 1117-1120.
- Wängler, H. H. 1961. *Atlas deutscher Sprachlaute*. Berlin: Akademie Verlag.
- Wierzchowska, B. 1971. *Wymowa Polska*. Warszawa: Państwowe Zakłady Wydawnictw Szkolnych.
- Wierzchowska, B. 1980. *Fonetyka i fonologia języka polskiego*. Warszawa: Zakład Narodowy im. Ossolińskich: Wydawnictwo Polskiej Akademii Nauk.
- Wiese, R. 1996. *The Phonology of German*. Oxford: Clarendon Press.

The Hungarian palatal stop: Phonological considerations and phonetic data

Christian Geng
ZAS Berlin

Christine Mooshammer
University of Kiel

This study examines the movement trajectories of the dorsal tongue movements during symmetrical /VCa/ -sequences, where /V/ was one of the Hungarian long or short vowels /i,a,u/ and C either the voiceless palatal or velar stop consonants. General aims of this study were to deliver a data-driven account for (a) the evidence of the division between dorsality and coronality and (b) for the potential role coarticulatory factors could play for the relative frequency of velar palatalization processes in genetically unrelated languages. Results suggest a clear-cut demarcation between the behaviour of purely dorsal velars and the coronal palatals. Moreover, factors arising from a general movement economy might contribute to the palatalization processes mentioned.

1 Introduction

The palatal stop has been a matter of debate for at least two reasons: first it is not clear whether it is articulated with a dorsal component, which would involve the specification as a complex segment. The concurrent specification would favour a classification as simple coronals and introduce at least one additional feature to separate them from the other [-anterior] coronals (Keating, 1991). Secondly, on a phonetic level, the relationship of the palatal stop to the velars remains unclear (Recasens, 1990).

The data analyzed in the phonological discourse related to the notion of distinctive features has mainly utilized static X-Ray-images, linguo- and palatograms to determine a featural description of palatal articulation. In contrast, the research work which has studied palatal articulation from the phonetically motivated viewpoint of lingual coarticulation has mainly relied on electropalatographic data, which has the merit of supplying time-varying data but only if there is a contact

and with the drawback of a very limited spatial interpretability. The first aim of this study is to give a descriptive account of the articulatory realization of palatal stops by making use of movement data acquired by the articulograph.

In recent years, the influence of phonetic mechanisms in shaping the paradigmatic contrasts of a language has been refocused by the scientific community, as articulatory economy, perceptual discriminability and the maintenance of contrast have been agreed on as the key phonetic factors shaping the sound inventories of the world's spoken languages. For the palatal articulation we are dealing with this involves the reference to the natural phonological process of (velar) palatalization, and therefore the comparison of the palatal stop articulation with the stop articulation at the velar place of articulation. Therefore, in the current study we relate the palatal to the velar stop articulation of the same language to explore the articulatory economy principles governing the palatal/velar or coronal/dorsal contrast

The introduction is structured as follows: we will first summarize phonetic approaches for explaining sound change in general, taking palatalization processes as a concrete example.¹ We will see that these refer to different parts of the phonetic band and partially make reference to a clear-cut-distinction between true assimilative change and coarticulatory variation. It will turn out that more recent approaches assign a more prominent role to coarticulatory variation than did the generative tradition. We will then summarize selected approaches to coarticulation and the different predictions they make about the concrete data we collected.

1.1 Accounts of phonetically induced sound change

Several accounts of phonetic sound change have been proposed in the literature. The main distinction between them lies in the identification of the factors by which they are driven: Is the primary principle rooted in articulation, acoustics/perception or a kind of transaction between these two kinds of influences. Let us exemplify these different standpoints on the basis of a subprocess of an extensively studied phonological process, palatalization. According to Bhat (1978), at least three distinct processes contribute to palatalization processes, and Bhat presents these as independent subprocesses of palatalization: These are tongue-fronting, tongue-raising and spirantization. Of special concern here is the the fronting of velars conditioned by the presence of a front vowel, i.e.

¹Note that the term palatalization has a different meaning in phonetics/phonology: it denotes the addition of a secondary palatal articulation to a primary articulation, like in palatalized labials. Here, the term palatalization is used in the way it is used in typology or historical linguistics, i.e. as a phonological process (Bhat, 1978).

[k] → [c] / _ [i]

As mentioned, accounts in the neo-grammarians tradition would seek for articulatory factors underlying the fronting of the velar in front vowel context. A contemporary theorist following this rationale is Recasens. He views palatalization as a gradual sound change mechanism proceeding from the unstable medio-postpalatal or medio-palatal articulations towards preferred alveolopalatal ones through an *increase in predorsal and laminal contact* (Recasens, 2003).

An approach relying primarily on acoustic properties of the speech signal is the theory and research program of acoustic invariance (Blumstein & Stevens, 1979), which was elaborated in a series of papers (Lahiri & Blumstein, 1984; Blumstein, 1986; Keating & Lahiri, 1993). This theory makes the claims that (1) there is acoustic invariance in the speech signal corresponding to the phonetic features of a language (Blumstein, 1986, p. 178) and that (2) the perceptual system is sensitive to these invariant properties. These invariant properties are seen responsible for the *natural processes in phonology*, and, in particular can account for why certain “assimilation rules” are more likely to occur. The conditions for the occurrence of an assimilation (Blumstein, 1986, p. 186) are (a) that the two contiguous segments must have some similar acoustic properties and (b) that the original sound and the modified sound must also share a number of acoustic properties. In the case of velar fronting the acoustic property of the segments involved are identified as the relative distribution of energy: Palatals exhibit a selective increase in energy between burst and voicing onset in the frequency band between 800 and 1800 Hz, whereas the distribution for velars is comparatively “flat”.

In other words, “the assimilation of [k] to [c] involves a true assimilation of the acoustic property of gravity from the vowel to the preceding consonant”² (Blumstein, 1986, p. 186). In short, this kind of theorizing views the sound change under consideration as a true assimilative change by making reference to revived Jakobsonian featural descriptions.

The third class of theories emphasizes the use of perceptual factors for sound change. Among these are -on a metatheoretical level- the scenarios of generalizing mini sound changes as elaborated by Ohala (e.g. Ohala (1983), Ohala (1993)) and Lindblom’s approach of adaptive dispersion (Lindblom et al., 1995). A concrete instantiation of Lindblom’s approach is carried out in the work of Guion (1998): She interprets the facts about velar palatalization in terms of a perceptual reanalysis of fast speech. On a large scale, we see the explananda in the potential influence and interaction of these factors in shaping language change, and for the present purpose, in an exploration of its articulatory antecedentia.

²/i/ has F2 und F3 in the same region.

1.2 Theoretical Approaches to coarticulation

Classical generative theory makes a clearcut separation between coarticulation and other context-dependent phenomena, such as assimilations: Coarticulation deals with “transitions between vowels and adjacent consonants, the adjustments in the vocal tract shape made in anticipation of a subsequent motion” (SPE:295, after Farnetani & Recasens (1999)) In contrast, assimilations involve operations on phonological features, and are accounted for by phonological rules, which map lexical representations onto phonetic representations. This kind of reasoning comprises the acoustic invariance approach relying on Jakobsonian ideas as well as the approach taken in SPE with extensive marking conventions. We skip a more thorough discussion as these approaches are not laid out to make more detailed predictions about articulatory surface behavior³: Coarticulation and other properties of phonetic implementation are assumed to follow from universal principles of speech physiology.

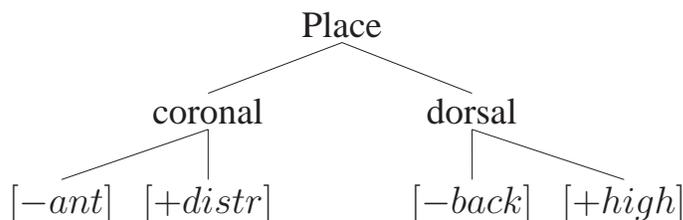
1.2.1 *The window model of coarticulation*

The “window model of coarticulation” elaborated by Keating (1990) blurs this clear-cut distinction between the grammar and the physics of speech⁴. In Keating’s view, the grammar has a phonological *and* a phonetic component. This phonetic component of the grammar has the function to specify whether a feature is relevant in the paradigm in a given language. For example, the voicing distinction is very systematic in English or German, but completely absent in Polish and Czech. This kind of facts has to be accounted for in the grammar. Altogether, there are three different ways to deal with underspecification: There is underspecification on the phonological level, which may persist in the phonetic domain. Then, there is phonetic underspecification which is conceptualized as a continuous notion. Additionally, unspecified features may be left unspecified or specified by rule. Now, if phonological assimilation rules assign a contextual feature to a segment, its associated window will be narrow before that context and the contour will have a plateau-like shape. If assimilation rules are not active, the key feature remains unspecified and the trajectories will be provided by “interpolation”. Furthermore, inter-language differences in coarticulation can be of phonological or phonetic origin. If phonological assimilation rules operate in one language and not in the other, they are of phonological origin. They are of phonetic origin, if different languages interpret an unspecified feature differently.

³but see Lahiri & Blumstein (1984).

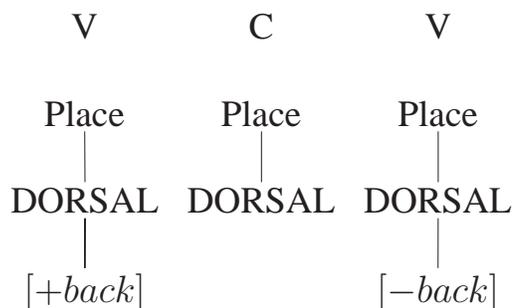
⁴It has to be mentioned though that this clear-cut distinction between the grammar and the physics of speech has been abandoned earlier by approaches stemming from within featural phonology, e.g. in approaches as the theory of feature spreading (Hammarberg, 1976), the look-ahead model, (Daniloff & Hammarberg, 1973), the model of coarticulatory resistance Bladon & Al-Bamerni (1976).

In a different series of papers, Keating (1988,1991,1993) elaborated the status of palatals and velars. In Keating (1988) it was proposed that palatals are complex segments involving both coronal and tongue body articulations, such that they would have a status like double-articulated labial-velars. The representation of the palatal has specifications on the dorsal as well as on the coronal node⁵:



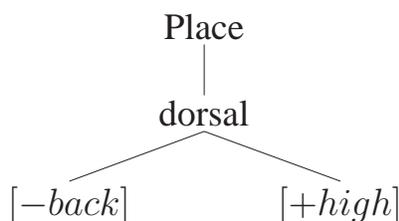
In yet another publication, Keating (1993) discusses the phenomena of velar fronting in terms of surface features for English:

With reference to the work of Houde (1968), she states that velars “lack inherent specification for Back ((Keating, 1991, p. 17))”:



“Stated another way, velar fronting is something that happens gradually over the course of the velar. Such temporal/spatial variation, or phonetic gradience, can be interpreted as a transparency effect on the velar with respect to backness.” (Keating, 1991, p. 17) This seems to tacitly assume a large “sliding” movement of the tongue during the closure interval, and, on the reverse, an absence of this kind of movements during the realization of the palatals. A point, which we will turn back on later.

For completeness, we reproduce the specification of the palatalized velars. In contrast to the velar specification cited above, palatalized velars have an inherent specification for Back.



⁵Keating (1991) offers another possibility for the representation of palatals: “However, another option in the representation of palatals is to treat them as simple coronals, and introduce at least one additional feature to distinguish them from the [-anterior] coronals. This is in fact what Halle (1968) does with his new features Lower Incisors Contact. Actually, both options could be exercised for more descriptive coverage (Keating, 1991, p. 45).”

1.2.2 Coproduction theory and derivatives

At this point, the necessity arises to review the very basic results of one of the most influential studies and the accompanying model of coarticulation: Öhman (1966) and Öhman (1967) proposed a model of vowel-to-vowel coarticulation, the basic empirical evidence of which was articulatory and acoustic analysis of Swedish VCV utterances produced in isolation, and similar speech material in American English and Russian. The material on the Russian data was different with regard to (secondary) palatalization. The major finding was that the consonantal transitions (V_1C and CV_2) depend on the identity of the transconsonantal vowel. But: this coarticulatory variability was reduced to almost random fluctuation in the case of Russian. Öhman interprets these findings as follows: The tongue is considered a system of independently operating articulators driven by invariant articulatory commands. The apical articulator is involved in the formation of apical consonants, the dorsal articulator in the formation of palatal and velar consonants and the tongue body articulator in the formation of vowels. The reduced coarticulatory variability for the palatalized F2-transitions is seen as the result of conflicting vowel commands on the tongue body, i.e. an [i]-like palatalization commands exerting a blocking effect on the following vowel.

On the basis of Öhman's work, the *coproduction* theory and *articulatory phonology* have been elaborated.

Fowler (1980) argues against speech production theories in general which take phonological features as input. The features used as input for the speech production mechanism are timeless, abstract and static and have to be translated into articulatory movement. As Farnetani & Recasens (1999, p. 51) put it: "In this translation process, the speech plan supplies the spatial target and a central clock specifies when the articulators have to move." In contrast, Fowler's intention is to overcome this dichotomy and she suggests to modify the phonological units of the plan: The phonological units become dynamically specified phonetic gestures, with an *intrinsic temporal dimension*. In speech, these gestures are implemented by *coordinative structures*, i.e. by temporary functional dependencies among the articulators contributing to the goal the gestures want to achieve. For example, in producing a bilabial stop, a temporal functional link is created between upper lip, lower lip and jaw. Several gestures are allowed to be coproduced. The amount of articulatory variability induced by this coproduction depends on the degree to which the gestures involved share articulators. The case of minimal gestural interference is the production of /VbV/-sequences, where vocalic and consonantal constriction gestures involve two independent sets of articulators. Conversely, a "gesture is defined along exactly the same set of tract variables and articulators as the flanking vowels", if the consonant is a velar (Saltzman & Munhall, 1989). However, the original work seems to make no reference to palatal articulation, but a series of papers by Recasens (1997, 2002) was explicitly designed to make these

predictions⁶: The goal of the DAC(Degree of coarticulatory constraint)-scale attempts to characterize phonetic segments according to the types of articulatory constraints involved in their production. These values then can be used to predict the “coarticulatory resistance” of the segments. As in Fowler’s theory, the DAC-model assumes that articulatory gestures associated with consecutive segments are coproduced and overlap to different degrees depending on their spatiotemporal properties, on prosodic factors and speech rate. According to DAC-scale, “consonants differ in DAC value according to the following order: dorsals (alveolopalatals, palatals, velars), lingual fricatives (/s/,/ʃ/), dark /l/, which can be assigned a maximum DAC value (DAC=3); dentals and alveolars such as /n/ and clear /l/ (DAC=2); and bilabials, with the lowest DAC value (DAC=1)(...) It is hypothesized that dorsal consonants are highly constrained based on the observation that their primary contact or constriction location stays relatively fixed in line with the large contact size involved and perhaps the sluggishness of the tongue dorsum. The same observation may even apply to velars provided that at least two targets in front and back vowel contexts are accounted for (Recasens, 2002)”.

1.3 Hypotheses

To sum up the selected theories presented and predictions concerning palatals and velars:

1. Keating: Palatals -if the description as complex segments is not outdated- are dorsals and coronals likewise. This amounts to relatively little sensitivity to vowel-induced contextual coarticulation for palatals in contrast to velars. At the same time, front velars should be distinct from palatals in their shape configurations, velars lacking the /i/-like “component”. If the surface underspecification Keating adopts for English is also valid for Hungarian velars, then the velars exhibit relatively large vowel-dependence in place of articulation. Furthermore, let us shortly review her statements on the phonetic implementation on contextual velar fronting: “Stated another way, velar fronting is something that happens *gradually over the course of the velar*”. As mentioned, Keating seems to view this as a “*transparency effect of the velar with respect to Backness.*” (Keating, 1993, p. 17) . We need to add here, that this implies that this transparency effect should be absent *over the the course of the palatal*. Putting this together, according to Keating, velars and palatals should be distinguished by the amount of coarticulatorily induced variability and their behavior *during* these stops.
2. In contrast, Recasens (1990) rejects Keating’s (1988) claim that palatals are complex segments produced with the blade and the tongue dorsum: “Contrary

⁶Recasens explicitly mentions his model as being a concrete instantiation of coproduction theory.

to Keating, we believe that those consonants which have been characterized as alveopalatals and front palatals in this paper belong to different articulatory classes. Moreover, front velars are back palatal articulations, and a mid palatal class is also needed in the light of the articulatory facts. (Recasens, 1990, p. 276)”.

This implies the following predictions: Contextual fronting of velars should amount in a convergence with the tongue shapes of palatals. Furthermore, according to the DAC-scale presented above, palatals and velars should be equally prone to effects induced by variation of the vowel context.

The aim of this study is to tease apart these contradictory predictions by a comparison of the palatal and velar stops in Hungarian with respect to (a) vowel-context-specific effects on tongue posture analyzed by factor-analytic techniques and (b) the vowel-context-specific movement patterns during the realization of these consonants.

2 Method

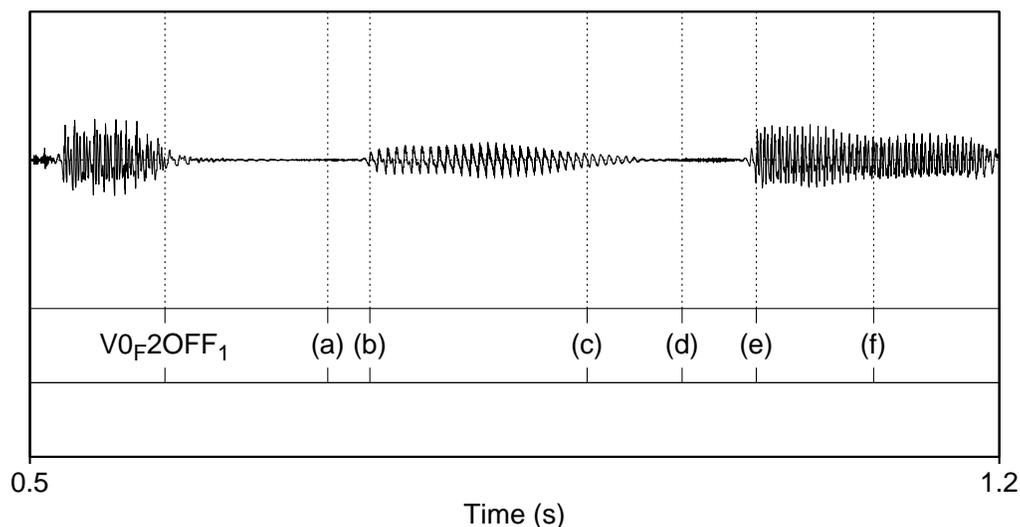


Figure 1. The figure shows an Illustration of segmentation criteria. For a more detailed description see text. The example sound is a voiceless palatal in the context of /i:/.

Tongue, jaw and lower lip movements of one female and two male speakers of Hungarian were recorded by means of Electromagnetic Midsagittal Articulography (EMMA, AG100, Carstens). The two male speakers natively were from Budapest, the female speaker came from the North of Hungary, from the region of Szeged. The choice of the Hungarian language arose from articulatory reasons, as the Hungarian palatal stops have been described as the palatal stops with the po-

tentially strongest dorsal colouring (Keating & Lahiri, 1993)⁷. Four sensors were attached to the tongue, one as far back as possible (TB), one approximately 1 cm behind the tongue tip (TT). The two middle sensors, tongue dorsum (TD) and tongue mid (TM) were located at equidistant points between them. Additional sensors were glued on the vermillion border of the lower lip (LLIP) and on the lower incisors (JAW). Two sensors on the nasion and on the upper incisors served as references for compensation of head movements relative to the helmet and definition of an intermediate coordinate system. The final coordinate system was defined by recordings of two sensors on a T-bar acquired in order to rotate the data to the occlusion plane for each speaker individually (Hoole, 1996). Original sample frequencies were 400 Hz for EMMA data and 48 kHz for the acoustical signal. For the analysis, the EMMA signals were low-pass filtered and downsampled to 200 Hz. The material consisted of /CVCa/ nonsense words with either velar or palatal voiceless stops as consonantal context and one of the long or short corner vowels /i,a,u/. Two realizations of the test words were embedded in the carrier sentence “Most a ... es a ... volt” (“This was a ... and a ... now”) and repeated six to 10 times. The following temporal landmarks were extracted manually from the acoustic signal by means of the software package PRAAT by Boersma & Weenink (1992–2004): (a) the burst of the initial consonant (b) the onset of the second formant of the first vowel, (c) the offset of the second formant of the first vowel, (d) the burst of the the medial vowel. (e) the onset of the second formant of the second vowel (f) the offset of the second formant of the second vowel. For the analyzes of the distances reported below -we will analyze the medial consonants of the /CVCa/-sequences- four different intervals of the VCa-sequence were defined: the movement from the midpoint of the first vowel to the onset of closure, where the midpoint of the vowel is defined as the central sample of the temporal landmarks of (b) and (c), the interval during closure defined as the time between (c) and (d), the interval between the burst and voicing onset as the time between (d) and (e) and, the interval during the /a/, where the vowel mid of the second vowel is defined in analogy to the interval of the first vowel. These intervals serve as the basis for the *articulatory* analyzes reported below. In our opinion, the use of acoustic landmarks seems satisfied, because the commonly applied definition of articulatory landmarks from the speed signal was not consistently possible in all vowel environments. The distances traveled during the intervals defined above are calculated by summing the distances traveled by selected sensors over the course of these trajectories.

⁷As far as we can see, diachronically it has not emerged from velar palatalization: Palatals are a rather rare outcome of velar fronting (Bhat, 1978).

3 Results

3.1 Preliminary qualitative evaluation on palatal stop realisations in Hungarian

There was been a long-standing debate whether the palatal stop in Hungarian is a stop or an affricate (see Siptár & Miklós (2000) for a summary). Realisations of the three speakers presented here showed yet another pattern: During the stop interval no full silence was achieved but the whole interval was accompanied by frication. Additionally we found a portion which we interpreted as a (residual) burst. This was followed by a second frication portion with a change in the spectral energy distribution. Two more speakers have been recorded, one with and without EMMA. Preliminary inspection gave evidence of a clear palatal stop realisation. Since for the former speakers frication during the closure phase only occurred for the palatal but not for the velar stop we assume that it is not an artefact of the recording procedure but a speaker-dependent allophonic variation. More thorough and detailed spectral analyses are needed before a conclusive categorisation of the observed patterns are possible.

3.2 Positional data

In the two top panels and the bottom left panel of fig.2, results of Principal Component analyses of the covariance matrices of the averaged articulatory configurations excluding the lips during (a) the initial burst, (b) the medial burst and (c) the vowel configurations at the midpoint of V1 are shown. Principal component analysis (PCA) involves a decomposition of a larger number of usually correlated variables into a (usually smaller) number of not directly observable uncorrelated variables. These are the so-called principal components. One drawback of these single-speaker factor solutions is the fact of rotational indeterminacy, i.e. there exists an infinite number of solutions which explain the same amount of variance. In the bottom right panel of fig. 2, the solution of the multispeaker factors analytic model PARAFAC which avoids this problem is shown⁸.

The aim of both methods is to reveal new meaningful underlying variables, in our case, we can reduce the articulatory configuration containing separate x- and y- positions for four tongue and one jaw sensor into a two-dimensional representation still representing the gross topology of the articulatory space analyzed. The first principal component accounts for as much of the variability in the data as

⁸PARAFAC is a type of multi-mode analysis procedure and therefore contrasting (PCA) or factor analysis, which are two mode representations. PARAFAC requires an at least three-dimensional data structure with the third dimension usually being represented by different speakers, i.e. if all speaker weights are fixed to be one, then PARAFAC reduces to PCA. The advantage of PARAFAC is that there is no rotational indeterminacy as in PCA, in other words, PARAFAC gives unique results. The PARAFAC-model was introduced in experimental phonetics by Harshman et al. (1977).

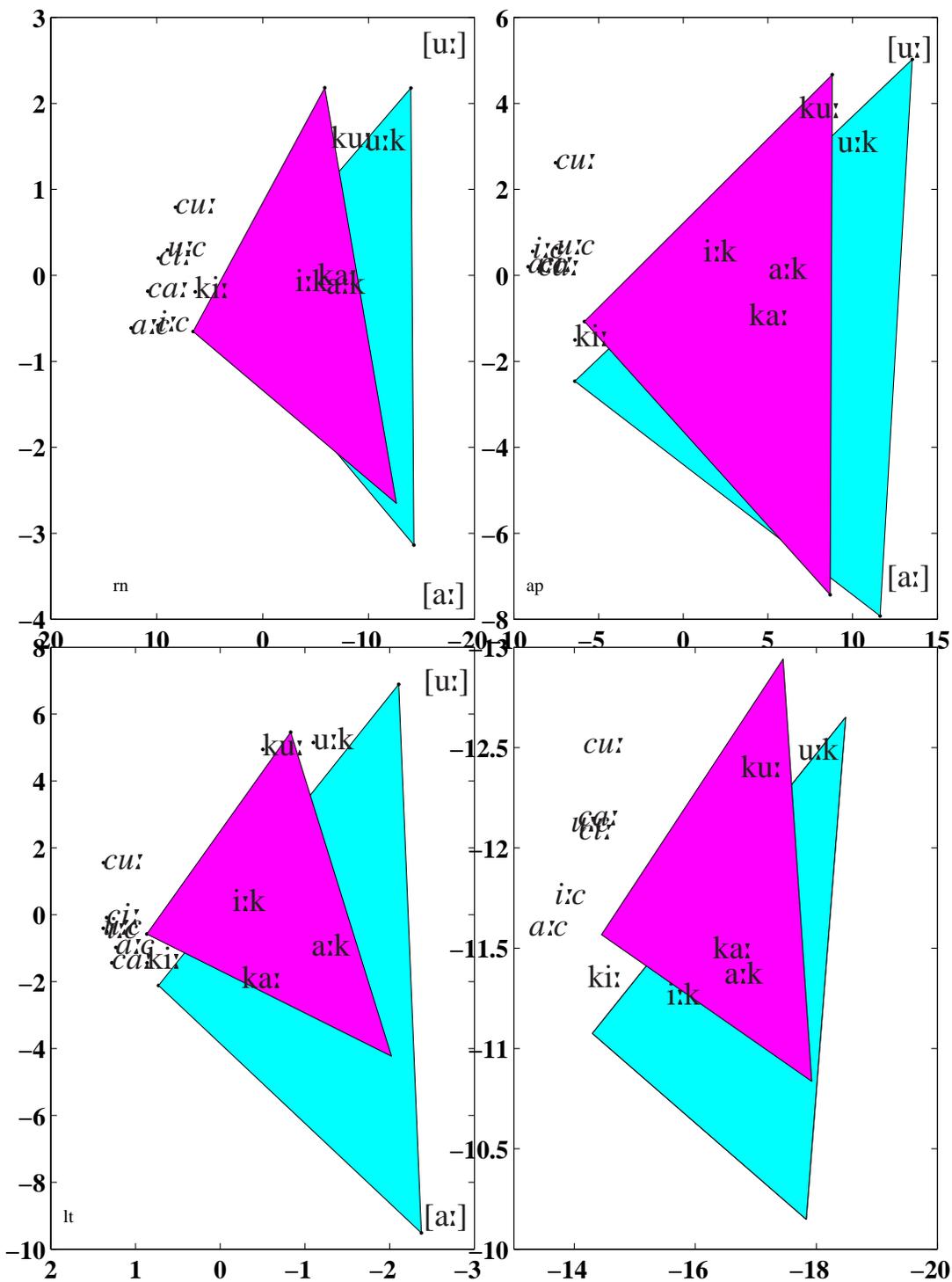


Figure 2. The two top panels and the lower left panel show speaker-dependent Principal Component analyzes of the mean tongue configurations at different temporal landmarks: the initial release, the medial release. The light gray triangles are the midpoints of the long corner vowels in the context of velar consonants, the dark gray triangles the corresponding projections of the long corner vowels in the palatal contexts. The lower right figure is a speaker-independent PARAFAC projection of the same data

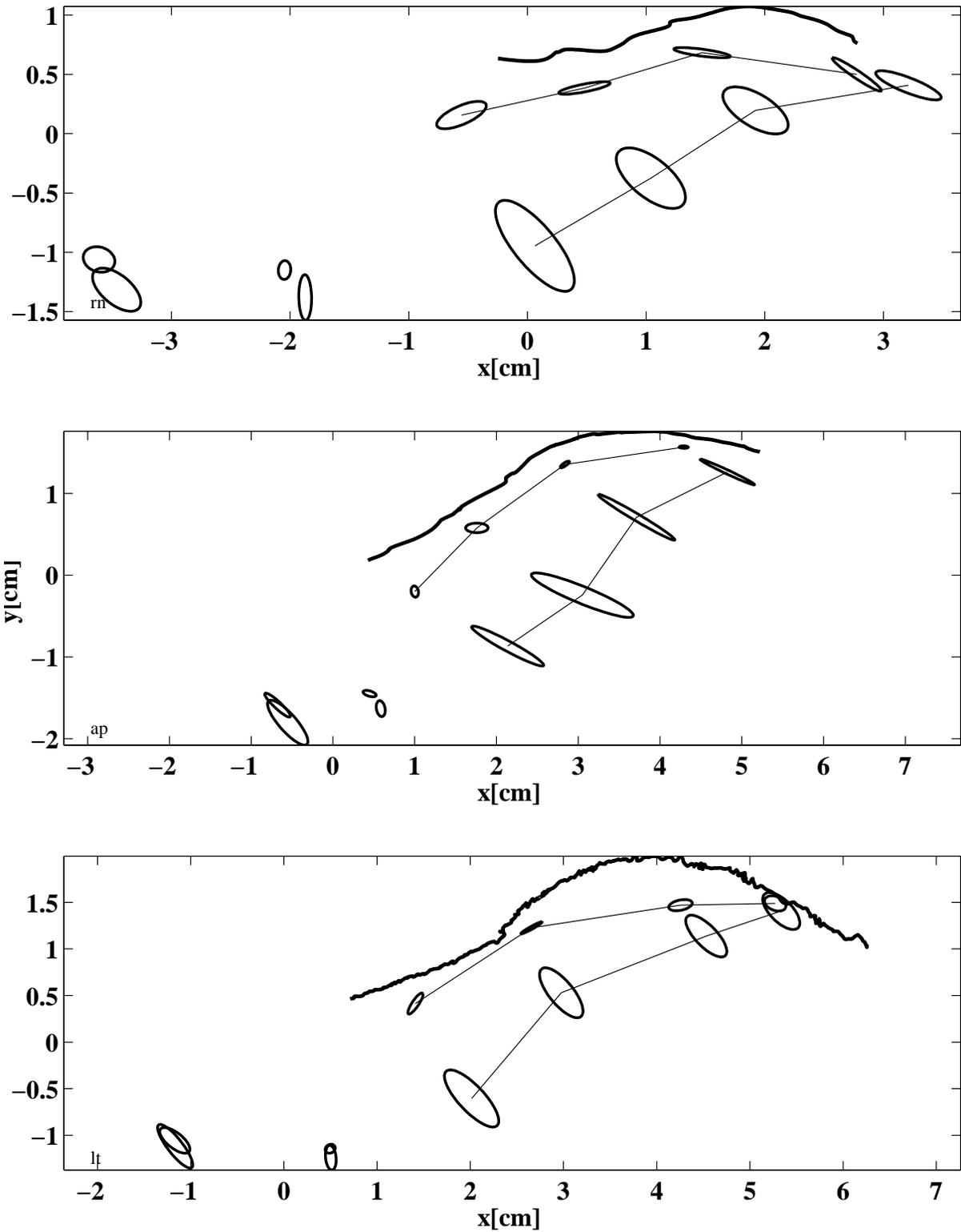


Figure 3. 1- σ -ellipses as indication for the variation of palatal and velar consonants as calculated by the means over all vowel contexts at burst time for initial and medial consonants. Small ellipses: Palatal configuration; large ellipses: velar configuration

possible, and so does each succeeding component. For the single speaker Principal Component Analyses, the first factors explain between 65 and 80% of the variance and the second factors between 17 and 27 %. The total amount of variance varies between 93 and 98%. This means that the most substantial portion of the variance is accounted for in the PCA representations, but the amount of variance explained is partitioned in different ways by different speakers. The grey “triangles” shown in fig. 2 are the vowel configurations at the midpoint of V1, with the dark grey triangle for the palatal context and the light grey triangle for the palatal context. The most salient pattern of all these plots seems that the velars show much more inclination to coarticulate with their vocalic environment in comparison with the palatals: For speakers ap and lt (top right and bottom left panel) all the palatals except the initial palatal in the context of /u:/ cluster close together, not far from the /i:/- edge of the “palatal triangle”. Compared to the place variation for the velar stop this gives the impression of a stable palatal configuration very close to the first factor score of /i/. This finding is further substantiated by the plot of $1 - \sigma$ -ellipse shown in figure 3. The velars substantially show more variation for all three speakers.⁹ In particular, only the second factor on the y-axis distinguishes between the palatals. The first factor is almost constant for the palatals and in line with the classic interpretation as “front-raising” to an /i/-like shape (Harshman et al., 1977). The most consistent pattern for the velars is that the shape of the velar in /u:/-context is close to the /u:/-corner. Of particular interest are the projections of the /i:/ in velar context and initial position. And here, the patterns do not seem to be fully conclusive: For ap, the /k/ is separated from the main palatal cluster as well as from the initial /c/ in /u:/-context. For speaker lt, the initial /k/ in /i:/-context is “dissolved” in the main palatal cluster. The same holds for speaker /rn/ although the whole single-speaker solution seems more diffuse than for the other speakers. So if we would have to be decisive about these factor solution, we would prefer to interpret the PARAFAC solution in the bottom right panel, because to be identified, certain assumptions have to be met: “The basic assumption is that, if a factor corresponds to some real organic unity, then from one study to another it will retain its pattern, simultaneously raising or lowering all its loadings according to the magnitude of the role of that factor under the different experimental conditions of the second study.” (Cattell & Cattell (1955), after Harshman & Lundy, 1984, p.151). This means that this solution has to be judged higher in terms of its validity. This is substantiated in a clearer interpretability: The medial palatal consonants in /i:/- and /a:/-context show the strongest coarticulation with the vowel /i:/ in the palatal context, a finding that could have been expected. Then, there is a second palatal cluster with initial /i/ and /a/ and medial /u/-context. This analysis substantiates and enhances the finding of the initial palatal /u/ as having a

⁹One could raise the objection that the place of maximal constriction is not reliably measured even by the EMA tongue back sensors, but the parallel orientation of the ellipses for the rearmost sensor indicates that the constriction is mostly caught.

bias towards the vowel /u/. And finally, the initial velar in /i/-context is the closest to the palatals of all velars. Moreover, another finding concerning the vowel triangles is remarkable: The sizes of the triangles are reduced in the palatal context, indicating, that the palatal consonants themselves exert a stronger coarticulatory influence on the vowels than the velars: in comparison with the velars, this results in a fronting of /u/ and a rising of the /a/.

3.3 Kinematic Characteristics

Figure 4 shows distances traveled during the four intervals which were defined earlier for the tongue dorsum sensor. Similar results were obtained for the tongue back sensor. The most salient aspect of these plots is that the total distances traveled by the TD sensor during the /VCa/-sequences are larger in the palatal contexts for /a/ and /u/-contexts; the reverse holds for the /i/-contexts.

Concerning the velar contexts, we found some relatively surprising patterns in comparison with the data we analyzed earlier. In particular, in Geng et al. (2003) we found consistently bigger total amplitudes in the /a:/-contexts. We interpreted this finding in agreement with Munhall et al. (1991), who observed a reduction in movement complexity after removing of the jaw influence. This tendency is weaker in this corpus and even reversed for speaker ap. So if the patterns for the /a/-contexts can at least partly be explained through a contribution of the jaw for *both* consonants, this explanation is not justified for the large total amplitudes for the palatals observed in the context of /u:/. If we cannot attribute these large movements to an influence of the jaw, then this pattern must be attributed to a strong movement component by the tongue itself. We will return to this point later. Another quite general observation in these plots is that the movement amplitudes *during* the stop, -i.e. the black parts of the bars- are usually larger for the velar consonants. This holds with the exception for the /a:/ and /u:/-contexts, where this finding is blurred, most probably due to large amplitudes aforementioned. Similar observations can be made for the distances traveled between the stop release and the onset of the second vowel /a/.

As a crude method for quantifying the *direction* the tongue paths travel during the closure interval, we weighted the distances the sensors traced in the closure interval by a “direction coefficient”, which was determined as the sign function of the difference between the x-coordinates of the first sample of the closure interval and the last sample of the closure interval, i.e., negative values indicate the tendency to make a movement in backward direction during the closure. The error bars in figure 5 indicate the standard deviations of this composite for the tongue dorsum sensor. Again, similar results were obtained for the tongue back sensor. Note that this is a very gross measure, in particular, a mainly vertical movement during closure would have the consequence of making the sign function, which is only based on the horizontal movement in this interval relatively arbitrary. So note

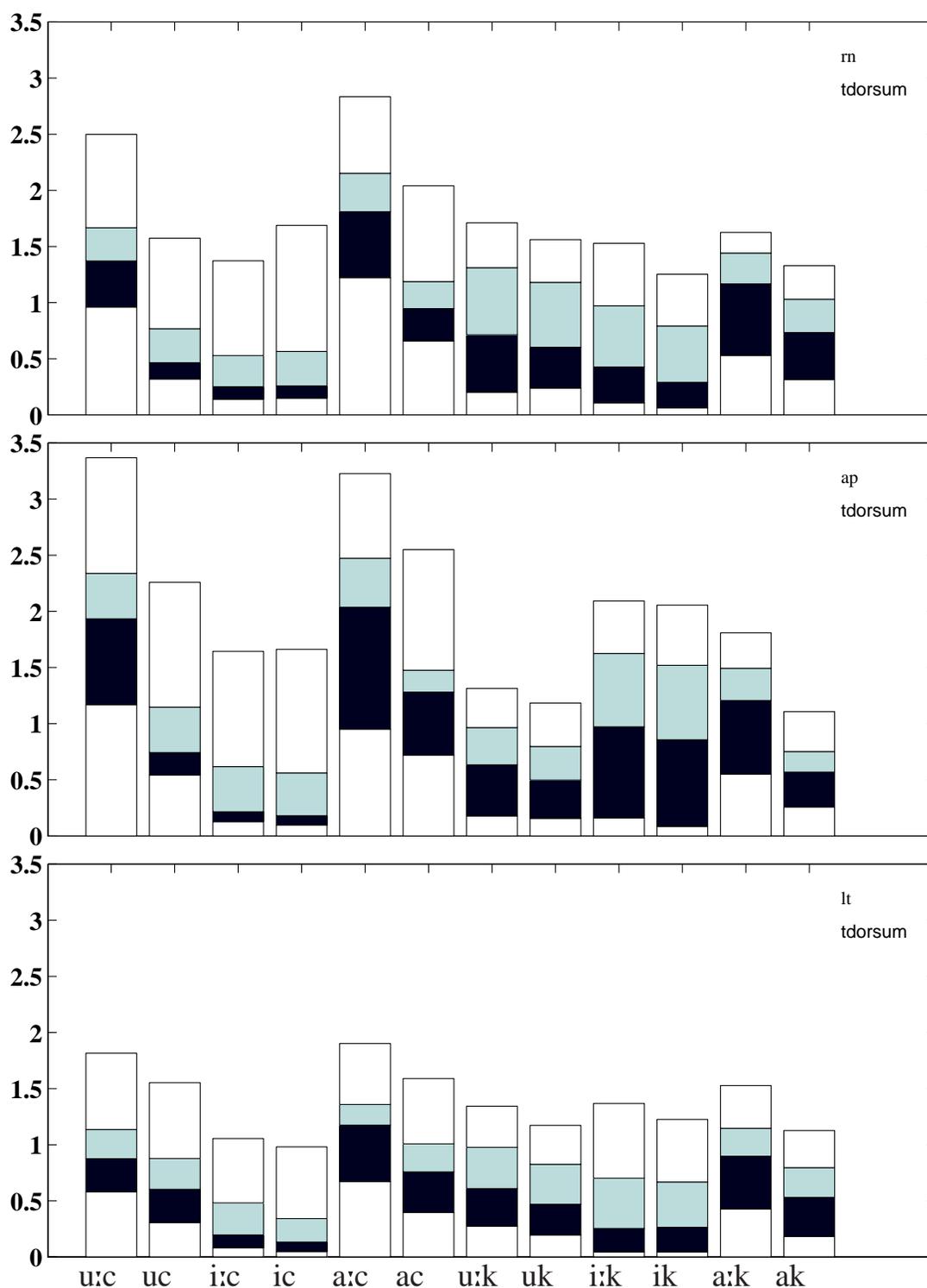


Figure 4. Distances the tongue dorsum sensor traveled during the four different intervals described. The stack bars indicate from bottom to top: first stack, white, distance traveled during the first vowel; second stack, black, distance traveled during oral closure; third stack, gray, distance traveled between stop release and the onset of the second vowel; fourth stack, white, distance traveled during the second vowel. Upper case characters: lax vowels, lower case: tense vowels.

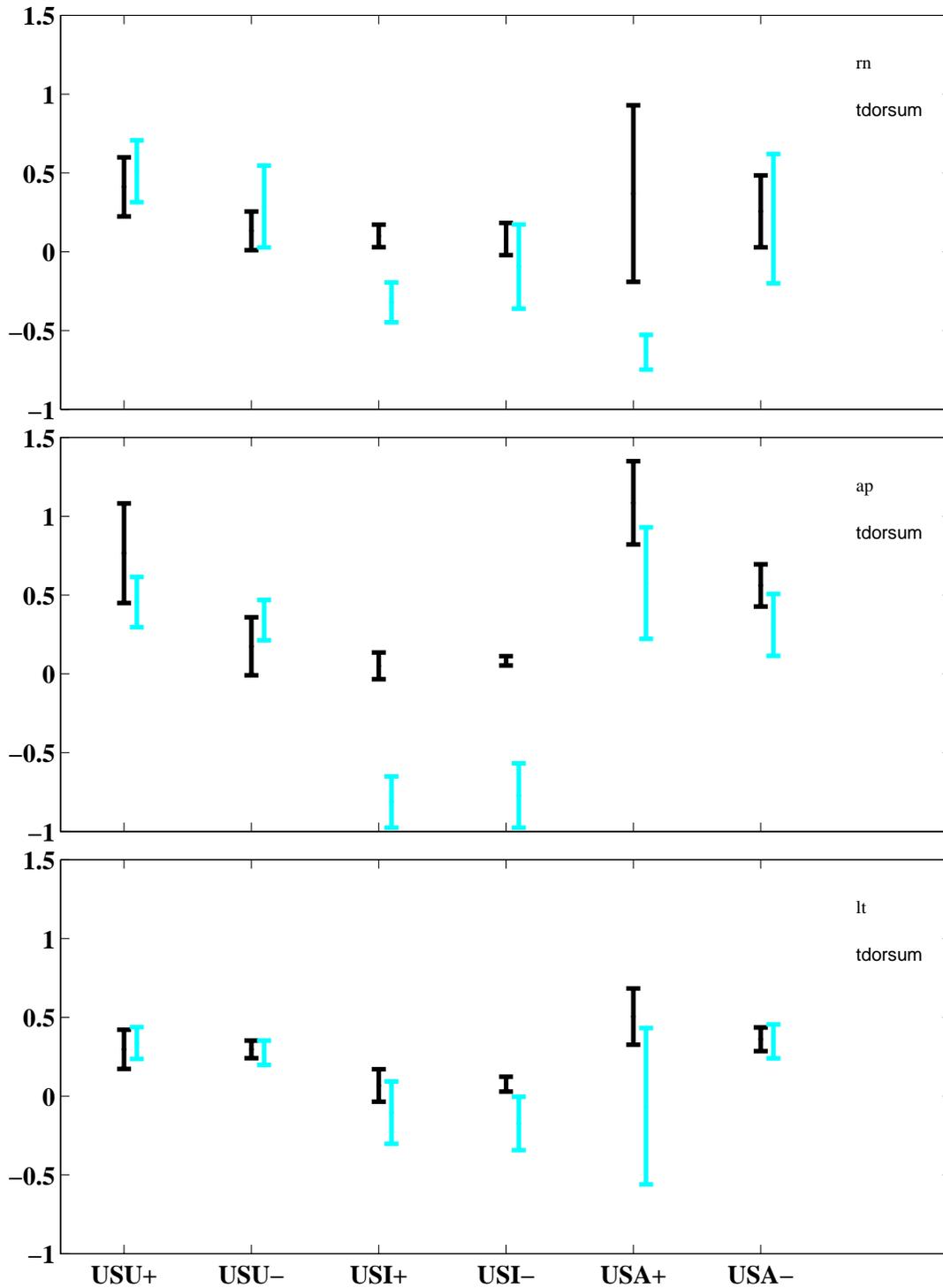


Figure 5. Distances the tongue dorsum sensor traveled during the the closure interval weighted by a direction coefficient. The direction coefficient was determined as the sign function of the difference between the first sample of the closure interval and the last sample of the closure interval, i.e., a negative value indicates the tendency to make a movement in backward direction during the closure.

that the partially large standard deviations in these plots underlyingly might represent (a) heterogeneity in the true movement as an expression of token-to-token-variability or (b) the predominance of a vertical component causing a “noisy” sign function. So we will limit ourselves to the results which do not have this uncertainty and exclude the patterns which have large standard deviations and values above and below zero. For example, this concerns the results for speaker It in the context of /a:/: As due to the influence of the jaw, a vertical movement dominates, which results in a noisy sign function and therefore in large standard deviations.

Following the front vowel /i/ (long and short) and short /u/, very little movement was found during the palatal closure. A higher degree of forward movement occurred during the palatal following long /u/ and /a/. For the velar stop movement during closure was consistently largest following long and short /u/. After /i/ two speakers showed almost no movement during /k/ which is consistent with the data on German (Geng et al. 2003) and a relatively large backwards movement for speaker AP. Movement directions following long /a/ varied inter- and intraindividually (e.g. large standard deviations for speaker LT) These results indicate that the palatal stop is produced close to the constriction location for /i/, therefore no movement is required during closure.

Figure 6 shows the correlations between the positions in the mid of the first vowel and the distances traveled during the stops for tongue dorsum sensor. Again, similar tendencies were observed for the tongue back sensor. The left panel shows the correlation of distance and x-position and the right panel of distance and y-position. For two out of our three speakers, the correlations for velar and palatal articulations have the same sign, i.e. the more anterior the sensor location during the following vowel, the larger the distance traveled during the stop closure (left panels), and, the higher the position of the sensor during the preceding vowel, the smaller the distance traveled during the preceding stop. Note that for RN and LT, the correlations between horizontal position and the distance traveled are spurious, only the height of the sensor during the preceding vowel seems to be related to the distance traveled during the stop. For one speaker, ap, the sign of the correlations is reversed for both sensors in the *velar* condition. This reverse in sign for the velar seems to contradict the results for German we presented in Geng et al. (2003). For the three speakers we presented there, no such sign reversal was observed¹⁰, rather similar patterns to those of the other two speakers of this study were found.

4 Discussion

The results of this experiment can be summarized as follows:

¹⁰note that the vowel environment was richer in this study.

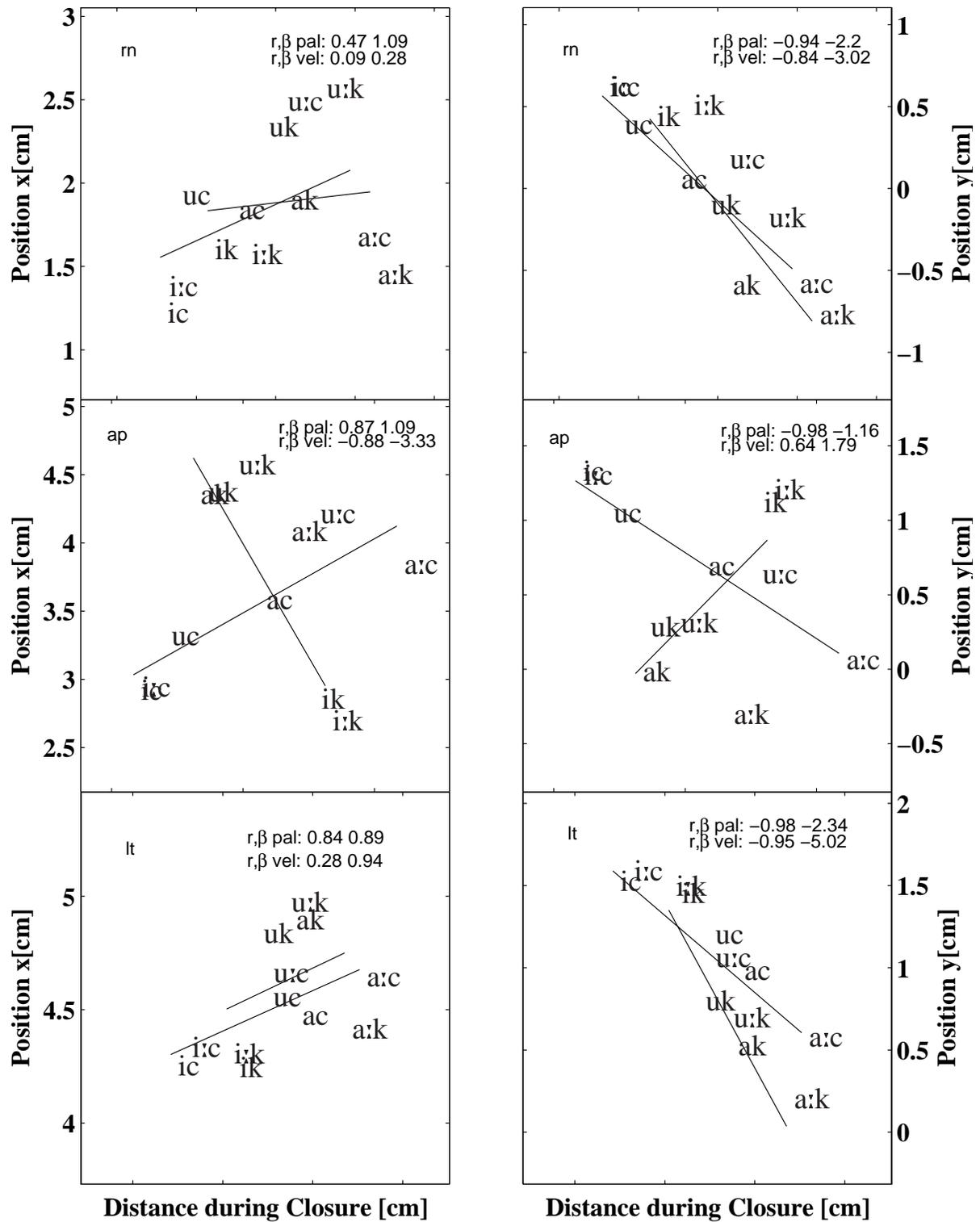


Figure 6. Correlations between the positions in the mid of the first vowel and the distances traveled during the stops, both for the tongue dorsum sensor. The left panel shows the correlation of distance and x-position and the right panel of distance and y- position.

- Coarticulatory resistance of the palatal stop: Not very prone to vowel-like influences (a) the palatal itself exhibits a very stable configuration in comparison to the velar. This stable configuration is similar to an /i/-like shape (b) the palatal itself exerts an influence on the vowel articulation inasmuch as the size of the vowel space is shrunk in comparison to vowel space in the context of the velar stop.
- The initial palatal in the context of the back vowel /u:/ gets isolated from the relatively stable configuration we observed for the remaining palatals.
- Another important finding were the extremely large distances the palatal has to travel in the context of the long /u:/ over the whole VCV-sequence.
- For two out of our three speakers, the movement amplitude of the palatal in /u:/-context *during* closure is among the highest amplitudes altogether.

In the introductory section, we contrasted predictions about the gross coarticulatory variability patterns made by different approaches: Keating's approach creates its prediction from underspecification at different phonological or phonetic levels. This amounts to relatively little sensitivity to vowel-induced contextual coarticulation for palatals in contrast to velars if the derivations we used are the intended ones for the Hungarian language. In contrast, the DAC-scale by Recasens, adapting ideas from the theory of coproduction proposed by Fowler, predicts similar patterns in the tendency to coarticulate with the vowel environment for both palatals and velars: It assigns palatals and velars the same values on the DAC-scale. On this macro-level, Keating's approach makes better descriptions. Here, the necessity arises to summarize the influences Keating sees at work in shaping coarticulatory patterns. These are the following factors: (a) production constraints, operating both within and across languages, (b) constraints deriving from language-specific phonological structure and (c) language-particular constraints, unrelated to production or phonology and therefore unpredictable. Farnetani & Recasens (1999, p.49) claim that cross-consonant differences and cross-language similarities in patterns of coarticulatory resistance (under-)specification held constant point to consonant-specific production constraints going beyond language peculiarities. We found the initial palatal stop in the context of /u:/ strongly coarticulating with the following vowel. We do not have cross-language data on palatal stops, but tried to relate this pattern to consonant-specific production constraints: To meet the objective of a palatal closure, the tongue has to travel a long distance in the context of /u:/. These were present in the form of long distances the tongue back and dorsum were moving. As these cannot be attributed to the influence of the jaw, we interpret them as consonant-specific production constraints for the palatal in back vowel context.

In the light of the data presented in this paper, Recasens' assignment of the same DAC-value for palatals and velars is unlikely because (a) the palatals themselves

are articulated more consistently, and (b) they exert a bigger influence on the neighbouring vowel than the velars. For further substantiating our finding of a higher DAC value for the palatal we plan to analyse the postconsonantal vowel /a/ in the /CVCa/ sequences. If our hypothesis is true then the vowel /a/ should vary with the preceding medial vowel quality to a greater degree for the velar as compared to the palatal. In our view, several reasons could be responsible for the lack of fit to the data the DAC exhibits: One reason could lie in a too strong emphasis on the idea of the coupling of articulators neglecting a general articulatory movement economy aiming at the avoidance of long traveling paths. This might be augmented by Recasens' preference of the palatographic acquisition method, which does not take into account contacts behind the hard palate and also movements leading to tongue-palate contact.

A second important claim we elaborated in the introductory section is that articulatory velar fronting is a "transparency effect of the velar with respect to Backness" and "something that happens gradually over the course of the velar" as stated by Keating. She explicitly refers to Houde (1968), who first observed "sliding" movements over the course of velar consonants. If these sliding movements were an effect of surface underspecifications, then these movements should be unobservable over the course of palatals. This claim was strongly disconfirmed by our data: The movement amplitudes during the palatals, -especially in the context of the vowel /a/- can exceed the amplitudes in velar context. Note that these movements, beforehand only observed for the velars and called "loops" (see Mooshammer et al., 1995), have stimulated a whole research program in speech production as it has been seen as a paradigmatic case for the most different influences simultaneously shaping planning of the movement trajectory. Accordingly, several competing explanations were given for the phenomenon: The looping patterns are seen as a passive forward movement of the tongue due to airstream mechanisms Kent & Moll (1972) or as a result of an active gesture aiming at the maintenance of voicing by Houde (1968), sometimes termed 'cavity enlargement' (Ohala, 1983). Löfqvist & Gracco (2002) to explain looping patterns in more general principles of motor control, postulating the entire movement to be planned in terms of cost minimization principles. A recent modelling study by Perrier, Payan, Zandipour and Perkell (2003) emphasizes the role of tongue biomechanics. We think that palatal stops can essentially contribute to this research program.

We began the introduction with reference to the mechanisms that might underlie diachronic velar fronting¹¹. Recasens (1990) argues that the front velar converges to a back palatal with respect to its articulatory configuration. Concerning this question, the results are ambiguous. While the speaker-independent factor solution and the single factor solution suggest a relatively strict separation in the shapes of the fronted velar and the palatal, this situation is ambiguous for

¹¹let us note again here that the Hungarian palatal is probably *not* the output of velar fronting.

the two other speakers. What we have more robustly, though, is the change in the articulatory configuration for the initial palatal stop in the context of /u:/, which we interpret as an excess of a threshold of maximum coarticulatory resistance. So if the diachronic process of velar palatalization has an articulatory grounding at all, this grounding could as well be afforded by a bidirectional articulatory change of contextual fronting of the velar and contextual backing of the palatal, rather than unidirectionally triggered by the velar.

5 ACKNOWLEDGEMENT

Work supported by German Research Council(DFG) grant GWZ 4/8-1, P.1

6 References

References

- Bhat D. (1978). A general study of palatalization. In: J. Greenberg (ed.) *Universals of Human Language. Vol. 2.: Phonology*, 47–92. Stanford: Stanford University Press.
- Bladon R. & Al-Bamerni A. (1976). Coarticulation resistance in English /l/. *Journal of Phonetics*, 4:137–150.
- Blumstein S. (1986). On acoustic invariance in speech. In: *Invariance and Variability in Speech Processes*, 179–197. Hillsdale, New Jersey, London: Lawrence Erlbaum Associates.
- Blumstein S. & Stevens K. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America*, 1001–1017.
- Boersma P. & Weenink D. (1992–2004). Praat, a system for doing phonetics by computer. www.praat.org. URL www.PRAAT.org.
- Cattell R. & Cattell A. (1955). Factor rotation for proportional profiles: analytical solution and an example. *British Journal of Statistical Psychology*, 8:83–92.
- Daniloff R. & Hammarberg R. (1973). On defining coarticulation. *Journal of Phonetics*, 1:239–248.
- Farnetani E. & Recasens D. (1999). Coarticulation in recent speech production theories. In: W.J. Hardcastle & N. Hewlett (eds.) *Coarticulation; Theory, Data and Techniques*, 31–65. Cambridge: University Press.
- Fowler C. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 8:113–133.

- Geng C., Fuchs S., Mooshammer C., & Pompino-Marschall B. (2003). How does vowel context influence loops? In: *Proceedings of the SPS6: Sydney*.
- Guion G. (1998). The role of perception in the sound change of velar palatalization. *Phonetica*, 55:18–52.
- Hammarberg R. (1976). The metaphysics of coarticulation. *Journal of Phonetics*, 4:353–363.
- Harshman R., Ladefoged P., & Goldstein L. (1977). Factor analysis of tongue shapes. *Journal of the Acoustical Society of America*, 69:693–707.
- Harshman R. & Lundy M. (1984). Data preprocessing and the extended parafac model. In: H. Law (ed.) *Research Methods for Multimode Data Analysis*, 216–284. New York: Praeger.
- Hoole P. (1996). Issues in the acquisition, processing, reduction and parameterization of articulographic data. *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München*, 34:158–173.
- Houde R. (1968). A study of tongue body motion during selected consonant sounds. Speech Communications Research Laboratory, Santa Barbara, SCRL Monograph 2.
- Keating P. (1993). Phonetic representation of palatalization versus fronting. *UCLA Working Papers in Phonetics*, 85:6–21.
- Keating P. & Lahiri A. (1993). Fronted velars, palatalized velars, and palatals. *Phonetica*, 73–101.
- Keating P.A. (1988). Palatals as complex segments: X-ray evidence. *UCLA Working Papers in Phonetics*, 69:77–91.
- Keating P.A. (1990). The window model of coarticulation: articulatory evidence. In: J. Kingston & M.E. Beckman (eds.) *labphon1*, 451–470. Cambridge: Cambridge University Press.
- Keating P.A. (1991). Coronal places of articulation. In: C. Paradis & J.F. Prunet (eds.) *Phonetics and Phonology 2: The special status of coronals: Internal and external evidence*, 29–48. Academic Press.
- Kent R. & Moll K. (1972). Cinefluorographic analyses of selected lingual consonants. *Journal of Speech and Hearing Research*, 15:453–473.
- Lahiri A. & Blumstein S.E. (1984). A re-evaluation of the feature coronal. *Journal of Phonetics*, 12:133–146.
- Lindblom B., Guion S., Hura S., Moon S., & Willerman R. (1995). Is sound change adaptive? *Rivista di linguistica*, 7:5–37.

- Löfqvist A. & Gracco V.L. (2002). Control of oral closure in lingual stop consonant production. *Journal of the Acoustical Society of America*, 111:2811–2827.
- Mooshammer C., Hoole P., & Kühnert B. (1995). On loops. *Journal of Phonetics*, 23:3–21.
- Munhall K., Ostry D., & Flanagan J. (1991). Coordinate spaces in speech planning. *Journal of Phonetics*, 19:293–307.
- Ohala J.J. (1983). The origin of sound patterns in vocal tract constraints. In: P.F. MacNeilage (ed.) *The production of speech*, 189–216. New York: Springer.
- Ohala J.J. (1993). The phonetics of sound change. In: C. Jones (ed.) *Historical linguistics: Problems and perspectives*, 237–278. London: Longman.
- Öhman S. (1967). Numerical model of coarticulation. *Journal of the Acoustical Society of America*, 41:310–320.
- Öhman S.E.G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, 39:151–168.
- Perrier P., Payan Y., Zandipour M., & Perkell J. (2003). Influences that shape tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study. *Journal of the Acoustical Society of America*, 114:1582–1599.
- Recasens D. (1990). The articulatory characteristics of palatal consonants. *Journal of Phonetics*, 18:267–280.
- Recasens D. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, 102:544–561.
- Recasens D. (2002). An ema study of vcv coarticulatory direction. *Journal of the Acoustical Society of America*, 111:2828–2840.
- Recasens D. (2003). Articulation and sound change in romance. In: *Proceedings of the 15th International Congress of Phonetic Sciences*, 231–234.
- Saltzman E. & Munhall K. (1989). A dynamical approach to gestural patterning in speech. *Haskins Status Report on Speech Research*, 99/100:38–68.
- Siptár P. & Miklós T. (2000). *The phonology of Hungarian*. Oxford: Oxford Univ. Press.