ZASPiL Nr. 52 – September 2010

Papers from the Linguistics Laboratory

Melanie Weirich & Stefanie Jannedy
(Eds.)

Table of Contents

Adresses of Contributors

Laura J. Downing
ZAS Berlin
Schützenstr. 18
10117 Berlin
Email: downing@zas.gwz-berlin.de

Susanne Fuchs
ZAS Berlin
Schützenstr. 18
10117 Berlin
Email: fuchs@zas.gwz-berlin.de

Scott Grimm
Department of Linguistics
Stanford University
Stanford, CA 94305-2150
USA
Email: sgrimm@stanford.edu

Stefanie Jannedy
ZAS Berlin
Schützenstr. 18
10117 Berlin
Email: jannedy@zas.gwz-berlin.de

Karsten A. Koch
ZAS Berlin
Schützenstr. 18
10117 Berlin
Email: koch@zas.gwz-berlin.de

Bernd Pompino-Marshall
Humboldt-Universität zu Berlin
Institut für Deutsche Sprache und Linguistik
Unter den Linden 6
Sitz: Hegelplatz 2
10099 Berlin
Email: bernd.pompino-marschall@rz.hu-berlin.de

Blake Rodgers
Department of Linguistics
University of Wisconsin-Madison
Madison, WI 53706
USA
Email: brodgers@wisc.edu

Melanie Weirich
ZAS Berlin
Schützenstr. 18
10117 Berlin
Email: weirich@zas.gwz-berlin.de

Marzena Żygis
ZAS Berlin
Schützenstr. 18
10117 Berlin
Email: zygis@zas.gwz-berlin.de

Glottal Marking of Vowel-Initial Words in German

Bernd Pompino-Marschall
Institut für deutsche Sprache und Linguistik der Humboldt-Universität zu Berlin
Zentrum für Allgemeine Sprachwissenschaft, Berlin

Marzena Żygis
Zentrum für Allgemeine Sprachwissenschaft, Berlin

Glottal marking of vowel-initial German words by glottalization and glottal stop insertion were investigated in dependence on speech rate, word type (content vs. function words), word accent, phrasal position and the following vowel. The analysed material consisted of speeches of Konrad Adenauer, Thomas Mann and Richard von Weizsäcker. The investigation shows that not only the left boundary of accented syllables (including phrasal stress boundary) and lexical words favour glottal stops/glottalization, but also that the segmental level appears to have a strong impact on these insertion processes. Specifically, the results show that low vowels in contrast to non-low ones favour glottal stops/glottalization even before non-accented syllables and functional words.

---

1 Introduction

Glottal stops and glottalization have attracted increasing interest from phonologists and phoneticians, probably due to their hardly definable, almost chameleon-like behavior.

Several phonetically oriented studies have reported huge variability of glottal stops and glottalizations in their acoustic realization. Moreover, their presence or absence – unless they have a phonemic status in a given language – is also subject to variation despite well-defined rules (see the results of the present study). This inter- and intra-speaker variability has been observed in a number of languages and shown to be dependent on several parameters such as: phrasal position, accented vs. unaccented syllable, segmental context, speech rate, dialect, speaker's gender, and others. (see e.g. *American English*: Umeda 1978, Pierrehumbert 1994, Pierrehumbert & Talkin 1991; Redi & Shattuck-Hufnagel 2001; *Chitwan Tharu*: Leal 1972, *Danish*: Haberland 1994; *Garo*:

Burling 1992, *Nootka*: Shank & Wilson 2000; *Tümpisa (Panamint) Shoshone*: Dayley 1989).

From a phonological point of view the discussion often focuses on phonological alternations, (e.g. /h/- vs. [ʔ]-alternation in *Pima Bajo*, Fernández 1996), grammatical functions of glottal stop insertions (e.g. glottal stop as a marker of the irrealis mood, i.e. future, imperative, and purposive verb forms in *Nhanda* Blevins & Marmion 1995), the appearance of glottal stops at prosodic constituent boundaries, (e.g. glottal stop appearance in utterance-initial position in *Anejom̃* Lynch 2000) or phonotactic restrictions, (e.g. in *Nootka* glottalized elements are banned from the coda Shank & Wilson 2000).

As far as German, the language of the main interest for the present study, is concerned it has not been extensively investigated as far as glottalizations and glottal stops are concerned. Phonetically, the topic was investigated by e.g. Krech (1968), Kohler (1994), Rodgers (1999) (see section 0). Phonologically, it has been stated that glottal stops are found as onsets of vowel-initial stressed syllables. Hall (1992) shows that the glottal stop occurs optionally at the beginning of a vowel-initial foot, cf. examples in (1). The same view is shared by Wiese (2000).

(1) Glottal Stop Insertion in German (Hall 1992:58).

arm	[ʔaʀm]	or	[aʀm]	'poor'
oft	[ʔɔft]	or	[ɔft]	'often'
Theater	[te.ʔáː.tɐ]	or	[te.áː.tɐ]	'theater'

An analysis presented by Alber (2001) refers to morphological boundaries and shows that a glottal stop appears in Standard German in two contexts: (i) at the left edge of a vowel-initial stressed syllable and (ii) at the left edge of a vowel-initial root or prefix. Alber (2001:6) concludes 'that stress can favor the presence of glottal closure, though it is less clear whether the difference between stressed and unstressed syllables is strong enough to be integrated into a phonological analysis.'

The aim of this paper is to investigate the appearance of glottal stops (and glottalization) in German spontaneous speech by looking what extent this appearance depends on other parameters such as speech rate, prosodic boundary, phrasal position, accented vs. unaccented segmental context, content vs. function words, and how large the inter- and intraspeaker variation is with respect to the presence/absence of glottal stops and glottalizations.

The paper is organized as follows. In section 2 we provide a brief overview of studies on glottal stops and glottalization focusing on three selected papers. Section 3 introduces the experimental design of the present study and section 4 provides results. Main conclusions of the paper are summarized and discussed in section 5.

2 Glottal stops and glottalizations: Previous studies

As mentioned in the previous section, glottal stops and glottalization have sparked much interest in phonetics and phonology. In this section we focus on three papers, namely Umeda (1978), Kohler (1994) and Redi & Shattuck-Hufnagel (2001) which in our view are of particular importance for the present study.

2.1 Umeda (1978)

The use of glottal stops in continuous speech in American English has been the subject of Umeda's (1978) study. In this study, five speakers read sentences excerpted from four short stories (each story contained about 30 sentences). The sentences were randomized and read in form of a list. In the second part of the experiment the informants read the stories in their original form. The data were analyzed with respect to the occurrence of glottal stops in dependence of several factors such as speaker's reading style, difficulty of the material, phonemes following the glottal stops, stress conditions, type of words (function vs. content words), and frequency of occurrence of words. In general, the results lead to the conclusion that the speaker's reading style and difficulty of material influence the insertions of glottal stops stronger than phonological and grammatical factors. The study emphasizes great inter-speaker variation related to individual reading styles (the staccato type of speaking shows more glottal stops than fluent speaking). Slow speech rate also induces a higher percentage of glottal stops than high speech rate (the slowest speaker showed the largest number of glottal stops, i.e. about 80% of the total occurrence of vowel-initial content words with the first syllable stressed, and the fastest speaker the least, i.e. about 25%). It was also shown that grammatical breaks between phrases of the same kind as e.g. listings of nominal phrases, breaks after adverbs *however*, *instead*, or clause boundaries induce a high percentage of glottal stop insertions even before function words.

The study furthermore shows that rare words are more frequently marked with glottal stops than common words. No significant differences in glottal stop assignment have been found regarding the effect of tongue height in vowels and the presence/absence of the pitch contour on vowels.

2.2 Kohler (1994)

Kohler's (1994) paper focuses on glottal stops and glottalization occurring in German colloquial read speech of a North German non-dialect variety. On the basis of an auditory and acoustic study of speech data from 36 speakers of the PHONDAT 90/93 Kiel database (IPdS 1994, 1995ff.), Kohler reports on the occurrence of glottal stops and glottalization as morphological boundary markers including word-internal boundaries on the one hand and as reduction phenomena of supraglottal stop articulations on the other hand.

The results show that glottalization is the most frequent realization of the canonical glottal stop in German. After pauses/silence, the presence of a glottal stop is substantially more frequent than its absence. Stressed vowels favour the presence of glottal stops/glottalization substantially more than unstressed ones in a continuous segmental context. The highest proportion of glottal stops is found before stressed vowels and after plosives (72 %). The preceding plosive context also triggers the highest proportion of glottal stops before unstressed vowels (35 % as opposed to 17%-20% in other contexts, i.e., preceding vowels, sonorants, and fricatives.) As far as word-internal boundaries are concerned it has been found that between prefix and stem as well as between components of compounds simple glottalization is the most frequent pattern (62%) followed by glottal stop and glottalization (24%) and glottal stop deletions (10%). The least frequent realisation was a single glottal stop (4%). Finally, Kohler (1994) shows that glottal stops and glottalization replace (i) stops /t/ and /k/ at the word boundary, e.g. Freita[k] abend 'Friday evening' is realized as Freita[ʔ] abend and (ii) syllable-final plosives, especially before and/or after a nasal, e.g. hi[ntən] is pronounced as hi[nʔn] 'behind'. The former case has been found in 4 instances and the latter in 48 (Kohler 1994: 45).

2.3 Redi & Shattuck-Hufnagel (2001)

This study reports on glottalization in American English observed in different contexts, for instance in word-initial vowels at intonation phrase onsets and at pitch accents as well as at the ends of utterances. The study also investigates glottalized voice quality occurring as an allophone of voiceless stops (in e.g. *butler)*. Speech material was obtained from two corpora of spoken utterances: Labnews consisted of read speech produced by six professional radio news announcers. A part of the materials was also produced by four nonprofessional speakers. The second corpus, called the ABC corpus, contained read speech produced by four nonprofessional speakers of AE.

The acoustic analysis shows that glottalization is realized as (i) aperiodicity (irregular periods), (ii) creak (lowering of fundamental frequency with near-total damping), (iii) diplophonia (alternation in shape, amplitude, or duration of successive periods) and (iv) glottal squeak (occurring adjacent to other types of glottalization, and perceived as a momentary shift to a relatively high-pitched and low amplitude voice quality). The results show a great range in the rate of glottalization in individual speakers' pronunciations. It appears that this rate is higher on words at the ends of utterances than on words at the ends of utterance-medial intonational phrases. Finally, it is shown that the rate of glottalization is higher at the boundaries of full intonational phrases than at intermediate intonational phrases.

3 Present study: Method

The glottal marking of vowel-initial words were analysed in historical recordings of prominent Germans (Konrad Adenauer, Thomas Mann, Richard von Weizsäcker) taken from the CD collection "Tondokumente zur deutschen Geschichte" (Audio documents of German history; Stiftung deutsches Rundfunkarchiv 2004ff.). Three recordings of each speaker were extracted from the audio CDs to *.wav files using CDex, leaving the sample rate at 44.1 kHz and segmented/annotated in PRAAT. Statistical analyses were conducted with the help of the STATVIEW software.

Our three selected speakers were Konrad Adenauer (1876-1967; first Federal German chancellor 1949-1963); Thomas Mann (1875-1955, famous writer and Nobel prize winner 1929) and Richard von Weizsäcker (*1920, Federal German president 1984-1994). The acoustic analysis included three speeches of each speaker:
(i) Adenauer: 1929, 1949a, b,
(ii) Mann: 1945, 1949, 1950,
(iii) von Weizsäcker 1984, 1989, 1992.
The analyzed material summed up to 29.04 minutes of recording time, cf. details presented in Table 1.

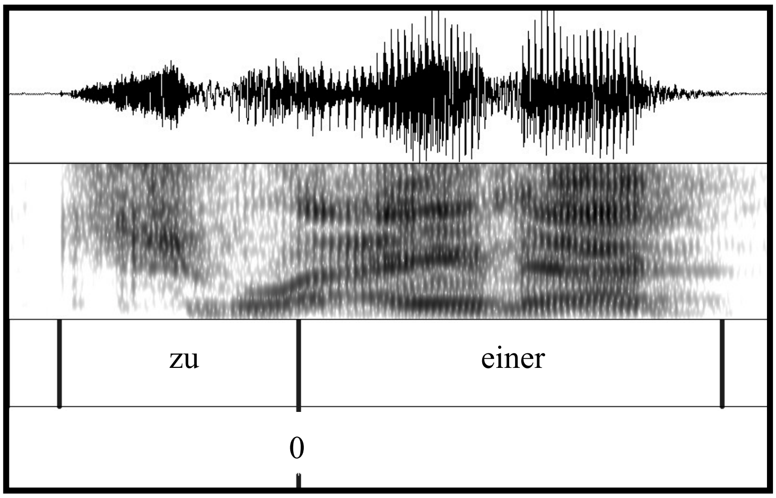
3.1 Procedure

The recordings were first transcribed orthographically and segmented with respect to pauses (>150 ms). Information on the duration of pause delimited utterances, the number of syllables produced and the duration of the following pause were stored in tabular form for the later analysis of speech rate. The orthographic transcript was then scanned for vowel initial words.[1]

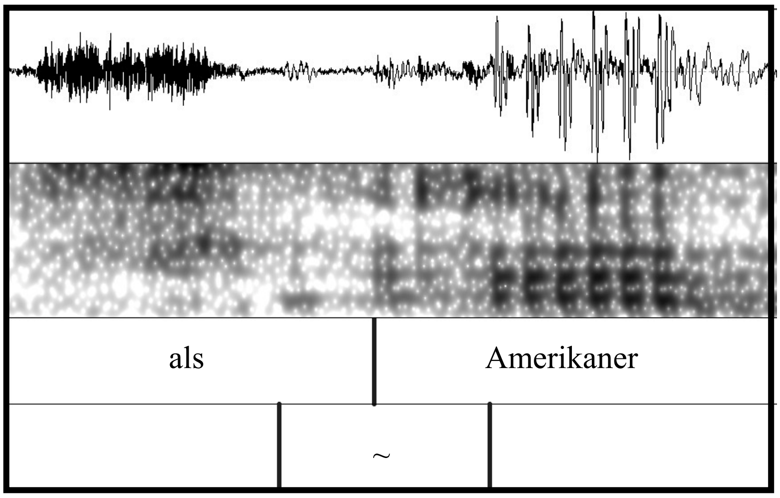
[1] Word-internal boundaries as after prefixes (e.g. *vereisen* [fɐʔaɪzən]) and in compounds (e.g. *Feierabend* [faɪɐʔaːbənt]) are excluded from the current analysis.

In a second annotation/segmentation step these word-initial vowels were classified as being unmarked, marked by glottalization (creaky voice) or by a glottal stop proper. Marked items were segmented with respect to the beginning and end of single pitch periods, i.e. glottalizations from the first to the last irregular pitch period (onset of modal voicing), glottal stops from burst/first affected pitch period to the onset of modal voicing (segmentation of the closed phase was difficult or impossible (i.e. utterance-initially). Figure 1 illustrates segmentation in examples from Mann (1950).

(a)



(b)



(c)

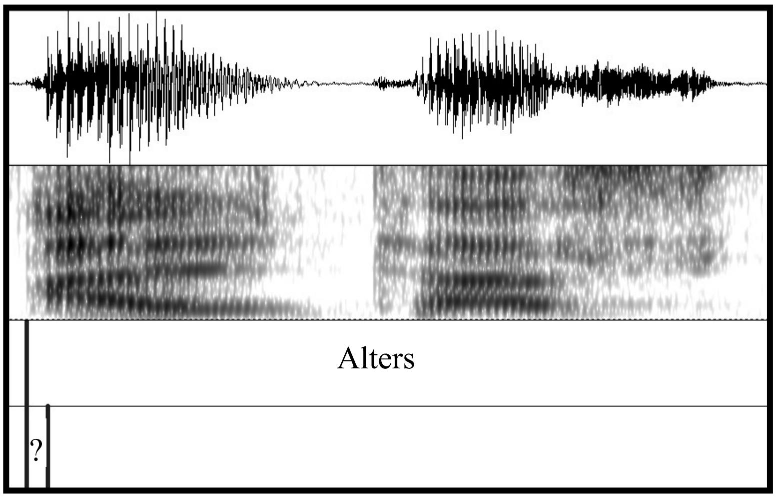


Figure 1: Sample segmentations of word-initial vowel markings in the material of Mann (1950): (a) unmarked, (b) marked by glottalization (61 ms), (c) marked by glottal stop (18 ms).

4 Results

In a first step the segmented material was analyzed with respect to its general temporal makeup as described in 4.1. In 4.2 we will then present the results with respect to glottal markings in vowel-initial words.

4.1 General

Table 1 gives an overview of the results with respect to recording durations[2] and speakers' recording durations (with and without intervening pauses), single utterance length, syllable counts, pause durations and speech rate (articulatory rate, i.e. syllables/utterance duration [s]) depending on speaker and recording. Table 1 also includes the statistical results of ANOVAs and post hoc Scheffe tests with speaker and take as independent variables, and utterance length, syllable count, pause duration or articulatory rate as dependent variables; *: $p < .05$; **: $p < .01$; ***: $p < .001$).

[2] Values differing from those given on the resource CD are due to initial and final silent parts in the original material and/or material excluded in the present analysis (e.g. reporter's voice in Adenauer 1949a).

The utterance durations are generally longest for Richard von Weizsäcker, significantly differing from those of Konrad Adenauer as well as Thomas Mann. This may be due to their nature as highly prepared Christmas broadcasts in TV. Whereas Richard von Weizsäcker as well as Thomas Mann show the same utterance durations in all their speeches, Konrad Adenauer's utterance durations are significantly longer in his governmental address (Adenauer 1949b), again possibly due to the nature of this highly prepared text. These results are paralleled by significantly larger syllable counts together with a marginal difference between two recordings of Thomas Mann. Pause durations show up more variable between and within speakers, Thomas Mann here showing the smallest values throughout. Speech rate turned out to be quite speaker–specific, with clear significant differences between the fastest speaker (Richard von Weizsäcker) and the others as well as between the following two, i.e. Konrad Adenauer and Thomas Mann. Only Thomas Mann exhibits a slower speech rate once, i.e. in his post-war BBC broadcast. This could possibly be due to technical reasons. For the more detailed analysis of glottal markings given below, we therefore decided to base speech rate dependent analysis on the quartile ranges (also given in Table 1 and depicted in Figure 2) for the three speakers separately.

Table 1: Statistical data of the analysed material in a general overview

speaker	Konrad Adenauer			Thomas Mann			Richard von Weizsäcker		
date	1929	1949a	1949b	1945	1849	1950	1984	1989	1992
recording duration [min’s”]	2’15”	2’40”	3’58”	2’17”	6’43”	4’34”	2’55”	1’44”	2’38”
speaker sum	8’53”			13’34”			7’17”		
speaking duration [min’s”]	1’30”	4’44”	2’44”	1’47”	4’44”	3’16	2’07”	1’13”	1’46”
speaker sum	6’00”			9’47”			5’06”		
utterance duration mean (sd) [s]	*** 1.50 (.848)	1.77 (.937) **	*** 2.34 (1.236) **	1.99 (1.068)	1.79 (1.022)	1.899 (1.109)	2.58 (1.256)	2.10 (1.267)	2.13 (.929)
speaker mean (sd) [s], N	** 1.90 (1.088) 239			1.86 (1.058) 458 **			** 2.29 (1.161) 258 **		
syllables per utterance mean (sd) range	*** 5.67 (3.261) 1-15	6.92 (4.077) 2-18 **	*** 9.04 (5.057) 1-21 **	* 8.11 (4.653) 2-20	* 6.35 (4.445) 1-26	6.43 (3.841) 1-18	12.61 (6.757) 3-31	10.34 (6.668) 2-33	11.18 (5.263) 1-29
speaker mean (sd) range	** 7.31 (4.457) 1-21			6.68 (4.331) 1-26 **			** 11.49 (6.237) 1-33 **		
pause duration mean (sd) [ms]	* 754 (501.0)	903 (486.0)	* 1064 (971.2)	* 552 (286.6) *	* 752 (524.7)	765 (492.7) *	1014 (576.6)	878 (761.3)	1035 (693.7)
speaker mean (sd)	** 916 (718.2)			** 722 (486.2) ***			986 (671.3) ***		
speech rate mean (sd) [syll./s]	3.79 (.817)	3.86 (.886)	3.81 (.804)	*** 4.05 (.742) ***	*** 3.49 (.906)	3.45 (.902) ***	4.83 (.657)	4.80 (.841)	5.16 (.983)
speaker mean (sd)	** 3.82 (.831) ***			**/*** 3.57 (.904)			*** 4.95 (.849) ***		
speech rate quartiles [syll./s]	3.305 3.890 4.179	2.837 3.729 4.334	3.192 3.844 4.376	3.544 4.101 4.598	2.795 3.431 4.131	2.786 3.347 3.946	4.378 5.000 5.207	4.305 4.925 5.260	4.525 5.166 5.756
speaker mean quartiles	3.305 3.831 4.317			2.889 3.522 4.131			4.410 5.010 5.394		

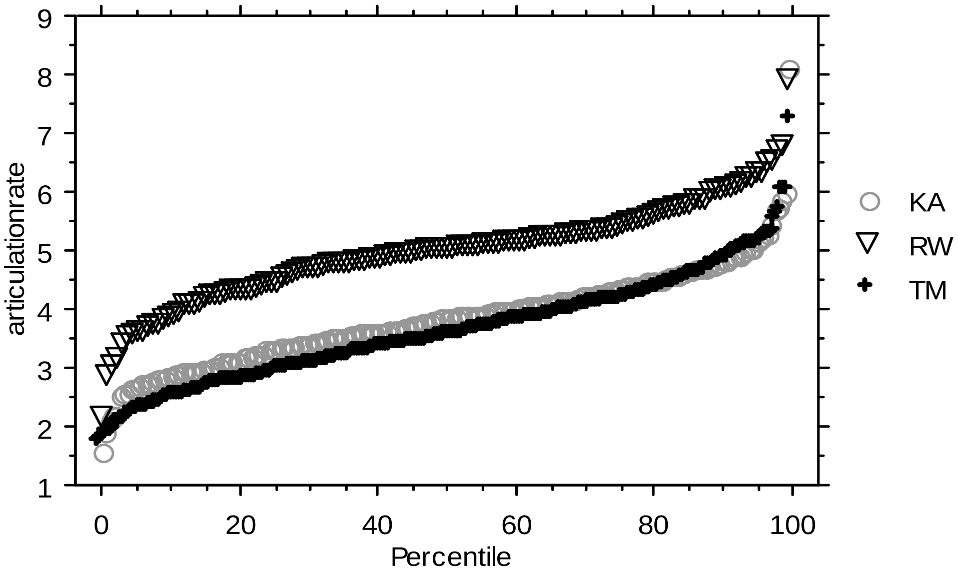


Figure 2: Percentile plot of the distribution of speech rates per utterance split by speakers

4.2 Marking of word-initial vowels

The distribution frequencies of the differently marked vowel-initial words were analyzed with respect to the speaker specific speech rates. So the nonmarked, glottalized and abrupt vowel onsets (i.e. those with preceding glottal stop) were counted separately in slow (i.e. first quartile of the speakers' rate range, cf. Figure 2), slow medium (second quartile), fast medium (third quartile) and fast rate (fourth quartile) utterances. The following tables always show the relative frequencies of marking type for each single condition (Table 2: speaker/tempo; Table 3: word type/tempo; Table 4: stress/tempo; Table 5: phrasal position/tempo; Table 6: vowel height/tempo) since absolute counts were not under control due to the experimental design. Note that further splitting up was not feasible since the number of occurences for different tokens would have become too small.

Table 2 depicts the speech rate dependent distribution of glottal markings in vowel-initial words for our three speakers (separately and in general). As can be seen, the glottal marking of word-initial vowels is generally diminishing with increasing speech rate: nonmarked items continuously rise in frequency from about 30% in slow speech (1^{st} quartile of the speakers' rate ranges) to more than 50% in fast speech; on the other hand, realisations of canonical glottal stop stepwise reduce from 48% in slow speech to ca. 16% in fast speech, parallel to an increase in markings by creaky voice only (from about 22% to about 31%).

Besides this general trend there are some interesting differences between the speakers: For Adenauer we can see a change from glottal stop marking at slower rates to a loss of marking in faster speech. For Weizsäcker, who is from a younger generation, marking by glottalization seems to be preferred at almost all rates (with the exception of equally frequent markings by glottal stop at the slowest and some more losses of marking at the fastest rate). Thomas Mann's behaviour, lying in between both these extremes, more closely resembles that of Adenauer in this respect.

Table 2: Rate-dependent relative frequencies of glottal markings of word-initial vowels (0: unmarked, ~: creaky voice, ʔ: glottal stop; speaker-dependent and overall; most frequent marking per rate (and speaker) highlighted in bold (with equal frequencies also in italics))

marking	speaker	slow		slow medium		fast medium		fast	
	KA	40.0		37.5		**50.0**		**64.6**	
0	TM	24.1	30.1	**45.8**	**38.9**	**43.2**	**40.0**	**53.9**	**53.2**
	RW	29.4		28.8		26.0		**44.4**	
	KA	08.6		15.6		11.5		22.9	
~	TM	22.2	22.0	13.3	23.4	24.2	30.7	26.1	30.6
	RW	***35.3***		**44.2**		**55.2**		43.1	
	KA	**51.4**		**46.9**		38.5		12.5	
ʔ	TM	**53.7**	**48.0**	41.0	37.7	32.6	29.3	20.0	16.2
	RW	***35.3***		26.9		17.9		12.5	

In Table 3 the relative frequency of different markings is given with respect to word type (content words vs. function words). The fact of diminishing glottal marking with increasing speech rate mentioned above clearly shows up to be dependent on word type: Content words are more resistant to the total loss of marking at higher rates in contrast to the generally unstressed function words.

Table 3: Rate-dependent relative frequencies of glottal markings in content words (c) vs. function words (f) (0: unmarked, ~: creaky voice, ʔ: glottal stop; most frequent marking per rate and word type in bold)

	slow		slow medium		fast medium		fast	
	c	f	c	f	c	f	c	f
0	15.4	36.9	23.6	**46.9**	22.4	**51.1**	25.7	**66.1**
~	20.5	22.6	32.7	18.9	**42.1**	23.7	**54.3**	19.1
ʔ	**64.1**	**40.5**	**43.7**	34.2	35.5	25.2	20.0	14.8

Therefore, in Table 4 the effect of stress on the relative frequency of glottal markings in vowel-initial content words is shown. Due to the smaller number of tokens the results are not as clearcut as before, but the general tendency of unstressed items to lose markings at lower rates than stressed ones is still visible.

Table 4: Rate-dependent relative frequencies of glottal markings in stressed vs. unstressed content words (f) (0: unmarked, ~: creaky voice, ʔ: glottal stop; most frequent marking per rate and word type in bold (with equal frequencies also in italics))

	slow		slow medium		fast medium		fast	
	+stressed	-stressed	+stressed	-stressed	+stressed	-stressed	+stressed	-stressed
0	10.7	27.3	11.4	***42.1***	21.8	25.0	12.5	**54.6**
~	14.3	***36.35***	42.9	15.8	**43.6**	**40.0**	**68.75**	22.7
ʔ	**75.0**	***36.35***	**45.7**	***42.1***	34.5	35.0	18.75	22.7

The effect of phrasal position on the nature of glottal marking, shown in Table 5, also shows a tendency of stronger marking in phrase initial items but this effect seems rather weak, i.e. only visible at slow medium rate, cf. Rodgers (1999).

Table 5: Rate-dependent relative frequencies of glottal markings in phrase-initial words vs. non phrase-initial words (f) (0: unmarked, ~: creaky voice, ʔ: glottal stop; most frequent marking per rate and position in bold)

	slow		slow medium		fast medium		fast	
	initial	non-initial	initial	non-initial	initial	non-initial	initial	non-initial
0	33.3	29.3	28.2	**42.2**	**50.0**	**37.2**	**59.2**	**51.6**
~	12.5	24.2	00.0	30.5	11.9	35.5	06.1	37.1
ʔ	**54.2**	**46.5**	**71.8**	27.3	38.1	27.3	34.7	11.3

So far we may conclude in parallel to the published results that the marking of vowel-initial words in German is realized (a) in a rate-dependent manner ranging from glottal stop insertion via glottalisation to no marking at all. The strength of marking is influenced by the intervening variables of (b) word type, (c) stress and (d) phrasal position: content words are more strongly marked than function words, stressed initial vowels more strongly than unstressed ones. Position within the phrase affected the speech rate dependent realization of glottal marking only at slow medium rate of articulation.

In contrast to Umeda (1978), who found no differences of glottal marking with respect to vowel identity in American English, our German material clearly shows a dependency of marking on vowel height (cf. Table 6).

Table 6: Vowel-height and rate-dependent relative frequencies of glottal markings of word-initial vowels (0: unmarked, ~: creaky voice, ʔ: glottal stop; most frequent marking per rate and vowel height highlighted in bold (with equal frequencies also in italics))

	slow			slow medium			fast medium			fast		
	low	mid	high	low	mid	high	low	mid	high	low	mid	high
0	07.7	25.0	**45.3**	13.5	***35.3***	**55.3**	18.6	**47.2**	**57.0**	25.6	**56.8**	**70.0**
~	15.4	**40.0**	20.3	28.8	29.4	17.5	39.5	30.6	22.6	**47.4**	32.4	19.2
ʔ	**76.9**	35.0	34.4	**57.7**	***35.3***	26.3	**41.9**	22.2	20.4	26.9	10.8	10.8

We could therefore add to our observations that (e) low vowels are more strongly marked by glottal stops and glottalization than non-low ones, cf. Krech (1968).

In addition to the rate-dependent distributional analysis of glottal markings, we finally analyzed the durational characteristics of glottal markings. The results of this analysis are given in Table 7 below.

An ANOVA of the duration of glottal marking (with marking type, rate and speaker as independent variables) revealed the expected[3] highly significant ($p < .001$) difference between glottalizations and glottal stops and a tendency for an interaction between marking type and rate with a significant ($p < .05$) but not systematic effect of rate for glottalizations (cf. Table 7).

Table 7: Rate-dependent durational characteristics of glottal markings (mean (sd) in ms, [N]; ~: creaky voice, ʔ: glottal stop)

glottal marking	slow	slow medium	fast medium	fast
~	68 (26) [27]	88 (36) [39]	93 (62) [85]	73 (29) [72]
ʔ	21 (14) [58]	20 (9) [63]	20 (9) [62]	19 (11) [38]

[3] Due to our segmentation strategy.

5 Summary and discussion

Glottal marking of vowel-initial words in spontaneous German speech again was shown to be remarkably variable. From a phonetic point of view it is best described as realizations along a continuous scale of vocal fold adduction/compression giving rise to exactly timed canonical glottal stop insertions before onset vowels at the extreme value of glottal activation changing to amalgams of glottal stop and glottalization as well as less strictly timed glottalization at word boundaries (cf. Figure 3) to only minor reflexes in fundamental frequency and no reflexes at all in the stream of voicing.[4]

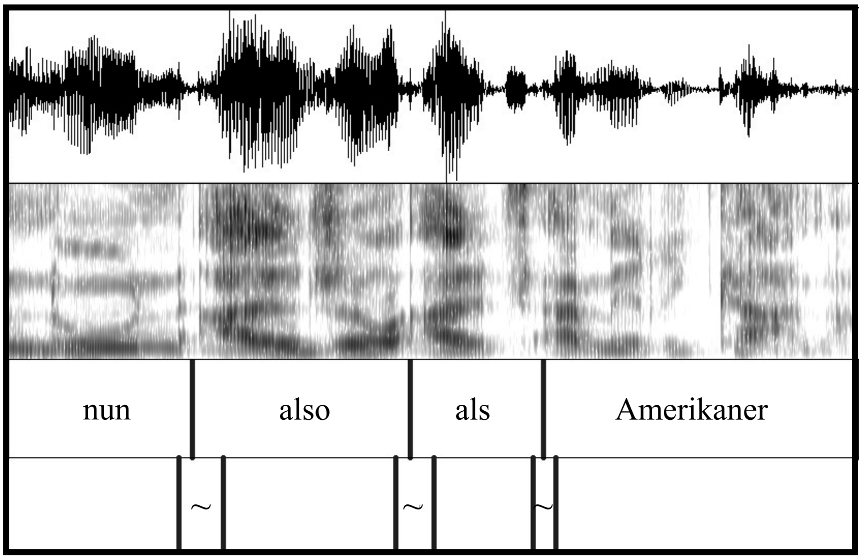


Figure 3: Three examples of glottalizations in sequence spanning the word boundary in Mann (1950)

In any annotation scheme this glottal behavior will result in a more categorical scale. As our results show, the distribution of the different marking categories (i.e. [ʔ], [~] and 0) is highly dependent on a number of factors. But, in general, it is clearly dependent on speech rate, irrespective of all other factors under consideration. This again speaks for its principally gradual nature. Parallel to the published results presented in section 2 we found that linguistic variables like word type, word accent and (to a lesser degree) phrasal position all influence the relative frequency of glottal markings in a rate-dependent way. This shows the importance or even superiority of speech rate over other parameters. On the other hand it could be shown that glottal marking is also dependent on the very

4 In strictly post-pausal position, this continuum is paralleled by hard, glottalized and smooth onset of the initial vowel.

nature (i.e. tongue height) of the marked vowel itself, cf. also Krech (1968). This finding has an important implication for prosodic research, namely, it shows that glottal stops/glottalizations are not only conditioned by prosodic boundaries but depend on segmental level as well. It remains to be answered by future research why glottal stops occur more frequently before low vowels and how this cooccurrence is to be understood in terms of the interplay of segmental and prosodic levels of representations.

Finally, the present study provides important material for phonological approaches to German prosody showing that the present accounts of glottalization and glottal stops need to be extended by taking into consideration the complexity of phonetic evidence.

References

Alber, Birgit (2001). Regional variation and edges: Glottal stop epenthesis and dissimilation in standard and southern varieties of German. *Zeitschrift für Sprachwissenschaft* 20, 3-41.

Adenauer, Konrad (1929). Grundsteinlegung zum Neubau der Universität Köln (26.10.1929). In: Stiftung deutsches Rundfunkarchiv (2005). *Tondokumente zur deutschen Geschichte*. Konrad Adenauer. 5. Januar 1876 - 19. April. 1967 Braunschweig: Archiv Verlag, Track 1.

Adenauer, Konrad (1949a). Feststellung der Annahme des Grundgesetzes durch Konrad Adenauer (23.05.1949). In: Stiftung deutsches Rundfunkarchiv (2005). *Tondokumente zur deutschen Geschichte*. Gründung der beiden deutschen Staaten. Braunschweig: Archiv Verlag, Track 3.

Adenauer, Konrad (1949b). Regierungserklärung des Bundeskanzlers Konrad Adenauer (20.09.1949). In: Stiftung deutsches Rundfunkarchiv (2005). *Tondokumente zur deutschen Geschichte*. Gründung der beiden deutschen Staaten. Braunschweig: Archiv Verlag, Track 5.

Blevins, Juliette & Doug Marmion (1995). Nhanda glottal stop. *Oceanic Linguistics* 34. 139-160.

Burling, Robbins (1992). Garo is a minimal tone language. *Linguistics of the Tibeto-Burman Area* 15.5, 33-51.

Dayley, Jon P. (1989). *Tümpisa (Panamint) Shoshone Grammar*. University of California. Publications in Linguistics vol. 115. University of California Press.

Fernández, Zarina Estrada (1996): *Prima Bajo*. München: Lincom.

Haberland, Hartmut (1994) Danish. In: König, Ekkehard & Johan van der Auwera (eds.). *The Germanic Languages*. London: Routledge, 313-348.

Hall, Tracy Alan (1992). *Syllable Structure and Syllable-Related Processes in German*. Tübingen: Max Niemeyer Verlag.

IPdS (1994), *The Kiel Corpus of Read Speech*. *Vol. I* [Kiel-CD #1]. Institut für Phonetik und digitale Sprachverarbeitung der Christian-Albrechts-Universität zu Kiel.

IPdS (1995ff.), *The Kiel Corpus of Sontaneous Speech*. *Vol. I* [Kiel-CD #2 (1995)], *Vol. II* [Kiel-CD #3 (1996)], *Vol. III* [Kiel-CD #4 (1997)]. Institut für Phonetik und digitale Sprachverarbeitung der Christian-Albrechts-Universität zu Kiel.

Kohler, Klaus J. (1994). Glottal stops and glottalization in German. Data and theory of connected speech processes. *Phonetica* 51, 38-51.

Krech, Eva-Maria (1968). *Sprechwissenschaftlich-phonetische Untersuchungen zum Gebrauch des Glottalschlageinsatzes in der allgemeinen deutschen Hochlautung.* Basel: Karger.

Leal, Dorothy (1972). *Chitwan Tharu Phonemic Summary*. Kirtipur: Summer Institute of Linguistics and Institute of Nepal Studies, Tribhuvan University.

Lynch, John (2000). *A Grammar of Anejom*. [Pacific Linguistics 507]. Canberra: The Australian National University.

Mann, Thomas (1945). "Deutsche Hörer!" – Rundfunkansprache über den deutschsprachigen Dienst der BBC (24.05.1945). In: Stiftung deutsches Rundfunkarchiv (2006). *Tondokumente zur deutschen Geschichte*. Neubeginn nach dem Zusammenbruch - 1945 bis 1949. Braunschweig: Archiv Verlag, Track 1.

Mann, Thomas (1949). Ansprache zum Goethe-Jahr in der Frankforter Paulskirche (25.07.1949). In: Stiftung deutsches Rundfunkarchiv (2007). *Tondokumente zur deutschen Geschichte*. Emigration und Rückkehr in die fremd gewordene Heimat. Braunschweig: Archiv Verlag, Track 5.

Mann, Thomas (1950). Vortrag "Meine Zeit" anlässlich seines 75. Geburtstags (21.05.50). In: Stiftung deutsches Rundfunkarchiv (2008). *Tondokumente zur deutschen Geschichte*. Große Reden - die Macht des Wortes. Braunschweig: Archiv Verlag, Track 3.

Pierrehumbert, Janet (1994). Prosodic effects on glottal allophones. In: Fujimura, O. & M. Hirano (eds). *Vocal Fold Physiology 8: Voice Quality Control*. San Diego: Singular Press, 39-60.

Pierrehumbert, Janet & David Talkin, (1991). Lenition of /h/ and glottal stop. *Papers in Laboratory Phonology II*. Cambridge UK: Cambridge Univ. Press, 90-117.

Redi, Laura & Stefanie Shattuck-Hufnagel (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29, 407-429.

Rodgers, Jonathan *(1999).* Three influences on glottalization in read and spontaneous German speech. *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK)* 34, 173-280

Shank, Scott & Ian Wilson (2000). An acoustic analysis of vowel formants in pharyngeal and glottal contexts in Nuu-chah-nulth. *University of Washington Working Papers in Linguistics* 19, 75-84.

Stiftung deutsches Rundfunkarchiv (2004ff.). *Tondokumente zur deutschen Geschichte*. Braunschweig: Archiv Verlag [continuously enlarged series of audio CDs].

Umeda, Noriko (1978). Occurrence of glottal stops in fluent speech. *JASA* 64, 81-94.

Weizsäcker, Richard von (1984). Weihnachtsansprache (24.12.1984). In: Stiftung deutsches Rundfunkarchiv (2007). *Tondokumente zur deutschen Geschichte*. Die Bundespräsidenten - Weihnachts- und Silvesteransprachen. Braunschweig: Archiv Verlag, Track 11.

Weizsäcker, Richard von (1989). Weihnachtsansprache (24.12.1989). In: Stiftung deutsches Rundfunkarchiv (2007). *Tondokumente zur deutschen Geschichte*. Die Bundespräsidenten - Weihnachts- und Silvesteransprachen. Braunschweig: Archiv Verlag, Track 12.

Weizsäcker, R. von (1992). Weihnachtsansprache (24.12.1992). In: Stiftung deutsches Rundfunkarchiv (2007). *Tondokumente zur deutschen Geschichte*. Die Bundespräsidenten - Weihnachts- und Silvesteransprachen. Braunschweig: Archiv Verlag, Track 13.

Wiese, Richard (2000). *The Phonology of German*. Oxford: Oxford University Pres

Articulatory and Acoustic Inter-Speaker Variability in the Production of German Vowels

Melanie Weirich
Zentrum für Allgemeine Sprachwissenschaft, Berlin

This study examines articulatory and acoustic inter-speaker variability in the production of the German vowels /i/, /u/ and /a/. Our subjects are 3 monozygotic twin pairs (2 female and 1 male pair) and 2 dizygotic female twin pairs. All of them were born, raised and are still living in Berlin and see their twin brother or sister regularly. We assume that monozygotic twins that are genetically identical and share the same physiology should be more similar in their articulation than dizygotic twins but that the shared time and social environment of twins, regardless of their genetic similarity, also plays a crucial role in the acoustic similarity of twins. Articulatory measurements were made with EMA (Electromagnetic Articulography) and the target positions of the produced vowels were analyzed. Additionally, the formants F1-F4 of each vowel were measured and compared within the twin pairs. Our data seems to point out the importance of a shared environment and the strong influence of learning over the anatomical identity of the monozygotic twins regarding the production of vowels. But, additional results suggest (1) the impact of physiology on the production of a vowel following a velar consonant and (2) the interaction of physiology and stress in inter-speaker variability.

---

1 Introduction

Acoustic and articulatory variability are essential topics and crucial parameters in the field of speech production, and inter- and intra-speaker variability is one of the hallmarks of communication. To answer the question how intra- and inter-speaker variability is influenced by different factors is one aim of this paper. At this, the non-linear relation between articulation and acoustics (Stevens, 1972) has to be kept in mind, since differences in articulation do not necessarily result in differences in the acoustic output. Additionally, articulatory variability can occur in terms of motor equivalence, since some speech parameters can be obtained by different alternatives in articulation and demonstrate "the capacity of a motor system to achieve the same end product with considerable variation in the individual components that contribute to that output" (Hughes and Abbs, 1976, 199). The same acoustic output necessary for the vowel /u/, i.e. low second formant frequencies, can be achieved by rounding the lips, lowering the larynx or moving the tongue backwards (Perkell et al., 1993, Savariaux et al., 1995).

In many research fields regarding the variability of certain human properties the impact of *nature* and *nurture* is discussed. However, what influence *nature* on the one hand and *nurture* on the other hand have on the acoustics and articulation in speech production, and how these determinants interact in terms of inter- and intra-speaker variability is less clear. The factors *nature* and *nurture* can be described and specified as biological determinants (i.e. genetics, physiology, biomechanics) and non-biological determinants (i.e. social environment, learning, linguistic factors). Biology and hence physiological and biomechanical factors play an important role in terms of motor control and articulatory targets in speech production (Lindblom, 1983). The factor learning and social environmental influences come to the fore when we speak of auditory goals (Perkell et al. 1997).

The question we are addressing is what roles these two different influence sources play in terms of variability in speech production. Thus, when do physiological constraints prevail the impact of our social environment, and when does as a result biology have an impact on articulatory variability? And on the other hand, in which cases are we free in choosing different articulatory strategies to reach an auditory goal, and nurture seems to be the determining factor?

In psychology, learning theories like the social learning theory of Albert Bandura (1977) emphasize that people in general learn by observing. In terms of language acquisition it can be assumed that children learn the syntactical structures, phonological rules and lexical parameters of a language through imitation of the people surrounding and talking to them. Moreover, also dialectal pronunciation and socio-linguistic parameters of the parents are observed and absorbed by the child. Therefore, social environmental factors play an important role in speech production. Speech acquisition has to be at least partly independent of individual differences in the physiology of the vocal apparatus as it is in general possible to learn and speak any existing language for a child that is young enough and does not reveal any speech, language or hearing impairment. In exemplar-based models (Pierrehumbert, 2001, Johnson, 1997), social environmental influences are also seen to be crucial, since it is assumed that more recently encountered utterances are stored with higher activation levels than older utterances. Hence, socio-linguistic variation may be explained partly by a change in nurture. Thus, some non-biological determinants of variability in speech could be shown, namely: social environment, observing and learning.

On the other hand, it cannot be denied that speech has a biological foundation. For example, the shape of the vocal tract influences its filter characteristics and therefore, speakers differ in their formant values (Fant, 1960). The speech signals we are able to produce and perceive are limited by our physiology (Fuchs et al., 2007). Regarding the economy of speech gestures, Lindblom (1983, 217) assumes that "languages tend to evolve sound patterns that can be seen as adaptations to biological constraints of speech production". A connection between biology and differences in articulatory token-to-token variability was found by Shiller et al. (2002). Physiological parameters seem to constrain the capability of our speech motor system. In their study an asymmetric relationship between jaw stiffness and kinematic variability could be shown: higher stiffness values were observed along the axis of jaw protrusion and retraction and go hand in hand with reduced kinematic variability. In addition, for high jaw positions, stiffness was greater and kinematic variability smaller. Perkell (1985) found in his study higher articulatory precision in the positioning of dorsal tongue points near the place of maximal constriction for /i/ and /a/ in a direction perpendicular to the vocal tract midline than in direction parallel to the midline and hence, supporting evidence of the physiological "saturation effects". Furthermore, studies by Mooshammer et al. (2004) and Brunner et al. (2005) have shown that variability appeared to be less when the amount of linguo-palatal contact was high, suggesting high biomechanical restrictions in the production of high vowels. In addition, velar stops are described to be strongly influenced by anatomical and physiological constrictions. Perrier (2003) investigated looping movements of the tongue during VCV sequences and showed in their study that the production of the consonant - which was a velar stop - is influenced by the vowels. They assume that biomechanical factors are at least partly responsible for the resulting looping movements in the production of a velar stop. That anatomical properties, in particular the shape of the palate, could also play a role in inter-speaker variability was shown in a study by Brunner et al. (2009). The authors assume that speakers with a flat palate are more constrained in their articulatory variability, since small variation of the tongue position has a larger impact on the area function and henceforth on the acoustics than in speakers with a dome-shaped palate.

It has to be kept in mind that when we talk about speech, first of all we talk about communication. And speaker-specific variability can also be seen in the light of communicative demands. Communication is a two-sided process with a speaker on one side and a listener on the other. The aim of the speaker should be to be understood by the listener with the least effort possible (parsimony of the system). The aim of the listener is to receive the information the speech signal carries. The speech signal itself consists of different segments with different importance. Words under focus and stressed syllables are the most crucial parts of the coded transferred information. Therefore, it can be assumed that these segments are spoken with more effort, and reveal larger articulatory gestures that are longer in time (de Jong, 1997). It may also be assumed that these stressed syllables correspond to learned auditory goals, and that the unstressed syllables are generally shorter in time, more influenced by coarticulation processes, less articulatory distinct and more variable. Thus, it can be hypothesized that unstressed syllables are more sensible to physiological factors, more influenced by the individual vocal apparatus and less oriented towards the learned auditory targets. Similarly, certain prosodic boundaries are lengthened and potentially strengthened in comparison to others. For example, Fougeron and Keating (1997) point to the process of articulatory strengthening at the edges of prosodic domains: the higher the prosodic domain, the more linguo-palatal contact was found. As a consequence of this, a stronger influence of physiology and anatomical parameters can be assumed at higher prosodic boundaries with more linguo-palatal contact. Hence, when we talk about articulatory and acoustic variability in speech, the factor stress and the parameter prosodic domain have to be kept in mind.

Different factors that can have an impact on variability in speech have been shown: the influence of social environment, learning, physiology and stress patterns. One possibility to investigate the interaction and impact of these factors and to describe the influence of nature and nurture are twin studies.

1.1 Twin studies

In order to study the influence of biological parameters (physiology, biomechanics) on the one hand and non-biological parameters (learning, environmental factors) on the other hand, we investigated inter-speaker variability in the speech of monozygotic twins (MZ, who are genetically identical) and dizygotic twins (DZ, who share approximately 50% of their genes). For example, if we find high inter-speaker variability in a certain speech parameter within a MZ twin pair that lives apart from each other, the influence of genetics and physiology on this parameter seems to be rather small. We will discuss our results in terms of articulatory and auditory targets. If underlying articulatory targets are assumed in speech production, MZ twin pairs should be identical in their articulation (and hence also acoustics) independent of the time they spend together. If auditory goals are assumed, they should differ in their acoustical output when living apart from each other. In addition to that, when a

DZ pair that spends most of their time together is very similar in their acoustical outputs of a certain speech parameter, auditory goals and not articulatory targets must be crucial in this parameter.

Twin studies are a common method in the field of psychology. The two influence factors investigated are (1) identical genes (and physiology) and (2) a shared environment. The latter factor refers to social environmental factors that contribute to the resemblance between individuals who grow up in the same family. The *Equal Environments Assumption (EEA)* assumes that MZ and DZ twins share the same amount of environmentally caused similarity, as studies of mislabelled twin pairs have shown (Spinath, 2005).

Comparing the similarity of DZ and MZ twins within the field of speech research is rather new and less common. Still, some studies regarding perceptual and acoustic differences within twin pairs have been conducted (for an overview see Loakes, 2006). Perception experiments revealed the striking similarity between twins' voices but also showed that twin pairs can be identified above chance by people familiar with their voices (Whiteside & Rixon, 2001). The most frequently investigated speech parameter in twins' speech is fundamental frequency, and results point to a great influence of physiology on this parameter as MZ twins reveal higher correlations than DZ twins (Przybyla et al., 1992, Debruyne et al. 2002). Also, voice quality characteristics analyzed by Van Lierde et al. (2005) such as perceptual voice characteristics, vocal performances, and the overall vocal quality by means of the Dysphonia Severity Index were very similar in MZ twins. For shimmer and jitter only, no significant correlation coefficient could be obtained. Smaller influence of identical genes and physiology and greater impact of environmental factors were found by Debruyne et al. (2002) for the variation of speaking fundamental frequency, as MZ and DZ twins revealed the same amount of similarity. Nolan & Oh (1996) showed similarities in coarticulation patterns of /l/ and /r/ but also differences in pronunciation alternatives of /r/. Regarding the production of vowels, the study by Loakes (2004) point to lax front vowels (F2, F3 of /I/) to be most speaker-specific regarding twins. There are still only few studies regarding inter-speaker variability in twins, but there is a greater lack of articulatory studies within this field. We are not aware of any study investigating articulation patterns in the speech of twins, although interesting findings regarding the impact of physiology on inter-speaker variability can be obtained by this.

1.2 Pilot Study

We have conducted a pilot study to reveal acoustic differences in the speech parameters of identical (MZ) and non-identical (DZ) twins (Weirich, 2009). This study should help to locate phonemes that show acoustic differences within twin pairs and therefore are promising to show also differences in articulation, although the relation of acoustics and articulation is not linear. To optimize the probability of finding differences within twin pairs the speech material should show in general high inter-speaker variance but low intra-speaker variance. 4 identical twin pairs and 1 non-identical twin pair were recorded with a large set of sentences repeated 5 times. The results revealed the vowels /u/ and /a/ to be most speaker-specific within all twin pairs. A chi-square test showed a significant influence of place of articulation ($X^2 = 4{,}879$, df = 1, $p < .005$): the central and back vowels [ɑ, ɑː, ʊ, uː, oː] revealed more significant differences than the front vowels [ɪ, iː, ɛ, eː, ʏ, ɣː]. The formant with most differences among F1-F4 was F3, followed by F1 and F4. F2 showed the least variation within all twin pairs. A clear discrepancy in the amount of differences in the formants could be found between non-related persons (twins from different pairs) and related persons (twin pairs). For the unrelated pairs the first three formants showed a relatively stable probability of over 50% of showing differences, F4 indicates a somewhat smaller value (45%), whereas the twins revealed an average probability of only 28%. Comparing the MZ with the DZ twins, the DZ twin pair reveals only in F3 a higher probability of showing differences. The differences between inter-speaker variability within DZ and MZ pairs were biggest in F1. Interestingly, here, the MZ pairs showed a higher probability of showing differences. In general, these results point to the importance of a shared environment over physiological identity and support the hypothesis of auditory goals as targets in speech production regarding the acoustics of vowels. Furthermore, our pilot study revealed the impact of the factor stress on inter-speaker variability in plosives: it was more likely to find differences in VOT (voice onset time) and VDC (voicing during closure) within all twin pairs, but especially in MZ pairs in stressed syllables than in unstressed syllables. Since unstressed syllables are perceptually less important, they can be more variable and thus, they do not show significant differences. The stressed syllables are less variable and the learnt auditory goals are more crucial here. Another explanation could be that physiology has a greater influence on speech production in unstressed syllables than in stressed syllables. Of course, the results of the pilot study have to be treated very carefully, as the DZ twins are presented by only one pair.

1.3 Hypotheses

The aims of our twin study are to explore, (a) to what extent speaker specific variability is influenced by the speaker's respective articulatory constraints and (b) to what extent speakers are influenced by their social environment and yet free to aim for a given auditory goal. The purpose of the present articulatory and acoustic study is to verify and cross-check the results of the pilot study regarding acoustical differences and similarities in vowels and additionally to examine the underlying articulatory patterns in the speech of MZ and DZ twins.

Our hypotheses are that (1) DZ twin pairs need not naturally show more differences in their acoustic outputs regarding vowels, as they adjust their speech production to each other and use auditory goals to produce vowels, but may reveal their anatomical and physiological differences in individual articulatory strategies, when this is possible. In addition to that, it is proposed that (2) the physiological influences of the tongue and the shape of the palate show a greater impact on the production of consonants (e.g. velar stops) and hence, also on vowels in their neighbourhood. Additionally, the factor stress is taken into account and it is hypothesized that (3) more inter-speaker variability can be found in stressed than in unstressed syllables within MZ twin pairs, mirroring the greater influence of physiology in the production of an unstressed syllable. Additionally, the DZ twins should reveal more differences in the unstressed syllables.

Therefore, an articulatory and acoustic analysis of inter-speaker variability within twin pairs was made of the vowel targets of /a/, /i/ and /u/ produced in a stressed syllable, and additionally of the vowel /i/ following a velar stop produced in the sequences /hagi/ (in a posttonic syllable) and /giba/ (in a tonic syllable).

2 Method

The presented study is part of a greater PhD project on acoustic and articulatory inter-speaker variability in the speech of twins. With respect to space limitation and for sake of simplicity, here, we will concentrate on our findings addressing vowels. Considerations regarding our speech material are based on earlier literature, studies that examined intra- and inter-speaker variability of speech parameters, and on the results of our pilot study. Because of the latter, the factor stress is taken into account. In the following section the experimental design, our speech material and subjects, and the acoustic and articulatory measurements that we conducted will be described.

2.1 Experimental set up

We investigated articulatory and acoustic inter-speaker variability within mono- and dizygotic twin pairs in the production of vowels. Articulatory recordings were carried out using a 2 D Electromagnetic Articulograph (Carstens AG 100). For the articulatory measurements two sensors, one at the upper incisors and one at the bridge of the nose, served as reference sensors to compensate for head movements. Three coils were glued to the tongue: one at the tongue tip, one at the tongue back and one between these two on the tongue dorsum. Another sensor was placed below the lower incisors in order to track jaw movements, and two further sensors were glued to the upper and the lower lip to record lip movements. As one aim of the study was to compare articulatory movements between speakers with a nearly identical physiology (within the MZ pairs), it was crucial to use the same positions for the coils on the tongue. Therefore, we measured the distances between the glued coils, took photos of the set up, i.e. the tongue with the glued coils on top and created a template for the first twin to use it as a reference for the second twin.

In addition to the articulatory recordings, the audio tracks were recorded for our speakers for further acoustic analysis (48 kHz sampling rate). After the recordings, the sensor of the tongue tip was removed and the contour of the palate was recorded by moving this sensor along the palate from back to front. This contour could be used afterwards to compare the shape of the palates within the twin pairs. Second, a silicone dental and palate cast was taken, and the shape and steepness of the palate was examined more closely. This was important to us to verify the assumption of identical physiology in MZ pairs. The articulatory data was preprocessed including correction algorithms for head movement, filtering of the data (low pass filter: bandwidth of 18 Hz with a damping of 50 dB at 52 Hz), rotation and translation of the position data and synchronization with the acoustic data.

2.2 Speech material

Our speech material consists of the stressed vowels /a/, /i/ and /u/ in a non-focused position. They were embedded in the verbs [‘li:bə] (to love), [‘vaʃə] (to wash) and [‘zu:xə] (to search for) which were part of the carrier sentences. In that way we could increase the amount of renditions of the vowels: each speaker repeated the target vowels 40 times in different carrier sentences presented to them on a monitor. For each subject we examined articulatory targets (tongue positions) and acoustic targets (F1-F4) and explored the inter-speaker variability within the twin pairs.

As it is hypothesized that physiology has a stronger influence on consonants than on vowels and especially on velar stops, we conducted a second analysis. We investigated the influence of a velar stop on the production and inter-speaker variability within twins of the following vowel /i/. Therefore, the non-words (names) /‘hagi/ and /‘giba/ that were presented in the carrier sentences were used to investigate articulatory and acoustic inter-speaker variability of /i/ in the syllable /gi/ in contrast to the analysis of the variability in the syllable /li/.

Additionally, the factor stress was taken into account to cross-check the results of our pilot study. We looked for an influence of stress on inter-speaker variability in vowels. Therefore, the production and variability of /i/ in a stressed syllable (/'giba/) and in an unstressed syllable (/'hagi/) were compared. Table 1 gives an overview of our speech material. The number of repetitions (N) differs slightly between the speakers due to some articulatory data that had to be excluded from the analysis, and therefore, mean values and standard deviations of the analyzed repetitions are given in the table. Moreover, according to our experimental set-up and time restrictions during the EMA-recordings, the vowels /a/, /i/ and /u/ were iterated approximately 40 times, while the vowel /i/ in the syllable /gi/ was iterated just 9 times in each stress condition. Note also that in contrast to the verbs containing the vowel renditions, the target words /giba/ and /hagi/ were in a focused position.

Table 1: Speech material

vowel	stress condition	target word		∅ N (std.dev.)
a	stressed	w**a**sche	/'vaʃə/	37.3 (3.1)
i:	stressed	l**ie**be	/'li:bə/	44.1 (5.2)
u:	stressed	s**u**che	/'zu:xə/	38.3 (2.6)
i:	stressed	G**i**ba	/'gi:ba/	9.3 (2.0)
i:	unstressed	Hag**i**	/'hagi:/	8.8 (1.2)

2.3 Subjects

Our subjects are 3 monozygotic twin pairs (2 female, 1 male) and 2 dizygotic twin pairs (both female) between 20 and 34 years. The genetic similarity (zygosity) of these twins was determined by a genetic laboratory by their genotypic comparisons based upon 16 different genetic markers[1]. We suppose that the physiological and biomechanical properties of the vocal apparatus are rather similar in the former and different in the latter. To verify this assumption, the shape of the palate was examined and compared within the pairs. The palate contours that we recorded after the experiment were rotated and adjusted within the respective siblings, as they could differ in their sitting position and posture of the head. This was done by matching the highest point of the palate of one twin to his sibling and then rotating the palate until it fitted.

[1] Monozygotic twin pairs are genetically identical. If a twin pair differs for any DNA marker, they must be dizygotic. When a reasonable number of markers reveal no differences, it can be concluded that the twin pair is monozygotic (Spinath, 2005).

All twin pairs were born, raised and are still living in Berlin, Germany, but differ with respect to the amount of time they recently spent together. The frequency of spending time together will additionally be considered as a factor influencing inter-speaker variability since the mutual influence of the twins and their shared social environment could play a role in shaping auditory goals. Another possible factor is the attitude towards being a twin. Therefore a separate interview with each subject was conducted and the twins made ratings on a 5 point Likert-scale from 1 (I don't like being a twin) to 5 (I like very much being a twin). The number 3 served as a neutral position with no positive or negative feelings about being a twin. All subjects showed a strong, positive attitude, and except for one pair concurred with their sibling. An overview of the characteristics of our subjects is given in Table 2.

Table 2: The twin pairs with the factors "shared time" and "attitude towards being a twin"

twin	zygosity	sex	age	Frequency of spending time together	Attitude towards being a twin (1-5)
SLCL	MZ	m	32	Twice a month	5 – 5
AFHF	MZ	f	34	Nearly every day	5 – 5
GSRS	MZ	f	26	Live together	5 – 5
LRSR	DZ	f	20	Live together	4 – 5
MGTG	DZ	f	20	Live together	5 – 5

2.4 Acoustic measurements

The vowels /a/, /i/ and /u/ in the target words mentioned in Table 2 were segmented and annotated in PRAAT and the formants F1-F4 were measured semi-automatically in the middle of the segmented interval in PRAAT with a positive time step of 0.01, a maximum number of 5 formants, a maximum formant value of 5500 Hz (for female) and 5000 Hz (for male), a window length of 0.025s and a real pre-emphasis from 50Hz. Figure 2 gives an example of a labelled vowel. Each measured formant value of every analyzed vowel was checked manually and corrected if necessary.

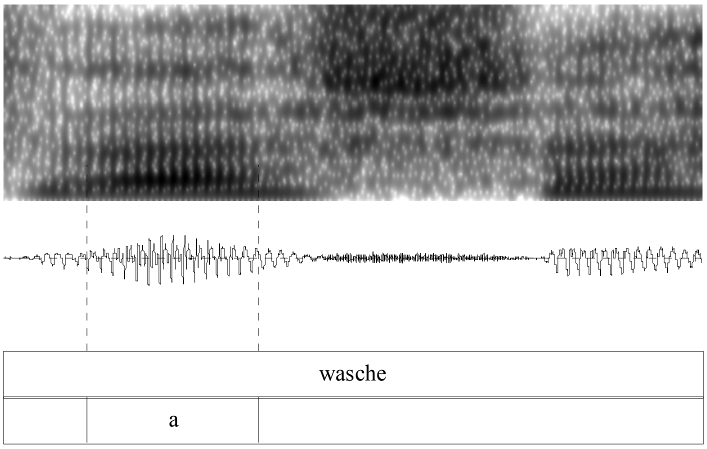


Figure 2: Example of a segmented and annotated vowel in Praat

For each twin pair, mean formant values were compared and the inter-speaker variability was examined. Scatterplots and dispersion ellipses (2 standard deviations) of F1-F2 variation were conducted for each subject and vowel using MATLAB 7.4.0. This was done by a principal component analysis with 2 main components: the highest amount of variability served to define the direction and length of the first axis, the second axis was perpendicular to the first. After that, statistical analyses in R (version 2.8.1) were done to find significant differences in mean formant values within the pairs. First we conducted an ANOVA over all speakers with the respective *formant* as dependent and *speaker* as the independent factor to look for a general effect of the speaker. After that we looked for differences *between the speakers of the same pair* with a Tukey Post Hoc test. The Post Hoc test has advantages over a normal t-test, since the Tukey Test adjusts the results to the amount of t-tests that are made.

2.5 Articulatory measurements

For each speaker and vowel, articulatory target positions were measured. In general, this is the point in time (and the position of an articulator) when the tongue or the jaw has reached a certain extreme position or maximum and the movement direction changes after reaching this target position. For each vowel a particular articulator was chosen to define the achievement of this articulatory target. Reaching the target position was assumed when the velocity of this particular articulator was minimal. The chosen parameters can be found in the following Table 3.

Table 3: Parameters defining articulatory targets

vowel	Parameter
/a/	Lowest vertical position and minimum in velocity of the jaw
/i/	Highest vertical position and minimum in velocity of the tongue dorsum
/u/	Lowest horizontal position (= maximal protrusion) and minimum in velocity of the upper and lower lip

Articulatory target positions of /a/ were easy to determine and the point of maximal jaw opening was in most cases congruent with the lowest horizontal position of all tongue coils. For the realizations of the vowel /i/ the tongue dorsum coil was significant, but less distinct, as the tongue was already up because of the preceding /l/. Still, target positions for /i/ could be defined certainly in most cases. Defining an articulatory target position of /u/ in the target word /‘zu:xə/ revealed to be the most difficult determination. Often, no minimum in the velocity of the tongue coils could be found, and therefore, the upper and lower lip and the orientation on the acoustical envelop were taken into account.

Again, scatterplots were made in MATLAB to visualize the articulatory targets of each subject, and vowel and statistical tests in R (ANOVA and Tukey Post Hoc Tests) were conducted to look for inter-speaker variability within the twin pairs in the target positions of each vowel.

3 Results

In the following section we will report our results addressing the analysis of the palates, acoustic and articulatory inter-speaker variability within the twin pairs, and the two factors *stress* and *velar stop* will be taken into account.

3.1 Similarities of the palates

Figure 1 shows the adjusted palate contours of the twin pairs (midsagittal tracing, face to the left). The vertical line in each graph marks the highest point of the palate, which was taken as a reference for the adjustments. The horizontal lines under the graphs indicate the lengths of the palates.

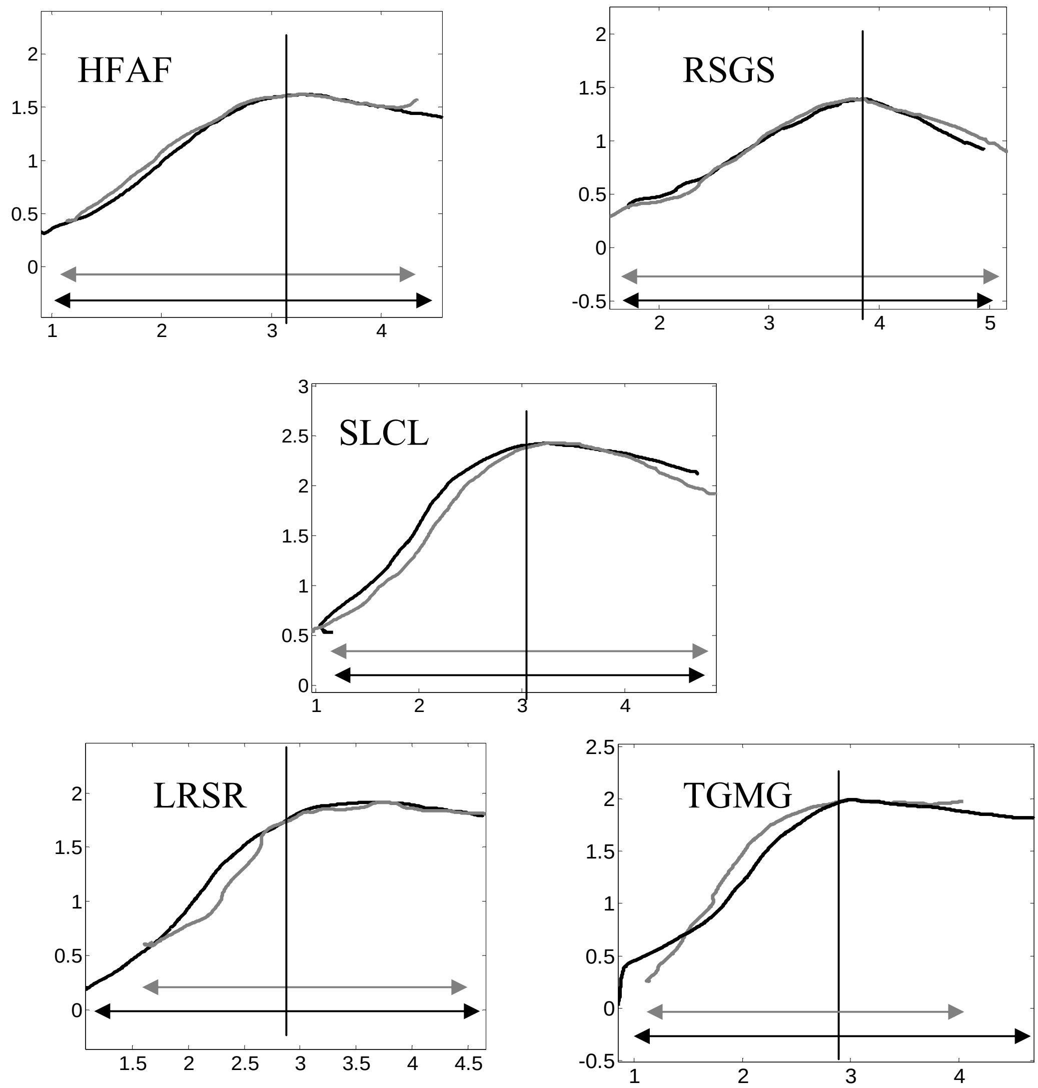


Figure 1: Palatal contours of each twin pair (generated in MATLAB), axis scales in cm, different colors show different subjects.

Hereby, the hypothesis of a more similar palatal shape of monozygotic twins than of dizygotic twins could be supported. The figure reveals the outstanding similarity of the palate contour of the 2 MZ pairs HFAF and RSGS. In contrast, the palate contours of the DZ pairs LRSR and TGMG are less similar and differ in shape and steepness. At a first glance the palate contours of the MZ pair SLCL also show some differences. But when you look at the size dimensions of the palate, both MZ twins reveal a remarkable high palate (2cm) and are identical in the distance from the highest point of the palate to the beginning of the incisors, whereas both DZ twins vary in these distances (cf. the different lengths of the lines under the respective figures). Especially the DZ pair LRSR varies in the distance from the highest point of the palate to the incisors and LR (grey) reveals a much smaller palate than here sister SR (black).

3.2 The vowel targets /a/, /i/ and /u/

Results regarding our first hypothesis, that is, differences in the formants F1-F4 and the articulatory targets of the vowels /a/, /i/ and /u/ within MZ and DZ twin pairs, are given in the following sections.

3.2.1 Acoustics

To get a first impression of the vowel spaces of each subject and the twin pairs in particular, F1/F2 scatterplots for each pair and the 3 vowels were conducted. The following Figures 2a and 2b show the scatterplots for the 5 twin pairs (Figure 2a shows the 3 MZ pairs and Figure 2b the 2 DZ pairs). Each measured F1/F2 value is marked by a single dot. Ellipses were calculated and drawn to visualize the intra-speaker variability of each vowel. The two colors (black and grey) distinguish the two speakers of a twin pair.

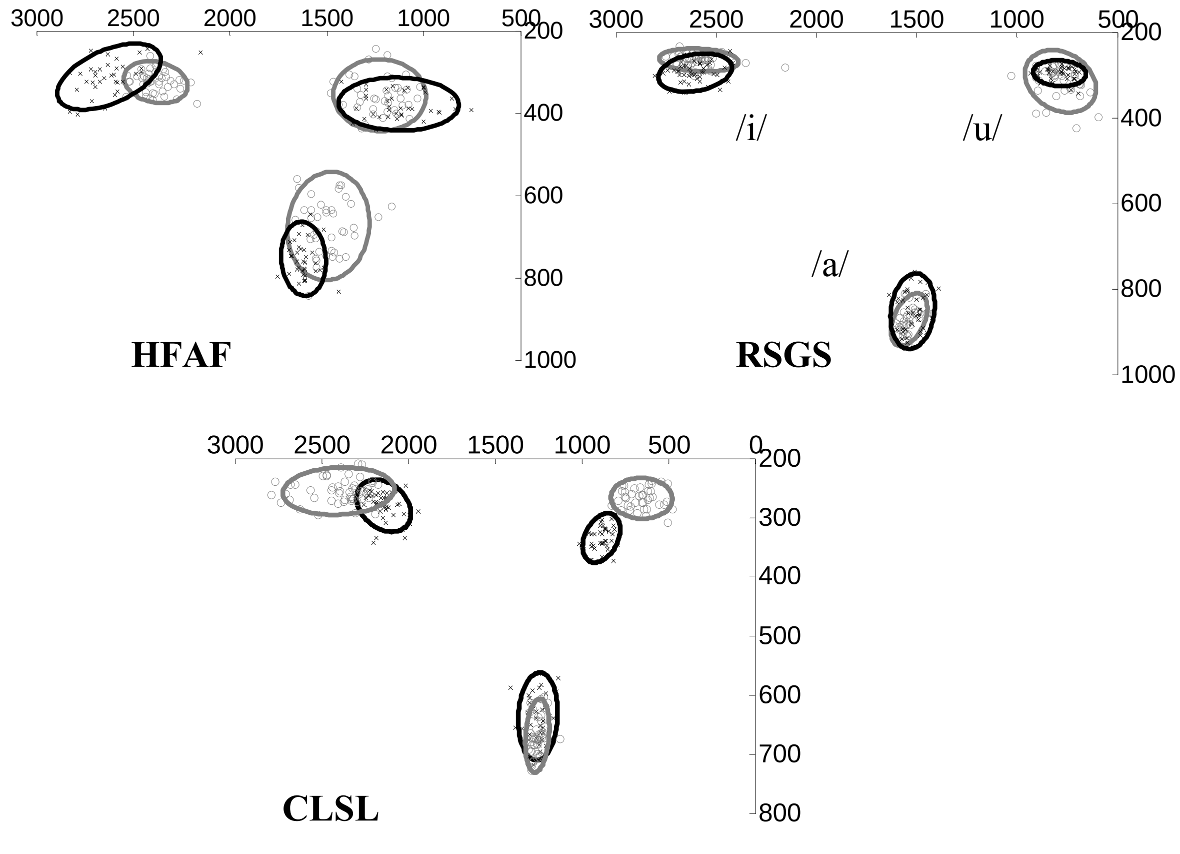


Figure 2a: Scatterplots of F1 (negative y-axis) and F2 (negative x-axis) for the female MZ pairs (HFAF, RSGS) and the male MZ pair (CLSL) and the vowels [a],[i:],[u:]

Figure 2a shows that the vowel spaces of the MZ pair RSGS are most similar; here, the ellipses overlap nearly 100%. For the MZ pairs SLCL and HFAF, differences within the pairs can be assumed in the mean formant values of /i/

and /a/ for HFAF, and /i/ and /u/ for SLCL, but similarities in the sizes of the ellipses and hence, the intra-speaker variability is strikingly apparent: Variability in F1 and F2 is small for SLCL and GSRS, but considerable for HFAF.

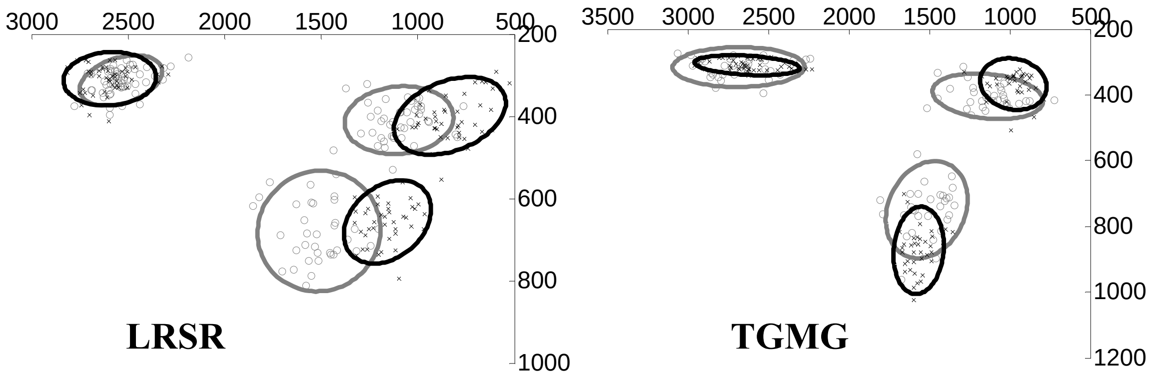


Figure 2b: Scatterplots of F1 and F2 for the female DZ pairs and the vowels /a/,/i/,/u/

In the scatterplots of the DZ twins (fig. 2b), the F1-F2-ellipses of /i/ are most similar within the pairs. For /a/ and /u/, differences can be seen, but especially for LRSR in F2. As was shown in figure 1 before (chapter 2.3), the DZ pair LRSR revealed differences in the sizes of the palates and it can be assumed that the different sizes of the vowel spaces are influenced by these anatomical differences: LR (grey), who showed a much smaller palate also reveals a smaller vowel space than her sister SR (black).

To look for statistically significant differences in formants within the twins, mean formant values of F1 – F4 of the three vowels were measured of each subject and compared with the corresponding sibling. Table 4 shows the significant differences found between speakers of the same twin pair. The MZ pairs showed on average 5.3 differences in F1-F4 of the three vowels, the DZ pairs 7 (of 12 (3 vowels x 4 formants) possible differences), but an ANOVA revealed no significant effect of the factor *zygosity*. Within the MZ pairs, the least inter-speaker variability in formants was found for the twin pair that shares genetics as well as environment (GSRS, as indicated in the scatterplots before), and the most differences were found for the MZ pair which sees each other only twice a month (CLSL). Concerning the amount of differences, the DZ pair LRSR (that lives together) even comes before this MZ pair. As in the pilot test, the results point to the a shared environment as the greatest impact factor on the acoustics of stressed vowels.

An influence of vowel height on acoustical variability as assumed in the introduction could not be found in terms of less inter-speaker variation in /i/ due to the strong influence of physiology on the production of this vowel. In fact, the MZ twins showed most differences in the formants of /i/ and least in /u/, whereby inter-speaker variability within the DZ twins was biggest in /u/ and smallest in /i/. A possible explanation could be that our investigated speech material consisted of stressed vowels and physiology seems to play a more important role in unstressed vowels.

Table 4: Significant differences in F1-F4 within the twin pairs of /a/, /i/, /u/ (Tukey Post Hoc Test in R, Significance Level <.01)

Twin pair	**zygosity**	**/u/**	**/a/**	**/i/**	**Amount of differences Total /12**	**% of F1/F2**
GSRS	MZ	F3	F4	F1, F4	4/12	25 %
AFHF	MZ	-	F1, F2, F3	F2, F4	5/12	60 %
SLCL	MZ	F1, F2	F3	F1, F2, F3, F4	7/12	57 %
SRLR	DZ	F2,F3, F4	F2, F4	F4	6/12	33 %
TGMG	DZ	F1,F2, F3, F4	F1, F3, F4	F3	8/12	37 %

For two of the three MZ pairs the inter-speaker variability is quite equally distributed between F1/F2 on the one hand and F3/F4 on the other, it is even higher for the lower formants. In contrast to this, within the DZ twins, F1 and F2 account for only approximately 35 % of the differences. Since size and form of the vocal tract have a strong influence on the speaker specific higher formants of a speaker, the MZ twins are expected to show less inter-speaker variability in F3 and F4 than in F1 and F2. Our results confirm the assumption that the higher formants are more dependent on physiological factors and less affected by the speakers' alternative articulation strategies than the lower formants, since the DZ twins show here more variability than the MZ twins. As indicated in the scatterplot of the DZ pair LRSR (see Figure 2b), there were no differences at all for F1, pointing again to a small impact of physiology on this formant[2].

3.2.2 Articulation

Table 5 gives an overview of the articulatory inter-speaker variability within the pairs in the production of /a/, /i/ and /u/. As described before, different coils were used as a reference for determining the articulatory targets of the vowels (cf. tab. 3). To compare the articulatory targets between the speakers we analyzed the positions of the tongue at the point in time of the determined targets. We are concentrating now on the positions of the tongue dorsum coil for

[2] Note that differences in F1 can also appear because of differences in loudness and hence a bigger jaw opening. Since we did not yet check for similar amplitudes in the speech signals, this has to be kept in mind for further experiments and assumptions.

/i/ and the tongue back coil for /u/ and /a/. The positions in a horizontal (X) and vertical (Y) direction of these coils at the point in time of the determined target positions were compared within the twins, and results with significant differences are given in the following table.

Table 5: Differences in target tongue positions (in cm) of the three vowels within the twin pairs, significant differences ($p < .01$) in bold.

Twin pair	zygosity	coil position	/a/ (tback)	/i/ (tdorsum)	/u/ (tback)
GSRS	MZ	vertical (Y)	**0.223**	0.029	**0.235**
		horizontal (X)	0.049	**0.168**	0.006
AFHF	MZ	vertical (Y)	**0.445**	0.044	**0.114**
		horizontal (X)	**0.250**	**0.193**	**0.165**
SLCL	MZ	vertical (Y)	**0.743**	**0.171**	**0.176**
		horizontal (X)	**0.226**	**0.428**	**0.326**
SRLR	DZ	vertical (Y)	**0.491**	0.011	**0.591**
		horizontal (X)	**0.235**	**0.126**	0.035
TGMG	DZ	vertical (Y)	0.057	0.077	0.049
		horizontal (X)	**0.824**	**0.469**	**0.583**

In general, most similarities in the target positions within the pairs were found for the vowel /i/ and least for the vowel /a/. This could be explained in terms of differences in intra-speaker variability (see Brunner et al. 2005, Mooshammer et al. 2004), with high vowels showing less articulatory variability than low vowels. If more articulatory intra-speaker variation can be expected for /a/ than for /i/, less inter-speaker variation could be assumed in siblings with similar palate shapes. In contrast to the similarities in the articulation in /i/, in the acoustical analysis, many differences were found for /i/. This points to the fact that in higher vowels less articulatory variance is necessary to achieve differences in the acoustical output. The pairs with the least inter-speaker variability in all vowels were the MZ pair GSRS and the DZ pair TGMG (but note that the differences in the vertical position of the tongue dorsum of /i/ of TGMG reached significance with $p < 0.02$). Most differences were revealed by the MZ pair SLCL. This seems to point to a minor role of similar physiology in our investigated vowels and again to the influence of a shared environment (as SLCL is the pair which spends least time together). A possible explanation could also be the factor stress, as all vowels were produced within a stressed syllable, and the effect of physiology is suggested to be stronger in unstressed syllables, and stressed syllables are supposed to be less variable and therefore more sensitive to differences. Also, the high palate of SLCL could lead to greater intra-speaker variability in general and hence, a greater variability between the two speakers.

3.3 Influence and interaction of the factors stress and velar stop

The next section contains our results regarding our second and third hypothesis , i.e. the factor stress and the influence of a velar stop on the articulatory and acoustic inter-speaker variability of the vowel /i/ are discussed. Additionally to the production of /i/ in the stressed syllable /li/, the vowel is analyzed and compared in the syllable /gi/ in a stressed and in an unstressed position.

3.3.1 Acoustics

Note that due to our speech material, the amount of repetitions differ between the three conditions. For the formants of /i/ in the stressed and unstressed syllable /gi/ only 9 repetitions per condition (on average) could be taken into account. The mean formants of /i/ in /liebe/ could be investigated in approximately 40 repetitions. Therefore, different sample sizes were used for the Tukey Post Hoc Test, and it has to be considered that these variations influence the probability of finding significant differences. Tests with larger sample sizes are more reliable, and it is more probable to find significance on a lower p-level. Thus, in the interpretation of the results this difference has to be kept in mind.

Table 6: Significant differences in formants of /i/ within the twin pairs for the three conditions: /i/ produced in the unstressed syllable /gi/, in the stressed syllable /gi/, in the stressed syllable /li/ (Tukey Post Hoc Test in R, Significance Level <.01)

Twin pair	Zygosity	Stressed /i/ in /liebe/	Stressed /i/ in /giba/	Unstressed /i/ in /hagi/
GSRS	MZ	F1, F4	F1	-
AFHF	MZ	F2, F4	F1	-
SLCL	MZ	F1, F2, F3, F4	F3 (.011)	F2, F3
SRLR	DZ	F4	-	F1
TGMG	DZ	F3	F1	F1, F2 (.012)

Comparing the groups of the MZ twins with the DZ twins, it is noteworthy that there is a clear favour of significant differences in the formants of /i/ produced in the stressed syllable /li/ for the MZ twins, but a quite equally distributed amount of differences in formants for all three conditions for the DZ twins. As we noted before, more differences were expected for the stressed /li/ condition because of the bigger sample size. In spite of this fact, the DZ twins revealed more differences in F1 and F2 in the /gi/ conditions, mirroring the stronger influence on physiology in the first two formants of a vowel following a velar consonant.

By comparing the two /gi/ conditions, results support our hypothesis of an interaction of physiology and the factor stress: The MZ twins GSRS and AFHF that revealed a strikingly similar palatal contour (see Figure 1) show no differences in the unstressed /gi/ condition, but in the stressed condition (F1). Both DZ twin pairs reveal more inter-speaker variability in the unstressed than in the stressed syllable, pointing to auditory goals to be crucial in stressed vowels.

As was said before, it has to be considered that due to the larger sample size, the overall probability of detecting significant differences is greater in the /li/ condition than in both /gi/ conditions. Moreover, the /i/ in /liebe/ is in a non-focused position, whereas the /i/ in /hagi/ and /giba/ is part of a word under focus. Nevertheless, here, we are concentrating on a comparison between inter-speaker variability of MZ and DZ pairs and not on one between the three conditions for all speakers, and thus, the requirements are equally balanced and comparable.

3.3.2 Articulation

Table 7 shows the results of the articulatory inter-speaker variability within the twin pairs in the production of /i/ in the three conditions: in the unstressed syllable /gi/, in the stressed syllable /gi/ and in the stressed syllable /li/. We will focus on the target tongue positions, and therefore, the horizontal and vertical position of the tongue dorsum coil, to compare the articulation between the speakers.

Table 7: Differences in target tongue positions (in cm) of the vowel /i/ within the twin pairs, significant differences ($p < .01$) in bold.

twin	zygosity	stress	Tongue dorsum Y	Tongue dorsum X
AFHF	MZ	Unstressed /gi/	0.0628	0.1618
		Stressed /gi/	0.0404	**0.4844**
		Stressed /li/	0.0446	**0.1931**
GSRS	MZ	Unstressed /gi/	0.1516	0.0763
		Stressed /gi/	0.0452	0.1189
		Stressed /li/	0.0290	**0.1677**
CLSL	MZ	Unstressed /gi/	0.1403	0.1887
		Stressed /gi/	**0.2351**	**0.4352**
		Stressed /li/	**0.1712**	**0.4280**
SRLR	DZ	Unstressed /gi/	0.1510	0.0092

		Stressed /gi/	**0.1312**	**0.2342**
		Stressed /li/	0.0383	**0.2590**
MGTG	DZ	Unstressed /gi/	0.0628	**0.3254**
		Stressed /gi/	**0.1699**	**0.3013**
		Stressed /li/	0.0728	**0.4700**

From the results given in Table 7, it can be said that the factor stress has an impact on the articulatory inter-speaker variability in the production of /i/ within twin pairs. None of the MZ pairs shows significant differences in their target positions of the unstressed /i/ (p <. 01). Of the two DZ pairs, one pair reveals significant differences in the horizontal position of the tongue dorsum. For the stressed condition, each pair revealed at least one significant difference.

Interestingly, for the DZ pairs the differences were more common in the stressed syllable /gi/ than in the stressed syllable /li/ (even though the sample size would favour significance in the /li/ condition). For the MZ pairs, there was either no difference in variability (AFHF, CLSL) or more differences in the syllable /li/ (GSRS) were found. This could be interpreted in terms of a stronger influence of physiology and biomechanics on articulation in the production of the vowel /i/ following a velar consonant (in the syllable /gi/).

4 Discussion

In the outline of the paper, 3 hypotheses were formulated. The first one addressed the great influence of a shared environment over the identical physiology of MZ twins in general, and assumed that:

(1) DZ pairs need not naturally show more differences in their acoustic outputs regarding vowels, as they adjust their speech to each other and use auditory goals to produce vowels, but may reveal their anatomical and physiological differences in individual articulatory strategies, when this is possible.

Our results support this hypothesis, as no differences in the amount of acoustic inter-speaker variability in the formants of the stressed vowels /i/, /a/, and /u/ between MZ and DZ twin pairs were found. The MZ pair that lives together revealed the least differences regarding formants, but one DZ pair that also lives together showed less acoustic inter-speaker variability than the MZ pair that sees each other only twice a month. Nothing can be said about individual articulatory strategies. Of course, the validity of the hypothesis is limited by our speech material (i.e. the vowels /i/, /a/, /u/ in a stressed position), and due to the time-consuming articulatory recording, the relatively small group of speakers, or pairs (3 MZ and 2 DZ twin pairs), respectively. Additionally, acoustic results may indicate that in stressed vowels MZ and DZ twins differ in their amount of variability in the higher formants. MZ pairs were less likely to show differences in F3/F4 than the DZ pairs, mirroring a stronger influence of physiology on higher formants.

To look more closely at the possible influence factors on acoustic and articulatory inter-speaker variability within twins, two more hypotheses were formulated and investigated:

(2) The physiological influences of the tongue and the shape of the palate show a greater impact on the production of consonants (especially velar stops) and hence, also on vowels in their neighbourhood.

Results suggest that articulatory and acoustic inter-speaker variability of the vowel /i/ seems to be influenced by the production of a preceding velar stop: MZ twins, who are supposed to have identical anatomy and physiology of the vocal apparatus, are more similar in their articulatory targets (vertical and horizontal tongue positions) and acoustic outputs (formants, especially F1 and F2) of the vowel /i/ when a velar consonant precedes the vowel. DZ twins that show differences in shape and steepness of the palate, reveal more inter-speaker variability in their articulatory targets and acoustic outputs of /i/ produced in the syllable /gi/ than in the syllable /li/. This result points to a stronger influence of physiology on the production of /i/ following a velar stop.

As our pilot study revealed that stress could be a possible influence factor on inter-speaker variability, our third hypothesis assumed that:

(3) More inter-speaker variability can be found in stressed than in unstressed syllables within MZ twin pairs, mirroring the greater influence of physiology on the production of an unstressed syllable.

We found supporting evidence that there is an interaction between physiology and the factor stress: Physiology seems to have a stronger influence on the production of /i/ when produced in an unstressed syllable. Both DZ twins revealed more differences in formants in the unstressed condition, and the 2 female MZ twins with the remarkable similar palatal shape showed more differences in formants in the stressed condition. In their articulatory targets the MZ pairs revealed no inter-speaker variability in the unstressed condition, but one of the DZ twins did.

To sum up, it can be said that a shared environment plays a very important role in inter-speaker variability in vowels. But, there are several factors that contribute to this variability and intensify the impact of the identical physiology of the vocal apparatus of MZ twins, namely, the production of a velar consonant preceding the vowel and the factor stress.

Acknowledgments

I would like to thank Jörg Dreyer for helping carrying out the 2D EMA recordings. Many thanks to Susanne Fuchs, Jana Brunner, Stefanie Jannedy and Bernd Pompino-Marschall for valuable comments on this paper. Thanks also to Jana Brunner for providing helpful scripts. This research has been supported by the Federal Ministry of Education and Research.

References

Bandura, Albert (1977). *Social Learning Theory***.** Prentice-Hall, Englewood Cliffs, N.J.

Brunner, Jana, Susanne Fuchs, & Pascal Perrier (2005). The influence of the palate shape on articulatory token-to-token variability. ZAS *Papers in Linguistics 42*, 43-67.

Brunner, Jana, Susanne Fuchs, & Pascal Perrier (2009). On the relationship between palate shape and articulatory behaviour. *Journal of the Acoustical Society of America 125* (6), 3936-3949.

De Jong, Kenneth (1997). An articulatory and acoustic analysis of English flaps: Evidence for attentional reduction in production. *Journal of the Acoustical Society of America 101,* 3178.

Debruyne, Frans, Wivine Decoster, Annemie Van Gijsel, & Julie Vercammen (2002). Speaking fundamental frequency in monozygotic and dizygotic twins. *Journal of Voice 16(4),* 466-471.

Fant, Gunnar (1960). *Acoustic theory of speech production*. The Hague: Mouton.

Fougeron, Cécile & Patricia A. Keating (1997). Articulatory strengthening at the edges of prosodic domains. *Journal of the Acoustical Society of America 101,* 3728-3740.

Fuchs, Susanne, Bernd Pompino-Marschall & Pascal Perrier (2007): Is there a biological grounding of phonology? Determining factors, optimization and communicative usage. *Proceedings of the ICPhS,* Saarbrücken. 219-224.

Hughes, Olive M. & James H. Abbs, (1976). Labial-mandibular coordination in the production of speech: Implications for the operation of motor equivalence. *Phonetica 33 (3)*, 199-221.

Johnson, Keith (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson and J. Mullenix (eds.): *Talker Variability in Speech Processing,* San Diego: Academic Press, 145-166.

Keating, Patricia, A. (1990). The window model of coarticulation: Articulatory evidence. In J. Kingston & M. Beckman (eds.): *Papers in Laboratory Phonology I*, Cambridge: University Press, 451-470.

Lindblom, Björn (1983). Economy of speech gestures. In P. MacNeilage (ed): *The production of speech*, New York: Springer Verlag, 217-245.

Loakes, Deborah (2004). Front vowels as speaker-specific: Some evidence from Australian Englisch. *Proceedings of the Australian International Conference on Speech Science.* Sydney, 289-294.

Loakes, Deborah (2006). A forensic phonetic investigation into the speech patterns of identical and non-identical twins. Ph.D. Dissertation, The University of Melbourne.

Mooshammer, Christine, Pascal Perrier, Susanne Fuchs, Christian Geng & Daniel Pape (2004). An EMMA and EPG study on token-to-token variability. *AIPUK 36*, 47-63.

Mooshammer, Christine, Pascal Perrier, & Susanne Fuchs (2008): Speaker-specific patterns of token-to-token variability. *Journal of the Acoustical Society of America 123 (5),* 3076-3076.

Nolan, Francis & Tomasina Oh (1996). Identical twins, different voices. *Forensic Linguistics 3*, 39-49.

Perkell, Joseph S., Melanie L. Matthies, Mario, A. Svirsky & Michael I. Jordan (1993). Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot “motor equivalence” study. *Journal of the Acoustical Society of America 93 (5),* 2948-2961.

Perkell, Joseph S., Melanie L. Matthies, Harlan Lane, Frank H.Guenther, Reiner Wilhelms-Tricarico, Jane Wozniak, & Peter Guiod (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models, *Speech Communication* 22, 227-250.

Perrier, Pascal (2003): Influences of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study. *Journal of the Acoustical Society of America 114 (3)*, 1582-1599.

Pierrehumbert, Janet (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In Bybee, J. and P. Hopper (eds): *Frequency effects and the emergence of linguistic structure*. Amsterdam: John Benjamins, 137-157.

Przybyla, Beata D., Yoshiyuko Horii & Michael H. Crawford (1992). Vocal fundamental frequency in a twin sample: looking for a genetic effect*. Journal of Voice 6(3),* 261-266.

Savariaux, Christoph, Pascal Perrier & Jean-Pierre Orliaguet (1995). Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube: A study of the control space in speech production. *Journal of the Acoustical Society of America 98*, 2428-2442

Shiller, Douglas M., Rafael Laboissière & David J. Ostry (2002). Relationship between jaw stiffness and kinematic variability in speech. *Journal of Neurophysiology 88*. 2329-2340.

Spinath, Frank M. (2005). Twin designs. In B. S. Everitt & D. C. Howell (eds.), *Encyclopedia of Statistics in Behavioral Science,* Chichester: John Wiley & Sons. 2071-2074.

Stevens, Kenneth N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In P.B. Denes & E.E. Davis (eds.), *Human Communication: A Unified View*. New York: McGraw-Hill, 51-66.

Van Lierde, Kristiane, Bart M. Vinck, Sofia De Ley, Gregory Clement & Paul Van Cauwenberge (2005). Genetics of vocal quality characteristics in monozygotic twins: A multiparameter approach. *Journal of Voice 19(4),* 511-518.

Weirich, Melanie (2009). *Twins in the lab: Reasons for speaker-specific variability*. Paper/Talk presented at Phonetics & Phonology 5, Cologne. (http://www.uni-koeln.de/philfak/phonetik/aktuell/PuP5_Praesentationen/Weirich_Zwillinge_im_Sprachlabor.pdf)

Whiteside, Sandra P. & Emma Rixon (2001). Identification of twins from pure (single speaker) and hybrid (fused) syllables: An acoustic and perceptual case study. *Perceptual and Motor Skills (1).* 319.

The Usage and Distribution of *so* in Spontaneous Berlin Kiezdeutsch

Stefanie Jannedy
Zentrum für Allgemeine Sprachwissenschaft, Berlin

In this paper I investigate the usage of the adverb and particle 'so' in spontaneous speech (interviews) collected from 21 speakers of the urban multi-ethnolectal youth language *Kiezdeutsch*. Speakers from the neighborhoods Kreuzberg and Wedding in Berlin are ranging in age from 14 to 18. The 1454 tokens of *so* available in the corpus (about 5 hours of speech) were classified into 10 different categories; some were structurally defined while others were defined along dimensions of meaning. Our current results indicate that there are differential usages patterns depending on the speaker's gender and age for some of these categories. Further, it appears that some patterns that have been attributed grammatical meaning may not appear frequently enough to establish a separate meaningful grammatical category. Rather, most instances of this kind of use of *so* appear to have a hedging function, indicating speakers' non-commitance to a specific circumstance.

1 Introduction

Labov's famous 1966 study on the social stratification of English in New York City marks the advent of urban sociolinguistics. As our world is becoming more and more global and connected in significant ways, migration and integration are challenges and chances at the same time as multi-cultural and multi-ethnical societies are emerging predominantly within urban areas. As multilingualism, cultural diversity and social integration are challenges to be mastered, we know that linguistic expression of individual style and group identity by young speakers in major urban areas are driving forces of linguistic innovation and language change which have lead to the emergence of new multi-ethnolects and distinct urban vernaculars. For example, Torgersen et al. (2006) showed that the locus of linguistic innovation and language change is inner-city East-London, an area with a large immigrant population. Even though urban areas in Germany are characterized by the multi-ethnicity of its population, differences in cultural heritage, and linguistic diversity, sociolinguistically informed quantitative investigations of multi-ethnolectal variation within and across different urban areas have not been undertaken in Germany.

In many European cities, researchers have noticed and studied the emergence of linguistic variation and the grammatical innovations introduced by young speakers from multi-ethnic urban neighborhoods (Multicultural London English: Torgersen et al., 2006; Kerswill et al., 2008; Straattaal (Netherlands): Nortier, 2001; Appel, 1999; Rinkeby-Svenska (Sweden): Kotsinas, 1992 and 1998; Bodén, in print; unpublished ms.; Kobenhavnsk multietnolekt (Denmark), Quist, 2005). However, most language research in Germany has largely neglected variation in speech based on sociolinguistic factors and predominantly focused on lexical or pronunciation variation along geographical and dialectal dimensions (e.g. Deutscher Sprachatlas: Herrgen, 2007).

Through waves of immigration during the 60s and 70s particularly from Turkey and Kurdistan (eastern Turkey), first generation speakers went through a more or less uncontrolled second language acquisition of German and learned a day-to-day variety of German simply by picking it up. Work on this so called *Gastarbeiterdeutsch* of the first generation immigrants was done (among others) by Keim (1978) and Pfaff (1981). Initially it was thought to be a variety or slang spoken by young people of Turkish decent only. Feridun Zaimoglu (1995) coined the in-group name *Kanak-Sprak* for the speech of adolescent males of Turkish descent which however appeared to have negative connotations for users outside of the group of speakers. Basic descriptive groundwork on the speech of second and third generation immigrants from Turkey and their mono-ethnic German peers has been laid by Androutsopoulos (2001) and Auer (2003) who used the terms *Türkendeutsch* and *Türkenslang* respectively to refer to this variety of German. Today there is a general agreement that an appropriate name for this variety of Germany needs to reflect the fact that speakers are young multiethnic urban speakers with a wide variety of language backgrounds (Auer & Dirim, 2004; Wiese, 2009; Krivokapic et al., 2010), so a term was coined that does not reference the speakers but the location where this variety is spoken: locally identified, tightly knitted neighborhoods all over Germany. The term *Kiezdeutsch* at best functions as a shortcut to invoking the notion of a highly stigmatized urban multi-ethnolectal youth language, often spoken in migrant communities in larger metropolitan areas in Germany which emerged on the basis of German and other languages such as Turkish and Arabic. For that matter, it is much more than an *inability* or *refusal* to speak (proper) German, it is more so an act of identity. It however completely neglects the fact that the term *Kiez* has a variety of different meanings, depending on location: in Berlin, a *Kiez* is a neutral term referencing a small local neighborhood, in Hamburg there is only one Kiez which happens to be the red-light district, and other urban centers such as Cologne or Munich do not have Kiezes at all. Thus, for the lack of a better term, the usage of *Kiezdeutsch* nowadays is common in the literature. *Kiezdeutsch* is not only characterized by phonetic or phonological alternations such as the realization of the palatal fricative /ç/ as post-alveolar fricative /ʃ/ (Auer, 2003; Mertins, 2010). While Turkish does not have palatal fricatives, we assumed that this sound is being substituted with a sound that is available in L1. However, our data also reveals alternations between the palatal and the post-alveolar fricative for some speakers, suggesting, that beyond substitution, other mechanisms for sound selection are at work. Other phonetic differences to German as spoken in Berlin or even more standard varieties are the avoidance of consonant clusters or differences in the realization of diphthongs. Auer (2003) and Wiese (to appear; 2009a) and others also describe various morpho-syntactic alternations, which we will not discuss here, the reader is advised to look at the references for literature on this issue and for examples of such alternations.

While collecting spontaneous lab-quality speech data through linguistic interviews from inner-city Berlin adolescents from Kreuzberg and Wedding, we noticed the pervasive use of the particle *so* 'like', occurring at the edges of phrases and phrase medially. It appears that *so* is being used in a wide variety of contexts and functions which will be explored in this study.

Example (1) shows instances of overuse of *so* by a 16-year old male German speaker of Turkish descent from Kreuzberg:

(1)	Ich	red	mit	dem	Mann	**so**	ganz	**so**	locker	spontan	**so**
	I	speak	with	the	man	so	very	so	cool	spontaneously	so

'I speak with the man like very cool and kind of spontaneously'

sehr	**so**	freundlich	und	**so**
very	so	friendly	and	so

'very like friendly and so on'

Another reason for classifying the usage and distribution of *so* is to model the duration of phrase-medial and phrase final *so* (Krivocapic et al., 2010) as to have means to correlate the duration data with the respective phrasal position. However, this work is discussed elsewhere. In this study, we set out to describe the usage-patterns, positions, and meanings of the particle *so* in this variety of German and will describe the distributions of the 1454 *so* in our database.

1.1 Meaning, Function & Classification of *so*

The particle *so* in German is multi-functional and the scope of meanings is difficult to capture which is evidenced by repeated attention it received during the last years (Umbach & Ebert, to appear; Hennig, 2007; Paul, 2008; Wiese, 2009a, 2009b). The Duden (2004) lists about nine different uses and meanings for *so*, among them as a deictic element, as an indicator of finality, degree or intensity, but also as a marker of a comparison or consequence. Hennig (2007) and Umbach & Ebert (to appear) address the problem of grammatical classification and part-of-speech membership of *so*. Hennig (2007) points out that the classification of *so* from reference works alone does not capture all meanings and that expressions containing *so* are hardly mentioned in theoretical works on this topic. She concludes based on her analysis of a random sample of roughly 50 tokens each of *so* from written text and spoken corpora, that a grammatical classification of *so* is rather difficult if not problematic because *so* often occurs in (idiomatic) expressions that are difficult to classify. She postulates the inclusion of the investigation of phonetic / intonational properties of *so* from spoken discourses to determine word-class membership and pragmatic meaning of this word. She notices that empirical work on such issues can point out problems and issues which would remain otherwise undetected in purely theoretical treatments of such a topic as we might not think of forms or usages of *so* that occur in spoken corpora of unscripted speech.

This however is a problem with many if not all empirically underpinned investigations (of part of speech classifications) from unscripted spontaneous speech: the corpus may not contain instances of all different kinds of use. We are well aware of these issues and by no means do we argue to have come up with an exhaustive list of occurrences and usage patterns of *so* in *Kiezdeutsch* in general. What we will show in section 2 of this paper is what seems to have emerged as sensible groupings from our corpus of spontaneous Berlin Kiezdeutsch. It is worth pointing out that *so* seems to be in some ways similar to the English *like* in that it can have grammatical function as an adverb but also discursive functions such as a discourse particle, a discourse marker or a quotative marker (Drager, 2010; D'Arcy, 2007).

Paul (2008), Wiese (2009a) and Wiese et al. (2009b) also recognize different usage of *so*: *so* can follow an argument as in *für Jugendliche so* 'like for adolescents', *mein Dings so* 'like my thing'; *so* can occur with prepositional phrases as in *so im Grünen* 'like out in the nature' or in *so aus Schöneberg* 'like from Schöneberg'; with adjective phrases *so blond so* 'like so blond' and it can occur with or precede an argument such as a bare noun as in *so Club* 'so club', *so Billardraum* 'so pool room', *so Naturtyp* 'so nature type'. Our ZAS-corpus of

Kiezdeutsch also shows such instances of use, thus we concur fully with what Wiese and colleagues state.

They further suggest though that some instances of use of *so* may serve to mark information structural prominence. In fact, they propose that in *Kiezdeutsch*, *so* is currently taking on a new and additional function, namely that of a focus marker". Wiese et al. (2009b:22) say about *so*:

> "[…] it can precede its argument […], follow it […], and it even occasionally brackets it […]. […], *so* in this usage is always combined with the focus constituent of the sentence, which carries the main accent. If one takes information-structural aspects into account, then, this seemingly erratic behavior can be subsumed under a unified account of *so* as a focus marker, a particle that attaches to the respective focus constituent in a sentence."

Thus, the authors attribute one of the functions or meanings of this particle to intensify the expression that is under the scope of *so* and lend it some kind of prosodic prominence. In fact, Wiese and colleagues have suggested such a function and claim the emerging or potentially grammaticalized function of *so* as a focus marker. They specifically mention the bracket construction whereby a *so* precedes and a second *so* immediately follows the argument ([*so* … *so*]).

Umbach & Ebert (to appear) provide a theoretical investigation of out-of-the-blue usage of *so* and argue that *so* is a demonstrative expression, combining with gradable and non-gradable expressions. They classify the usage of the German demonstrative *so* into three different groups: 1. deictic and anaphoric *so*; 2. intensifying *so*; and 3. hedging *so*. They suggest that *so* has an intensifying meaning that can be compared to *sehr* 'very' if it precedes gradable expressions such as adjectives as in (their example 3) *er ist so groß* 'he is so tall'. They further observe that *so* can combine with non-gradable expressions such as nouns (their example 4) *Ich möchte so Klammern* 'I want like clips'. In this usage, they propose, *so* expresses hedging and some kind of uncertainty about the appropriateness of the selected term. Consider the minimal-pair type example in which the *so* is unaccented and the last accent falls on the utterance final adjective *blau* as in 'Der Himmel ist so blau.' 'the sky is so blue' versus 'Das Kleid ist so blau.' 'the dress is like blue'. In the latter example, *so* is much more likely to receive a hedging interpretation.

Even though both groups of authors identify an intensifying meaning of *so*, they do not seem to agree on the meaning of *so* before non-gradable expressions. Thus, the interpretation of *so + noun* or any other type of argument (plus a following *so*) by Wiese and colleagues is in direct opposition to the interpretation proposed by Umbach & Ebert (to appear). Even though we have not classified occurrences of *so* according to gradable or non-gradable arguments, we have found and coded bracket constructions in our data. In this paper, we will offer a phonetic-phonological argument as to why we find the reasoning on the information-structural function and meaning of *so* offered by Wiese and colleagues not convincing.

The discussion in the existing literature was helpful for establishing our own classifications of *so*, however, it was still challenging to attribute meaning and function to all of the occurrences of *so* that we have found in our corpus. It is not our aim to add to the discussion on part of speech classification of *so*, we have merely looked to that body of literature to help us set criteria for our own classifications. These we deemed necessary to establish a level of description of the overall functions and meanings of *so* in this multi-ethnolect. In this study we will quantitatively investigate actual discourse usage patterns of *so* in a multi-ethnolect which is - among other features - characterized also by the over-use of this particle. The amount of data that we have collected from this multi-ethnolect allowed us to evaluate specific claims brought forward in the literature.

2 Methods

2.1 *Gminer*

To get a handle of the massive amount of spontaneous speech data, all recordings were first orthographically transcribed with a freeware audio-transcription tool *Transcriber* (version 1.5.1). The transcriptions are time-aligned with the audio-signal and anonymized programmatically. The transcription conventions such as the usage of punctuation (“,”, “.”, “-“ etc.) for different types of pauses were custom developed for this type of spontaneous data and adjusted on a need basis (Mertins, 2010). The output of *transcriber* plus the associated audio files were then uploaded into a browser based database search tool installed on a virtual server. This data mining tool is based on the *ONZE-Miner* (Fromont & Hay, 2008) which was originally developed to search through hundreds of hours of historical recordings of (the **O**rigins of) **N**ew **Z**ealand **E**nglish. The tool that we have used was localized for use with German data and we have named our data mining tool the *Gminer* (**G**erman miner).

The *Gminer* provides customizable search-spaces for adding speaker specific meta-information associated with that particular interview such as the age, gender, native languages, attended type/level of school, ethnicity, or neighborhood etc. of the speaker. Further, integrated into the *Gminer* is the German CELEX-dictionary, allowing for automatic canonical tagging of the lexical forms contained in the interview that was uploaded: automatically given are the phonological representation of each word form, the syntactic category, morphological structure, the overall word frequency and several other parameters available through CELEX. This meta-information can be displayed on separate layers in the transcription in the browser. The *Gminer* allows for sophisticated searches across words and across different layers whereby custom layers with specific annotations can also be added. A great advantage of this tool is the capability of downloading sound files associated with the search results for further annotation or segmentation in acoustic analyses software such as Praat. Krivokapic, Fuchs & Jannedy (2010) have used this functionality to first search, and then download hundreds of files and measure the duration of the /s/ and /o/ of the particle 'so' in phrasal-final positions to evaluate a data-driven analysis of different levels of the prosodic hierarchy.

Depending on the research question, search results (across several words and layers of annotation) as well as the associated meta-data (speaker information) can be exported into a spreadsheet and further marked up with relevant linguistic information (e.g. if the particle *so* is preceded by a noun or if it is following a noun etc.). This marked-up spreadsheet can be easily imported into *R*, a powerful statistics work package suitable for graphical and statistical exploration of large amounts of data.

2.2 Speakers

For the purpose of this study, we have extracted all instances of *so* from 21 speakers of Kiezdeutsch. 18 speakers were from Kreuzberg, only 3 from Wedding, thus, at this point we are not able to look for differences rooted in their local neighborhoods. As we have recorded 10 male and 11 female speakers, we are able to look for gender differences in the distribution of the data. Speakers were distributed across 5 different age groups ranging from 14 through 18 (1 x 14; 6 x 15; 8 x 16; 4 x 17; 2x18). The data was also coded also for factors such as *school form* attended, *native language* and *country of birth*. These factors however could also not considered at this point.

2.3 Categorization of *so* into ten different groups

In total, 1454 instances of usage of *so* have been extracted from the database. Each token was further tagged and annotated by hand for usage and function by the author and colleagues. We have abstained from theoretical assumptions of the use or grammatical group membership of *so* and tried to capture the actual meaning or the structural surroundings of this word. In accordance with syntax- and semantics experts and the existing literature, the following categorization criteria were established. It may be argued that in several ways these categorizations are oversimplifictions which gloss over more complex differences between the instances of use encountered in the data. Nevertheless have decided to used these criteria and categorization to make the crude point that usage patterns of *so* can either be structurally or semantically/pragmatically be defined. Further, it should be noted that we have not counted the same instance of use in different categories but made a choice what group to include this token with.

2.3.1 Categorization according to meaning

With this initial investigation, we have subsumed what Umbach & Ebert (to appear) call the hedging-*so* and the intensifying-*so* in a category that we have named *degree-modifier*. Instances of use in this group modify the degree of its argument. Examples for this category are *Türkisch kann ich auch nicht so gut* 'I can't speak Turkish so well' or *sind so viel Fragezeichen* 'there are so many question marks'. It is planned to further investigate these cases because naturally occurring language from spontaneous interviews may have forms that are not taken into consideration in theoretical deliberations on use and function. For example, there is an abundance of cases where the particle *so* appears after the argument and before a phrasal break, thus, clearly referencing the preceding material. However though, at this point we were not able to fully dissect this category into further subgroups.

All instances of *so* that occurred as reference to an object to which an entity was compared to were categorized as *comparison*. Examples are *so wie meine besten Freunde* 'just like my best friends', *ich fühle mich so als, als Berlinerin* 'I feel like a Berliner' or *so wie ein Deutscher* 'like a German'. Items were categorized as *correlate* when they related one state to another as in or *bei uns ist es so, dass* 'at our place it is like this' or as in *er will halt nicht so, dass ich Kopftuch trage* 'he does not want me to cover my head'. All cases of *so* that referenced something were categorized as *referential*. Examples of this category are *ich gucke sie immer so an* 'I look at her like that' or *war doch so!* 'it was like that!'. We were left with a category of miscellaneous items (*misc*) that were not readily classifiable. These occurred for example before pauses or in utterances that are characterized by false starts and such. Examples are *ja so ja äh.* 'yes, so yes uhm' or *so mh eigentlich so* 'so uhm, actually so'. We are well aware of the problem of potential ambiguities between categories and will have raters naïve to the purpose to this study as well as semantics experts reconfirm or dispute current judgments. We do expect however, that due to the relatively large sample size for spontaneous data, the overall distribution of categories in the data will remain relatively stable.

2.3.2 Categorization according to structure

Categorizations of *so* based on structure was somewhat less complicated than categorizations according to meaning. The five categories described here are mainly based on structural or co-occurrence descriptions, thus, they seem fairly straight forward with little room for dispute. We have classified all instances of *und so* where *so* occurs in immediate proximity following *und* 'and' as *coordinative* and all instances of *so* in immediate proximity following *oder* 'or' as *alternative*. The instances when *so* was directly followed by direct speech or quotations (orthographically marked in the transcribed data by adding a colon and quotation marks) were classified as *quotative* usages of *so*. Examples are *Ich so: "was macht ihr denn hier?"* 'I was like, what are you doing here?' or *da denkst Du so: "äh?!"* 'you're thinking like 'huh'?!'

Brackets are structurally defined through the occurrence of *so* before and after a word, sequence or argument as in the following examples: *die sind alle* ***so*** *verteilt* ***so*** *in der Türkei* 'they are all very dispursed like in Turkey'; *der geht* ***so*** *Berg hoch* ***so*** 'he like walks up a mountain'; *ist nicht* ***so*** *schwer* ***so*** 'is not so difficult like that' or ***so*** *Männergespräche* ***so*** 'like male talk'. The usage of *so* in structures like brackets supposedly combines with the focus constituent of the sentence (Wiese et al. 2009) and thus mark focus. Expressions such as *so oder so* 'so or so', *so und so* 'so and so' and *ach so* 'oh really' are subsumed in the category of 'bracket'. Note that the bracket construction of *so*___*so* was only counted as one instance of use of *so*.

2.4 Statistics

We conducted contingency table- and goodness-of-fit tests (chi-square analyses) with *age* and *gender* as independent factors and the number of counts produced for each of the ten categories of *so* as dependent variable. (The overall structure of our data generally allows for further analyses with factors such as *school-form*, *neighborhood*, *country-of-birth,* or *mother-tongue*. However, currently, some of the cells in the table were empty due to not having enough data and thus, no analyses were performed.

In those cases were the chi-square approximation calculation generated warning messages due to low counts in some cells, we ran several *monte carlo simulations*[1] with 10000 runs each. Each of these simulations generated different p-values, yet, the simulations consistently resulted in p-values that were reliably significant. Therefore, we can be sure that overall, the comparisons that involve

[1] Also see the R-help pages for chi-square analyses: *help(chisq-test)*. The actual R-command line is: *chisq.test(<table.name>.tab,simulate.p.value=TRUE,B=10000)*

cells with low counts are significant. Since no degree of freedom (df) is reported in these calculations, it can easily be identified where we ran the additional simulations. Further, in instances where we wanted to test for significant differences between factor levels (e.g. differences in usage of a particular *so*-usage-category by 16- vs. 17 year olds), we used a procedure, testing if the proportions are the same in different groups of data (R: prop.test).

3 Categorization Results & Usage Patterns

In the following section, we will show the distribution of the *so*-tokens into the 10 categories by showing raw counts in the graphs. Figure 1 shows a barplot for the overall distribution of the data into the categories. The categories are discussed in order of frequency of occurrence of the pattern in the data.

Most instances of use (548; 37.7%) of *so* are to modify the **degree** of its argument in cases such as *so schlimm* 'so bad' or *nicht so oft* 'not that often'. This use is also very well attested for standard varieties of German and seems to be the default use for *so* in German. The second largest group is comprised of instances where speakers have a **referential** use of *so* (330; 22.7%) in instances such as *bei uns ist so* 'with us that is the way it is' or in *so ein weiße Mütze* 'such a white cap'.

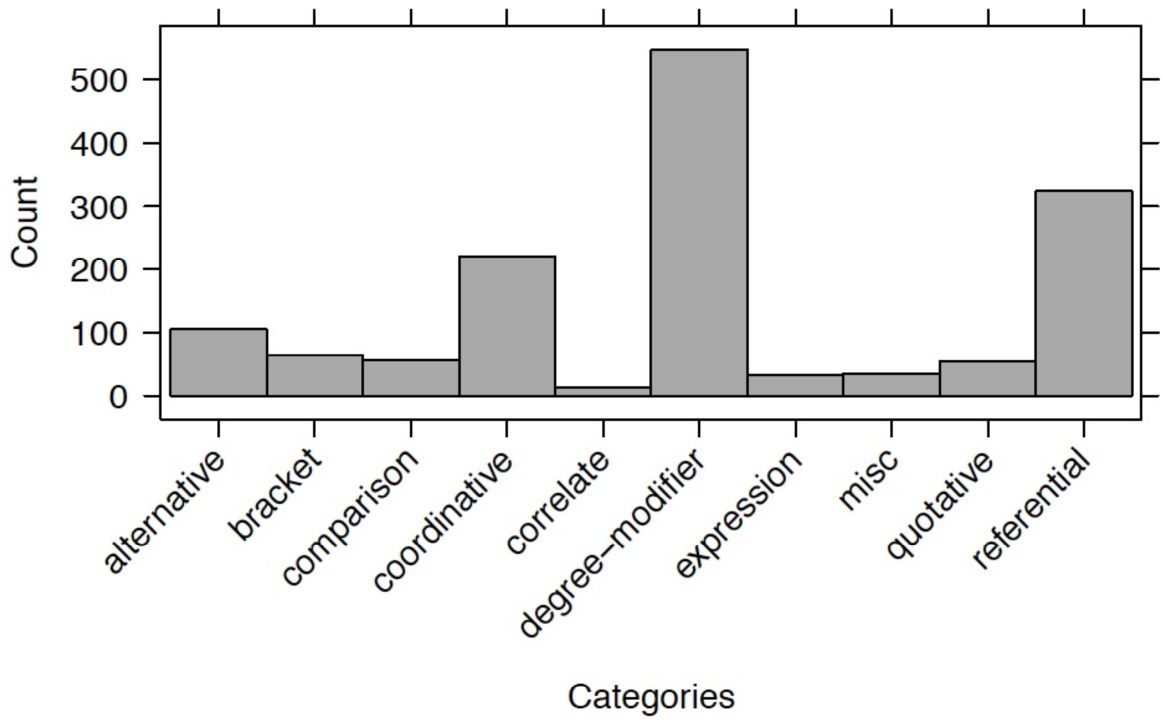


Figure 1: Cumulative graph of the differential use of *so* in the ZAS- Kiezdeutsch spontaneous speech database (raw numbers).

The use of *so* in a **coordination** *und so* ‘and so on’ occurred 222 times (15.3%), examples are *auf der Straße und so* ‘in the street and so on’ or *Schule, Universität und so* ‘school and university and so on’. Closely related is the category is the group we termed **alternative** *oder so* as it also combines a conjunction with the particle. We find 105 instances of use in the data (7.2%), examples are *dritter Monat oder so* ‘third month or so’ or *Türke oder so* ‘Turk or so’. To mark **comparisons** as in *wie so ein Tuschkasten* ‘like a paintbox’ or *Ach, Potsdam ist wie so ein Dorf ‘well, Potsdam is like a village’*, *so* was used 57 times (3.9%). The structurally defined **bracket** category occurred 66 times in the corpus (4.5%) – this category will be discussed in more detail below. **Quotative** constructions such as ... *und ich so:”Oh mein Gott”* ‘I was like: ‘Oh my God’ or *ich dachte so: “Nein!”* I thought like: ‘No!’ made up 3.7% (54 tokens) of the data in the corpus. The remaining three groupings are **correlates** (11 cases, 0.8%) like *so, dass* ‘so that’, **expressions** (32 cases, 2.2%) like *ach so* ‘I see’ or ‘ and a miscellaneous category that contained unclassifiable instances of *so* (29, 1.9%).

Dividing the data by gender reveals an overall effect with males generally using more instances of *so* than females (Pearsons χ^2 = 58.6765, df=NA, p<.001 with simulated p-value based on 10000 replicates).

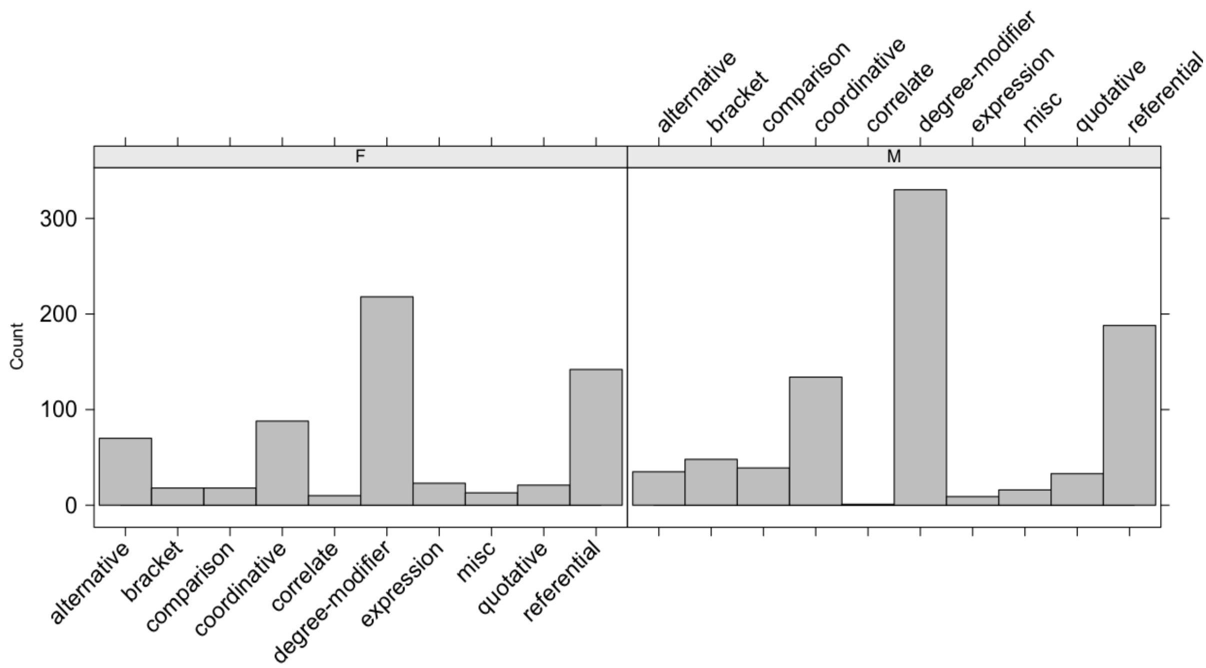


Figure 2: Cumulative graph of the differential use of *so* divided by the factor *gender* (raw numbers).

Individual comparisons for each category type by gender revealed significant effects for *alternative* (χ2= 25.501, df = 1, p< .001) and *expression* (χ2= 10.1885, df = 1, p< 0.002) with female speakers producing more tokens in these categories than males. The data also shows that males produce more *brackets* ($\chi2$= 4.7806, df = 1, p=0.028) than females, suggesting that this structure may be an innovation predominantly used by males.

Comparing the data by age reveals an overall significant effect ($\chi2$= 20.3349, df = NA, p< 0.001) with regard to the use of the bracket category. Data split up by category and age group is shown in Figure 3. This is not surprising given that we have already found a significant effect for gender and the group of 17 year olds is merely comprised of female speakers, under-using this linguistic innovation. A comparison of the proportions of usage of the bracket category by different age groups fails to reach significance between the 15- (13 instances of use) and 16- (44 instances of use) year olds ($\chi2$= 3.3147, df = 1, p = 0.068). A comparison between the 16- and 17-year olds (4 female speakers, 0 instances of use of the bracket category), shows a significant difference ($\chi2$-squared = 7.9202, df = 1, p< 0.01).

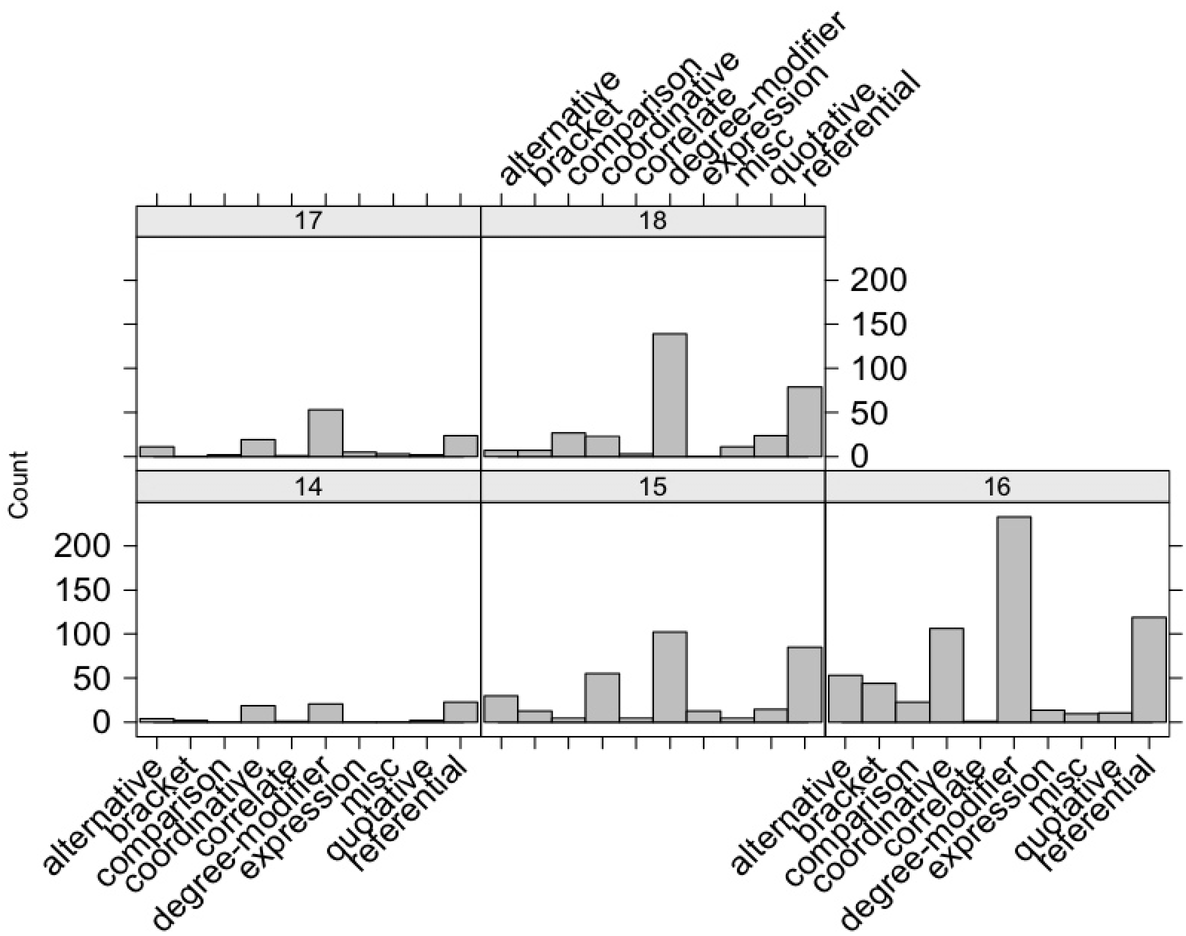


Figure 3: Cumulative graph of the differential use of *so* divided by the factor *age* (raw numbers).

The data suggests that usage of the bracket category (second bar from the left in each of the graphs in Figure 3) is gendered, it seems to be propagated especially by 15- (3M/3F) and 16- (5M/3F) year old males, whereas none of the 17 year old girls (0M/4F) produced such a token. The 18-year old male produced more bracket structures than the 18-year old female. Whether the data is truly age graded remains to be seen though, currently there is not enough robust evidence to make such claims given that at the time of analysis, there was only data from one 14-year old male and 2 18-year old speakers. As we are constantly adding data to our corpus, eventually we will be able to generalize over these age groups, too.

4 The 'bracket' Category

In total, there are 66 tokens in the 'bracket' category. Only 12 of the 21 speakers that we sampled for this study produced such a pattern. The structurally defined 'bracket' category itself is subdivided into three categories as shown in Figure 4: 1. coordinative structures (24) such as *so groß und so* 'so tall and such'; 2. focus structures in which (according to Wiese, 2009; Wiese et al., 2009; and Paul, 2008) the particle *so* is proposed to serve as a focus marker (36) in this multi-ethnolect of German. An example is *Mit meinem Vater red ich eher über so Männergespräche so* 'with my father, I speak about male topics if anything' and further potential focus structures.

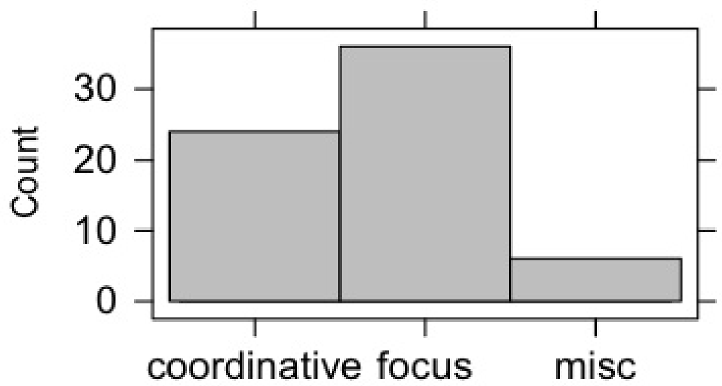


Figure 4: Cumulative graph of the differential use of *so* in the bracket category *so*___*so* in the Kiezdeutsch spontaneous speech database (raw numbers).

A third group includes all remaining structures that does not have great commonalities and is thus grouped in the miscellaneous category (6). Examples are expressions such as *so oder so* 'so or so' and *so und so* 'so and so' and other non-classifiable items such as *so weiter so machen* 'continue to do so'. Figure 4 shows the entire set of bracket structures classified into these three sub-categories. A proportions test reveals that a single speaker produces significantly more focus structures than all other speakers together ($\chi 2 = 7.6768$, df = 1, $p < 0.01$). Thus, the extent to which we find *so___so* bracket structures that potentially fulfill the criteria for receiving a focus interpretation seem to be due to a speaker effect.

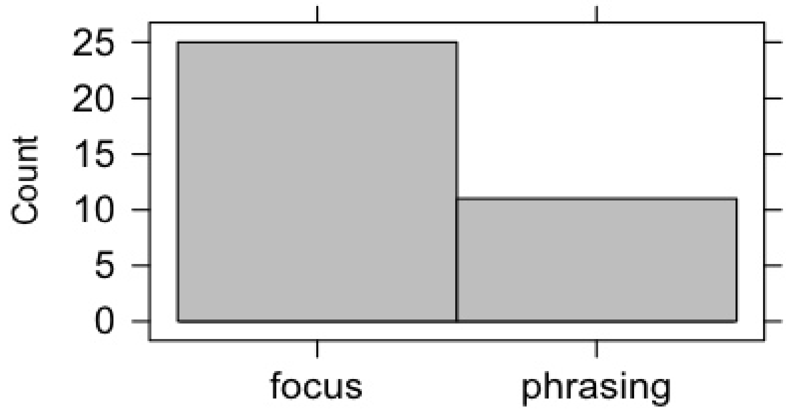


Figure 5: Cumulative graph of bracket category so___so in the Kiezdeutsch spontaneous speech database (raw numbers), divided by phonological phrase breaks.

A closer prosodic analysis of the 36 tokens that did satisfy the description of being a focus structure revealed that 11 of the utterances contained either an intermediate or an intonational phrase break either after the preceding *so* and the argument or between the argument and the following *so*, leaving 25 instances. Example (2) shown in Figure 6 shows a bracket structure which contains a phrase break after the initial *so* indicated by the lengthening:

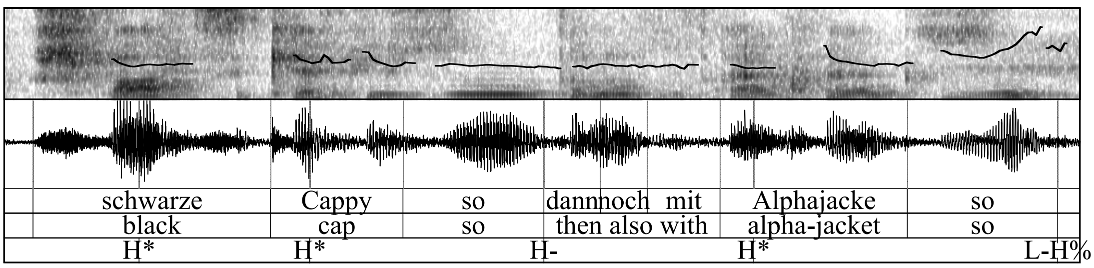


Figure 6: Spectrogram, waveform and f0-track for a bracket-sequence *so___so*, showing a phrase break as indicated by the lengthening of *so*.

(2) wenn man **so** // mit schwarze Cappy **so**
when one like - with black cap so
"[what impression does it leave] if you're with black cap

dann noch mit Alpha-jacke **so**
then also with Alpha-jacket so
and then also with an alpha-jacket on"

Another example that indicates phrasing includes *wenn man* ***so*** *arabisch oder türkisch //* ***so*** *laut spricht, dann haben die Angst [...]* ‘when one speakers Arabic or Turkish rather loudly, they they are afraid’.

Some of the remaining 25 utterances are doubtful examples to the focus theory, too: *der geht so Berg hoch so* ‘he walks like up the mountain’ where the argument in focus would presumably be *Berg hoch* ‘up the mountain’ but where the accent in this case is located on *hoch* and *Berg* is unaccented. A second example is more or less untranslatable: *Und äh immer* ***so*** *Dings* ***so*** *halt* so ‘and aeh, always something like’ where the Argument in focus would be either the unspecified noun *Dings* or even the unspecific discourse particle *halt*. Given this non-specificity, it is unlikely that the speaker attempts to draw attention to the content of these lexical items. In ***so*** *erstmal* ***so*** *Ferien* ‘so like first vacation’, the focused element would be *erstmal*, a particle that also is very unlikely to have much attention drawn to it.

5 Summary & Discussion

While we have collected hours worth of data, at the present time, the database is not yet well balanced with regard to having comparable numbers of speakers for each age group, gender, neighborhood and native language. We recognize this as our shortcoming and ongoing and future work will remedy some of these issues: We are still recruiting speakers, interview them and add their data to our corpus for future analyses. While we are open to the criticism that our data does not adequately reflect all of the morpho-syntactic, lexical and phonetic/phonological variation that occurs in day-to-day interaction in the streets, we are confident that by now we have enough material that is of good enough quality that it lends itself to corpus analyses of spontaneous speech data and allows for generalizations over groups of speakers, especially in the domain of phonetics and phonology. For example, some of the data that we have collected is used for cross-dialectal perception studies (Jannedy, Weirich, Brunner & Mertins, 2010). Previous perception work on this topic was conducted with specifically created or recorded stimuli (see for example Niedzielski, 1999; Brunelle & Jannedy, 2010). Since our interviews also capture the interactional styles of the speakers, the data naturally lends itself also to investigations on the interface of morphosyntax and phonology which to a small degree we have exploited in this paper by evaluating the occurrences of *so* in bracket structures with regard to the phonological structure of the utterance as a whole.

We have found empirical evidence for the multiple uses, meanings and functions of the adverb and particle *so* in multi-ethnolectal Kiezdeutsch German. In about 500 minutes of spontaneous speech, there were 1454 occurrences of *so* (corresponding to an average almost 3 *so* per minute). Empirical evidence strongly suggests that there are gender, age and speaker effects in the usage of *so* whereby most usage patterns are well attested in standard varieties of German, too. One of these structures, the *so*___*so* bracket construction was proposed to function as a focus marker.

In examples like *Mit meinem Vater red ich eher über* ***so*** *Männergespräche* ***so*** 'with my father, I speak about male topics if anything' or in *Ähm, ich will* ***so*** *über Islam* ***so*** *[Vorträge geben]* 'I want to [give talks] about Islam', both taken from our corpus, it seems that an accent or at least some kind of acoustic prominence would fall on *Männergespräche* and also on *Islam*. This could be taken as evidence that within the bracket, there is accentual marking of focus. It is however the case that accents in languages like English and German often go on the last accentable constituent of an utterance. In a way, this is a default position for accent since it is much less marked than an accent early against a longer unaccented post-nuclear tail. It ought to be noted that not very accent marks a focus, and thus, in examples of the type given, there may be an accentual prominence which is unrelated to the pragmatic focus.

Moreover, we showed that a great proportion of the bracket structure was produced by a single speaker, calling into question the wide-spread distribution of this pattern or its rise to a grammaticalized pattern to indicate focus. All in all, of the 1454 occurrences of *so*, only 66 satisfied the structural description of 'brackets'. Within these 66 brackets, 36 satisfied the 'focus' structure (*so* .... *so*). And of these 36 that satisfied the focus structure, 25 had no phrase break (prosodic boundary) between the initial *so* and the argument and ultimately satisfied the structural description of these focus constructions. As some of the examples showed though, not all material enclosed in the *so* bracket is really meaningful. Further, even instances that structurally and prosodically fulfill the criteria may ultimately just do so because the default accent location is late in an utterance, thus, an accent on the argument enclosed in the *so* bracket that occurs late in an utterance may just receive a default accent rather than a focal accent. All in all, just about 25 of 1454 (1.7%) utterances contained the bracket-focus structure in this corpus. We call into question that this manifests a pattern, especially since the data shows that many of the bracket structures were produced by a single speaker.

There is a list of issues that have not been considered for the current scope of the paper. In the future though, we hope to address these: the categorization into *so* plus a gradable versus *so* plus a non-gradable expression; the phonetic-phonological categorization and implementation of *so* – when is it accented, when not, is the material following *so* always accented, is it only sometimes accented? If so, does it correlate with a specific structure or meaning? Based on the data that we have found and analyzed, we are not convinced of the emerging function of *so* as a focus marker in this multi-ethnolect. Rather, in most instances that do not have a clear meaning or function (quotative use or *so* before adjectives), we have associated a hedging interpretation with this particle, where the speaker refrains from being more specific about the argument and leaves much of the interpretation to the addressee of the discourse. This for example can also be tested in perception/rating tests with naturally collected data. Due to the pervasive use of *so* in Kiezdeutsch for some speakers, this multi-ethnolect lends itself well to an investigation of the durational properties of this particle in various prosodic positions within the utterances (Krivokapic, Fuchs & Jannedy, 2010). This work is in progress and will be discussed elsewhere.

Acknowledgements

I kindly acknowledge the diligent work and patience of my research assistant Micaela Mertins without whom none of the transcripts would have been created and tripple-checked. I thank my colleague Melanie Weirich, although much busy with her own PhD-research, who commented on this research at various times, and who was never too busy to discuss issues that arose with this work. Members of our lab made themselves kindly available for any help that was needed and their advice was and still is invaluable. The ZAS-Beirat and my colleagues spent time with me discussing this work, I thank them. Last but not least I thank our interviewees for their time, their stories and efforts - without them we would not have had any data to analyze. All errors remain my own.

References

Androutsopoulos, Jannis. (2001). *Ultra korregd Alder!* Zur medialen Stilisierung und Popularisierung von „Türkendeutsch". *Deutsche Sprache* 4/2001, 321-339.

Appel, René (1999). *Straattaal*. De mengtaal van jongeren in Amsterdam. Toegepaste Taalwetenschap in *Artikelen* 62;2,39-55.

Auer, Peter (2003). ‚Türkenslang'. Ein jugendsprachlicher Ethnolekt des Deutschen und seine Transformationen. In: Häcki-Buhofer, A. (ed.). *Spracherwerb und Lebensalter*. Tübingen. Francke, 255-264.

Auer, Peter & Dirim, Inci (2004). *Türkisch sprechen nicht nur die Türken: Über die Unschärfebeziehung zwischen Sprache und Ethnie in Deutschland*. De Gruyter.

Boersma, Paul & Weenink, David (2009). *Praat*: doing phonetics by computer (version 5.0.47) [Computer program]. http://www.praat.org.

Bodén, Petra (in print). Pronunciation in Swedish multiethnolect. In Svendsen, B. A. & Quist, P. (eds), *Linguistic Practices in Multiethnic Urban Scandinavia*. Multilingual Matters.

Bodén, Petra (unpublished manuscript). Adolescents' pronunciation in multilingual Malmö, Göteborg and Stockholm. In Lindberg, I. & Källström, R. (eds).

Boudahmane, Karim, Manta, Mathieu, Antoine, Fabien, Galliano, Sylvain & Barras, Claude (2008). *Transcriber*. A tool for segmenting, labeling and transcribing speech (version 1.5.1). [Computer program]. http://trans.sourceforge.net.

Brunelle, Marc & Jannedy, Stefanie (accepted). The Cross-Dialectal Perception of Vietnamese Tones: Indexicality and Adaptability. In Hole, Daniel and Elisabeth Löbel (eds.). *Linguistics of Vietnamese – an International Survey*. Berlin/New York: Mouton de Gruyter (TILSM series).

D'Arcy, Alexandra (2007). *Like* and language ideology: Disentangling fact from fiction. *American Speech* 82 (4), pp. 386–419.

Dirim, Inci & Auer, Peter (2004). *Türkisch sprechen nicht nur die Türken. Über die Unschärfebeziehung zwischen Sprache und Ethnie in Deutschland. Berlin*. De Gruyter.

Drager, Katie (2010). Sensitivity to Grammatical and Sociophonetic Variability in Perception. *Laboratory Phonology* 1(1), pp. 93-120.

Fromont, Robert and Hay, Jennifer (2008). ONZE Miner: the development of a browser-based research tool. *Corpora* 3(2), pp. 173-193.

Hennig, Mathilde (2007). So, und so, und so weiter. Vom Sinn und Unsinn der Wortklassifikation. *Zeitschrift für Germanistische Linguistik* 34, pp. 409-431.

Herrgen, Joachim (2007). From Dialect to Variation Space: The "Regionalsprache.de" Project. In The National Institute for Japanese Language: Geolinguistics around the World. *Proceedings of the 14th NIJL International Symposium*. Tokyo, pp. 75-80.

Jannedy, Stefanie, Weirich, Melanie, Brunner, Jana & Mertins, Micaela (2010). Perceptual Evidence for Allophonic Variation of the Palatal Fricative /ç/ in Berlin German. Poster to be presented at the Meeting of the *Acoustical Society of America* in Cancun, Mexico.

Kerswill, Paul, Torgersen, Eivind, Fox, Sue (2008). Reversing "drift": Innovation and diffusion in the London diphthong system. In *Language Variation and Change*, 20 (2008), Cambridge University Press, pp. 451-491.

Kotsinas, Ulla-Britt (1992). Immigrant adolescents´ Swedish in multicultural areas. In: Palgren, Cecilia, Lövgren, Karin & Bolin, Goran (eds.). *Ethnicity in Youth culture*. Stockholm. Stockholms Universitet, pp. 43-62.

Kotsinas, Ulla-Britt (1998). Language contact in Rinkeby, an immigrant suburb. In: Androutsopoulos, Jannis & Scholz, Arno (eds.). *Jugendsprache – lague des jeunes – youth language*. *Linguistische und soziolinguistische Perspektiven*. Frankfurt a.M. Peter Lang (=VarioLingua 7), pp. 125-148.

Keim, Inken (1978). *Gastarbeiterdeutsch - Untersuchungen zum sprachlichen Verhalten türkischer Gastarbeiter*. TBL Verlag Gunter Narr, Tübingen.

Krivokapic, Jelena, Fuchs, Susanne & Jannedy, Stefanie (2010). Prosodic Boundaries in German: Final Lengthening in Spontaneous Speech. Poster presented at the Meeting of the *Acoustical Society of America* in Baltimore, MD.

Mertins, Micaela (2010). *Phonetische Besonderheiten im Kiezdeutsch. Der palatale Frikativ*. MA-Thesis, Humboldt Univ. of Berlin.

Niedzielski, Nancy (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18(1): 1-18.

Nortier, Jacomine (2001). „Fawaka, what´s up?" Language use among adolescents in Dutch mono-ethnic and etnically mixed groups. In Hvenekilde, Anne & Nortier, Jacomine (eds). Meeting at the Crossroads. Studies of Multilingualism and Multiculturalism in Oslo and Utrecht, 61-73. Oslo. Norvus Forlag.

Paul, Kerstin (2008). Grammatische Entwicklungen in multiethnischer Jugendsprache. Untersuchung zu Verwendung und Funktionen von *so* in Kiezdeutsch. Universität Potsdam.

Pfaff, Carol W. (1981). Incipient Creolization in Gastarbeiterdeutsch? An experimental sociolinguistic study. *Studies in Second Language Acquisition* 3:2.165-78.

Quist, Pia (2005). New speech varieties among immigrant youth in Copenhagen – a case study. In: Hinnenkamp, V. & Meng, K. (eds.) Sprachgrenzen überspringen. Sprachliche Hybridität und polykulturelles Selbstverständnis. Tübingen, Gunter Narr, 145-161.

R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, version 2.10.1. http://www.R-project.org.

Torgersen, Eivind, Kerswill, Paul, & Fox, Sue (2006). Ethnicity as a source of changes in the London vowel system. In F. Hinskens (ed.) *Language variation – European perspectives*. Selected papers from the third international conference on language variation in Europe (ICLaVE3), Amsterdam: Benjamins, pp. 249-263.

Umbach, Carla & Ebert, Cornelia (to appear). German demonstrative *so* – intensifying and hedging effects. *Sprache und Datenverarbeitung* (*International Journal for Language Data Processing*). (Draft Version 2009).

Wiese, Heike (to appear). The role of information structure in linguistic variation: Evidence from a German multiethnolect. To appear in Gregersen, Frans , Parrott, Jeffrey & Quist, Pia (eds.) *Proceedings ICLaVe5,* John Benjamins Publishing Company.

Wiese, Heike (2009a). Grammatical Innovation in multiethnic urban Europe: New Linguistic practices among adolescents. *Lingua* 119, 782-806.

Wiese, Heike, Freywald, Ulrike, Özçelik, Tina & Mayr, Katharina (2009b). Kiezdeutsch as a Test Case for the Interaction between Grammar and Information Structure. Interdisciplinary Studies on Information Structure (ISIS) 12. Universitätsverlag Potsdam.

How Intraoral Pressure Shapes the Voicing Contrast in American English and German

Blake Rodgers
University Wisconsin-Madison

Susanne Fuchs
Zentrum für Allgemeine Sprachwissenschaft, Berlin

This study examines intraoral pressure for English and German stops in bilabial and alveolar place of articulation. Our subjects are two speakers of American English and three speakers of German. VOICING is the main phonological contrast under evaluation in both word initial and word final position. For initial stops, a few of the pressure characteristics showed differences between English and German, but on the whole the results point to similar production strategies at both places of articulation in the two different languages. Analysis of the pressure trajectory differences between VOICING categories in initial position raises questions about articulatory differences. In the initial closing gesture, time from start of gesture to closure is roughly equivalent for both categories, but the pressure change is significantly smaller on average for VOICED stops. Final stops, however, present a more complicated picture. German final stops are neutralized to a presumed VOICELESS phonological state. English final /p/ is broadly similar to German /p/, but English /t/ often shows no pressure increase at all which is at odds with the conventional account of phonation termination via pressure increase and loss of pressure differential. The results raise the question of whether the German final stops should be considered VOICELESS or some intermediate form, at least as compared to English final stops.

1 Introduction

The aims of this study are to investigate the phonetic realization of the phonological voicing contrast in word initial and final position in American English and German by means of temporal acoustic measures in combination with intraoral pressure. A secondary aim is to develop a set of measures to enable automated measurement of pressure characteristics.

In American English and German, both VOICED[1] and VOICELESS stops may be realized without phonation during closure in word initial position (for German: Jessen 1998, Fuchs 2005). For word final position, German is particularly interesting as it is the prototypical language for final neutralization of stop VOICING in coda position (Brockhaus 1995), while English maintains a clear acoustic distinction in this position.

Intraoral pressure is a crucial parameter for phonation because it results from both laryngeal and supralaryngeal mechanisms which are difficult to study in combination. Therefore, the focus here will be on intraoral pressure changes and temporal acoustic measures to characterize word initial and word final stops in these two languages.

Moreover, word position in the sentence will be evaluated by comparing tokens in sentence internal position with tokens in sentence final position. We expect that sentence position could be a crucial factor influencing voicing contrast. First, glottalization or glottal stops are often found at the end of a sentence (Kohler 2001). Second, subglottal pressure, and consequently intraoral pressure, may be reduced due to the air consumption at the end of a sentence. This should in turn strongly influence the transglottal pressure difference and hence, phonation patterns.

In the following sections, we will first describe the phonetic realization of the voicing contrast in American English and German in more detail. Second, we will provide a brief overview of the literature on final devoicing. Third, we will focus on the role of intraoral and subglottal pressure with respect to phonation and finally, we will present the working hypotheses of the current study which are based on the previously reported results from the literature.

1.1 Phonetic realization of the voicing contrast in American English

Kent and Moll (1969) noted that the difference between VOICED and VOICELESS categories for stops was due to volume changes and the timing of the glottal opening relative to the upper vocal tract gestures. Specifically, they found that for VOICED stops, vocal tract volume was higher due to expanded pharyngeal walls and lowering of both the hyoid bone and the larynx. Labial and lingual motion for bilabial and apical stops was found to be similar for both categories.

Westbury (1983) observed a collection of articulatory differences between VOICED and VOICELESS stops in American English. VOICED stops were

[1] We used uppercase following Docherty (1992) to indicate phonological contrast and not actual presence or absence of phonation. 'VOICING' is used to designate the abstract phonological category.

generally characterized by an increase in supralaryngeal volume, and VOICELESS stops occasionally had a decrease in volume. The larynx was always in a lower position for VOICED relative to VOICELESS stops. The tongue root was more advanced for VOICED stops with the exception of utterance initial /b/.

Perkell (1969) found that the larynx rose for /t/ as compared to /d/, which is consistent with the volume observations by Westbury (1983). The pharynx began more constricted for /d/ and then expanded to a greater volume as compared to /t/ during closure.

These studies taken together indicate that the volume-expanding gesture seems to be an integral part of VOICED stops in English. This gesture will give rise to at least two of the acoustic parameters that signal VOICING in English, namely voicing during closure and how rapidly phonation is initiated or terminated. The studies also indicate that the expansion gesture can be accomplished with several different strategies.

1.2 Phonetic realization of the voicing contrast in German

In his comprehensive acoustic study on the realization of the voicing contrast in German, Jessen (1998) described a series of peculiarities related to the occurrence of the obstruent in a given word (Note that Jessen called the phonological voicing contrast a tense-lax distinction which would be primarily based on durational differences). In his study he considered three positions and found that the stops /b d g/ and /p t k/ can be distinguished primarily by aspiration duration rather than voicing in utterance initial stressed position (##_V) and also in the post-voiceless position (#_V). The phonologically voiced stops /b d g/ are often devoiced in these positions. Devoiced is a phonetic term which describes a token that is realized without significant voicing during oral closure. Different degrees of devoicing can occur. The stops /b d g/ and /p t k/ can be distinguished by voicing during closure and by aspiration duration in intervocalic V_V position. In Fuchs (2005) the articulatory realization of the voicing contrast was investigated in word initial, word internal and word final position. Her data provided evidence that glottal abduction was mainly a property of VOICELESS stops in word initial position. In all other positions, supralaryngeal correlates and temporal differences played the major role.

To summarize, the production of the voicing contrast in German varies consistently with the position of the obstruent in the word. Voicing contrast may be a somewhat misleading term, since in word initial position the contrast is primarily based on aspiration. Voicing during closure may occur rarely.

1.3 Final devoicing

Final obstruent devoicing or so called "Auslautverhärtung" in German has been a major issue in phonological debates and one of "most popular of German phonological rules" (Giegerich, 1989, p.51). The phonological rule itself refers to the loss of voicing for VOICED obstruents in word, morpheme or syllable final position, for example:

Orthography and English translation	Phonetic transcription
Rad (wheel) vs. Rat (advice)	[ʀa:d̥] vs. [ʀa:t]
lies (read!) vs. ließ (let)	[li:z̥] vs. [li:s]

Historically it seems to have emerged in the transition between Old High German and Middle High German.

Final devoicing has been explained as feature changing in the Sound Patterns of English (SPE) tradition, reduction or spreading (Mascaró 1987) and licensing (Lombardi 1991, 1995, 1999) (for a comprehensive overview, see Brockhaus 1995). From a phonetic standpoint final obstruent devoicing is related to the question of whether VOICED stops are produced in similar fashion to their VOICELESS counterparts (full neutralization) or some differences can be observed (partial neutralization). The question arises as to how much phonetic detail should be taken into account to speak for one or the other. Several phoneticians investigated the issue of neutralization by means of acoustic data. Evidence was provided in both directions – for full (e.g. Fourakis and Iverson 1984) and for partial neutralization (e.g. Port and Crawford 1989, Port and O'Dell 1985), and the debate became more and more centered on methodological issues. Results were dependent on statistical methods, speech corpora (read versus natural speech, isolated words versus words in a frame sentence, nonsense words versus real words, frequently occurring words versus not frequently occurring words) and experimental set-ups (the subjects were/were not aware of the research question, hyperarticulated versus more natural speech conditions etc.). Regional variations could also be found, e.g. South German speakers clearly maintain the contrast (Piroth & Janker 2004).

It is challenging to get a comprehensive picture of all articulatory motions involved in the voicing contrast. A more indirect measure directly driving the phonation patterns may be to investigate intraoral pressure, since it is the consequence of the orchestra of articulatory motions.

1.4 Intraoral pressure and the voicing contrast

Measuring changes in intraoral pressure are of particular advantage to investigate the voicing contrast, since they directly mirror the combination of the manifold articulatory actions at the laryngeal and supralaryngeal levels. For instance, intraoral pressure will automatically increase when the vocal tract is closed by the lips and the velar port. Moreover, it will increase rapidly up to the level of subglottal pressure (ceiling effect) when the glottis is open and there is continuous air supply from the lungs. In contrast, intraoral pressure may be relatively low when the glottis is closed or the supralaryngeal articulators are moved in such a way that the oral cavity is enlarged (so called cavity enlargement). Various articulatory actions are reported in the literature (for a review, see Westbury 1983, Fuchs 2005) which describe the enlargement of the oral cavity leading to the same acoustic result – the maintenance of voicing. These processes are known under the term ‘motor equivalence’. They are often speaker- or even phoneme-specific and can make an investigation of the voicing contrast difficult. The orchestra of articulatory maneuvers is directly reflected in the changing intraoral pressure and the respective phonation. Phonation is primarily driven by the transglottal pressure difference (difference between subglottal and intraoral pressure) and the vocal fold tension. There are few studies investigating intraoral pressure measures with respect to the voicing contrast.

Müller and Brown (1980) investigated intraoral pressure and the voicing contrast in American English and developed a measure of the concavity of the intraoral pressure increase that compares initial slope of the pressure increase to final slope. [alpha] is the initial slope and [beta] is the final slope. The measure [delta] is defined as [alpha] – [beta]. [delta] greater than zero implies a convex shape, while [delta] less than zero implies a concave shape. The authors found that VOICELESS stops had higher values of [delta] than VOICED ones. They also found, however, that in many cases, it was difficult to define this parameter due to irregularly shaped pressure curves.

Müller and Brown (1980) built on the vocal tract pressure model developed by Rothenberg (1968), and concluded that the only way to achieve a convex pressure trajectory was to include an active volume expansion in the model. Koenig and Lucero (2008) observed that intraoral pressure can provide insights into such expansion gestures, if present, and more broadly provide insights into pressure control for the purpose of maintaining voicing during stop closure.

1.5 Subglottal pressure and the voicing contrast

Since air is consumed during speech production, subglottal pressure typically decreases over longer periods of speech. A lower subglottal pressure at the end of a sentence in comparison to the beginning or middle of a long sentence should also affect the upper limit of the intraoral pressure in obstruent production, since intraoral pressure equalizes subglottal pressure when the glottis is open. Moreover, the transglottal pressure difference crucial for phonation may be compromised too, if the intraoral pressure is influenced by vocal tract closure and does not decrease at the same rate as subglottal pressure. Thus, one could suppose that devoicing of VOICED phonemes may occur even more frequently or more rapidly in sentence final position due to the lower subglottal pressure. Interactions with other mechanisms, like the decrease of vocal fold tension in sentence final position, may be plausible as well, and will influence voicing during closure.

1.6 Assumptions and predictions

1. Since previous studies of American English and German reported no differences with respect to phonation in word initial stops, we predict that the pressure increase (slope) and pressure peak will behave similarly for VOICED and VOICELESS stops and for German and American English in syllable initial position.

2. If German word final stops are, in fact, neutralized, we assume no statistical difference between the two voicing categories. However, from a phonetic point of view, it may still be possible that some minor residue of the voicing contrast remains which could speak for a partial neutralization from a speech production point of view.

3. For American English we expect the maintenance of the voicing distinction in word final position with a shallower intraoral pressure rise and lower pressure peak for the VOICED stops and a steeper pressure rise and higher pressure peak for the VOICELESS ones. In addition, we suppose that if German final stops neutralize, they will resemble the English final VOICELESS stops.

4. We expect that sentence position will influence the realization of the voicing contrast.

2 Methodology

2.1 Instrumentation

To record intraoral pressure (IOP), an experimental setup was designed whereby a piezoresistive[2] pressure transducer Endevco 8507C-2 measuring about 2.4 mm in diameter and 12 mm in length was affixed to the rearmost portion the hard palate. The sensor measures the difference between intraoral and atmospheric pressure, the latter being obtained via a tube passed through the teeth. This arrangement offers several advantages: It permits simultaneous recording of acoustics and intraoral pressure, it is not affected by saliva blocking the tube, it is more comfortable for speakers than inserting a tube or catheter through the nose, and it allows recording of a high quality acoustic signal in comparison to a Rothenberg mask. Data was collected for bilabial, alveolar and velar stops. Data for velar stops was found to be unreliable due to the placement of the sensor in the vicinity of velar occlusion. Pressure can rapidly drop to zero if occlusion occurs aft of the sensor, so velar data was not included in the analysis.

2.2 Speakers and speech material

All together four German and four American English speakers (two males and two females per language) were recorded at the phonetics laboratory at ZAS, but only 3 American speakers and 2 Germans have been analysed so far. All German speakers are from the northern parts of Germany and Berlin. The American English native speakers are from different areas in the US (California, Missouri, Delaware).

Table 1: Speech material of American English target words

	Word initial (onset) / Word final (coda)				
	_a	**_u**	**_i**	**_ae**	**_ʌ**
p	pop	poop	peat / deep	pack / tap	pup
b	bog / gob	boot / tube	bead / hebe	bat / cab	butt / tub
t	top / cot	tube / boot	teak / peat	tap / bat	tub / butt
d	dock / god	dude	deed	dad	duck / cud

Monosyllabic target words were selected to investigate the voicing contrast in onset and coda position of American English (see Table 1). In one case no real

[2] This device changes its electrical resistance in response to changing mechanical pressure, which allows measurement of a voltage output that is proportional to intraoral pressure.

word existed in the lexicon, so a phonotactically comparable word was created ('hebe').

Moreover, the VOICED and VOICELESS stops were followed by different vowel contexts, but these will not be investigated at this point. All target words (X) occurred two times in a frame sentence, once in sentence medial position and once at the end of the sentence ("Get THIS new X for me, no, get THAT new X."). Sentence focus was placed words other than the target words in order to avoid an enhancement of the voicing contrast due to emphasis. The emphasis was placed on the two words written in capital letters (THIS and THAT).

For the German speech material we tried to make the corpus as similar as possible to the English characteristics. However, monosyllabic words are not as frequent in German as in American English, so it was sometimes difficult to find real words. Three vowels were chosen, both tense and lax, for a total of six vowels. The target words are listed in Table 2. All target words were embedded in the frame sentence: "Ich meine ZWEI X am Morgen, nicht DREI X". (I mean TWO X in the morning, not THREE X) with the accent on TWO and THREE. We did not find any combinations of lax vowels ending with a voiced stop, except from Klub (which is a loan word from English club).

Table 2: Speech material of German target words (# = gap in this particular combination)

	Word initial (onset) / Word final (coda)					
	_a	**_u**	**_i**	**_A**	**_U**	**_I**
p	Paar/Paap	Pup	Piek/Piep	Pack/Kapp	Putz/Stupp	Pils/Kipp
b	Bad/Stab	Bub	Biest/Dieb	Bass/#	Bus/Klub	Biß
t	Tag/Tat	Tuch/Gut	Tief/Reif	Taft/Patt	Tusch/Schutt	Tisch/Pitt
d	Dart/Bad	Dur/Sud	Dieb/Lied	Dach/#	Duft/#	Ding

2.3 Labeling procedures and further calculations

The acoustic signal had a sampling frequency of 44 kHz and the pressure signal is sampled at 1375 Hz. The pressure signal is first low-pass filtered using a 61 tap FIR filter with a passband of 6.875 Hz. The filtered signal is used for calculating the derivatives and the slope of the pressure rise. The first and second derivative of this signal is determined, and the second derivative is filtered with the same low-pass filter. First derivative of pressure signal will be referred to as pressure velocity, and second derivative of pressure signal will be referred to as pressure acceleration.

Start and end of pressure excursion were determined by peaks in the pressure acceleration signal. For tokens that did not present a well defined peak, the start was marked as the first upward spike in the pressure acceleration, and similarly, the end was marked as the last upward spike. The markings in these cases were confirmed by inspection of the smoothed pressure curve and subjective determination of beginning or end of pressure excursion.

Time of closure was estimated using the zero crossing of the pressure acceleration. Müller and Brown (1980) marked closure as the time when oral airflow dropped to zero. They also observed that the closure event generally causes a change in the sign of the pressure acceleration from positive to negative. As we did not measure airflow in the present study, the second method will be used to estimate closure time. A small subset of VOICED tokens did not show a sign change, and closure time was estimated by visual inspection.

TF32 (Milenkovic 2005) was used for determination of voicing offset, voicing onset and release. For voicing onset and voicing offset, a high-pass filtered pressure signal was used in conjunction with the spectrogram to determine where voicing was present. Voicing was defined as presence of a coherent oscillation in the pressure signal. This method proved to be more sensitive than phonation data from either the acoustic plot or the spectrogram. In other words, pressure ripples were detected beyond either the acoustic ripples or the voicing bar. Release was marked as onset of frication noise when present, and subjective determination of a pressure knee or sudden drop in the pressure curve when frication noise was not present.

The pressure knee was defined as the peak for VOICED tokens and the end of the initial rise for VOICELESS tokens. The knee was defined in this way to enable a determination of the initial slope of the pressure curve. VOICED tokens typically (though not always) showed a steady rise in pressure until release. VOICELESS tokens typically showed a rapid rise to the voicing offset point (Figure 1).

Time points defined for measurement are (1) start of pressure excursion (t0), marked as peak in smoothed pressure acceleration curve, (2) closure (tclos), which was approximated by the point where the pressure acceleration curve changed sign (following Müller and Brown, 1980), (3) voicing offset and (4) release (trel), which was determined from the acoustic burst. The time point of the knee (tknee) was defined as the offset of voicing for VOICELESS stops and release time for VOICED stops. The knee measurement is intended to capture the first major break in the smoothed pressure trajectory. Pressure at knee relative to pressure at closure allows a rough determination of the overall initial pressure slope.

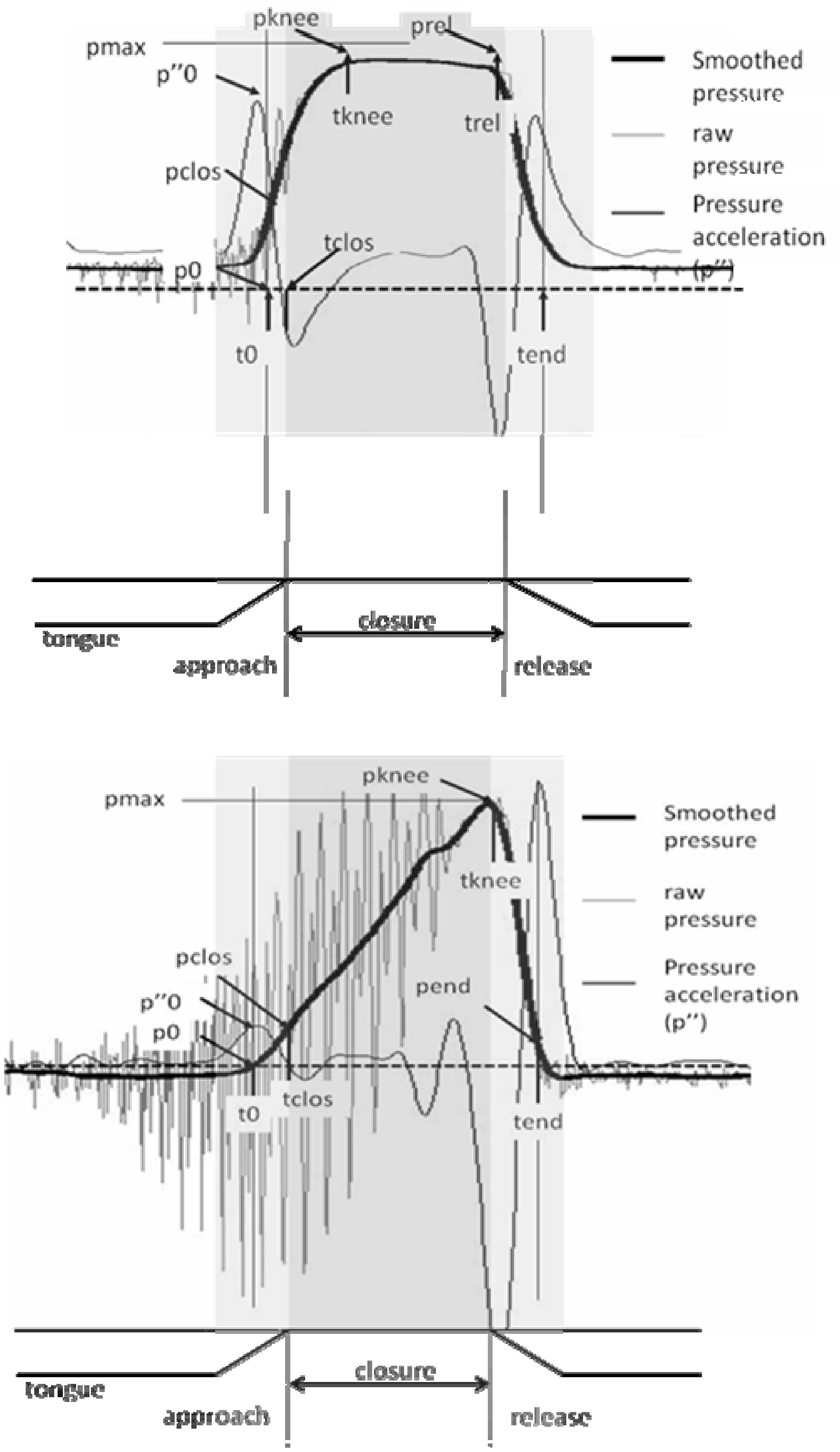


Figure 1. Pressure plots for word initial position a) VOICELESS stop and b) VOICED stop. Raw pressure signal is gray, and smoothed pressure signal is the heavy black curve.

2.4 Statistical design

The open software package R (R 2010) was used for statistical computing. Since voicing contrast is the major topic we are interested in and speakers may behave differently, we first split all the data according to speaker, position of the consonant in the word, place of articulation and sentence position and ran several t-tests with a temporal or pressure parameter as the dependent variable and voicing contrast as the independent factor. We considered all differences with $p \leq 0.05$ as significant, but additionally accounted for multiple comparisons by applying a Bonferroni correction (0.05/ number of tests run).

As a dependent variable we selected one of the temporal or pressure parameters, as fixed effects we chose voicing contrast, language, and sentence position and as random effects speaker and place of articulation. Random effects also have an influence on the variance in the data, but these effects are eliminated in the analysis. Data for word initial position will be presented first in the results section, followed by data for coda position.

3 Results

3.1 Initial stops

German and English stops in syllable initial position were found to be qualitatively similar, but there were significant differences in several pressure measures that will be discussed later. Initial VOICELESS stops are characterized by four phases. The first phase begins with a pressure increase. This is brought about by initiation of the tongue gesture possibly influenced by the glottal opening. Müller & Brown (1980:358) found in their modeling that the size of the glottal opening in the initial phase of VOICELESS stops influenced the slope of the initial pressure rise. Curvature of the pressure trajectory is positive until closure, which is the start of the second phase. At this point curvature becomes negative. Pressure velocity typically increases for the first portion of the pressure rise, and then decreases until phonation ceases. The third phase has no phonation. Pressure is relatively flat, but can increase or decrease slightly, typically at a steady rate. The fourth and final stage begins at release of occlusion. The pressure drops back to ambient, and phonation typically resumes as pressure nears ambient for VOICELESS stops. Figure 2 shows average pressure plots for four speakers.

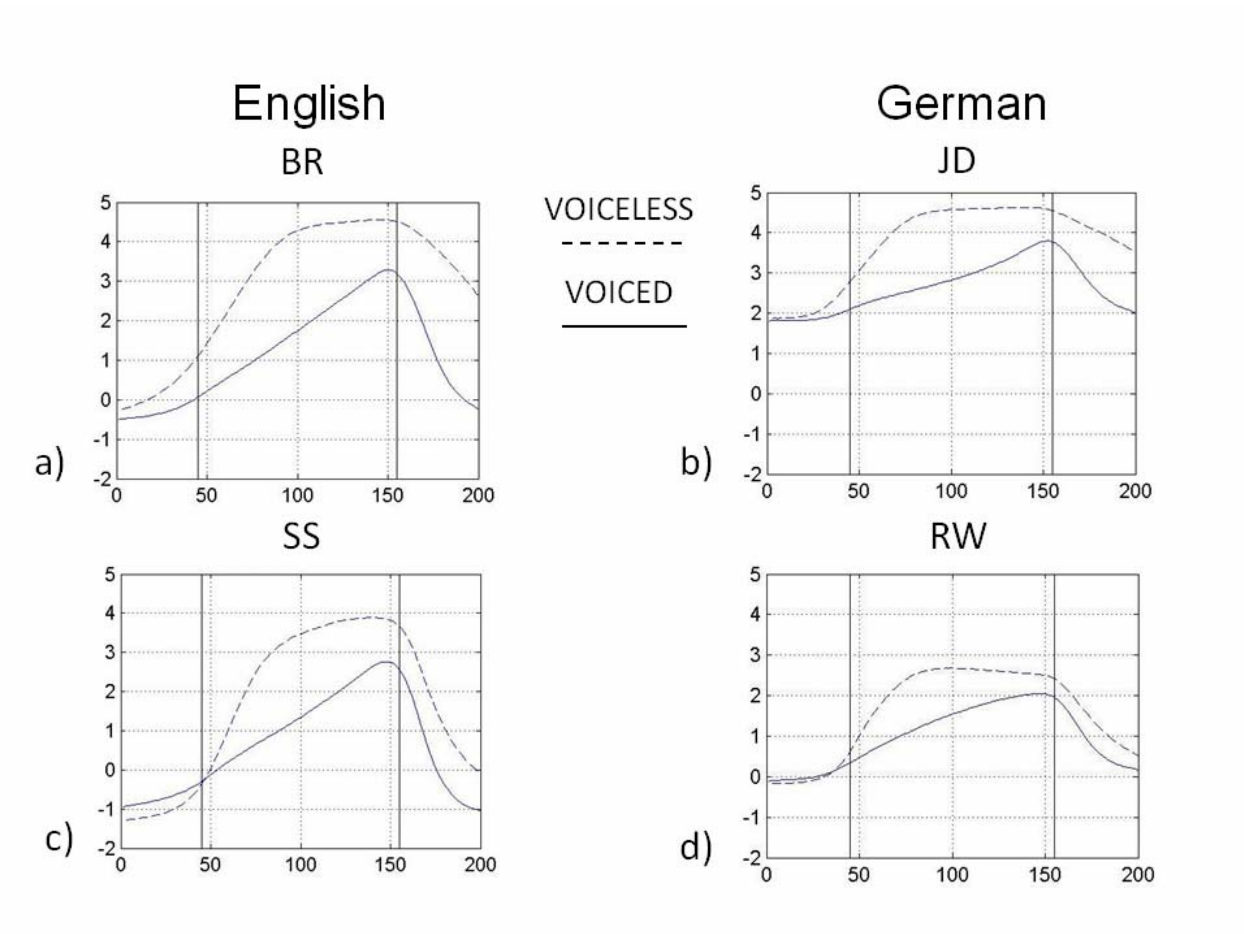


Figure 2: Average pressure plots for initial alveolar stops for individual speakers. Time alignment is at closure time and release time.

Speaker RW showed a consistently concave pressure plot for VOICED tokens, with the exception of a few samples. This can also be seen in his average plots, which are more concave for VOICED stops than those of the other speakers, English or German. This might suggest a weaker volume expansion gesture.

Initial VOICED stops are characterized by three identifiable phases. Like VOICELESS stops, they begin with the start of the tongue trajectory toward occlusion, and move to the second phase at occlusion. VOICED stops typically have no phase corresponding to the third phase of VOICELESS stops, instead having a second phase characterized by rising pressure until the third stage is initiated at release. The rate of pressure increase is smaller than that for VOICELESS stops, and the rate of pressure change is typically flat or decreasing. Phonation typically continues until release, at which point pressure falls back to ambient. In some rare cases, phonation stops for a brief interval prior to release.

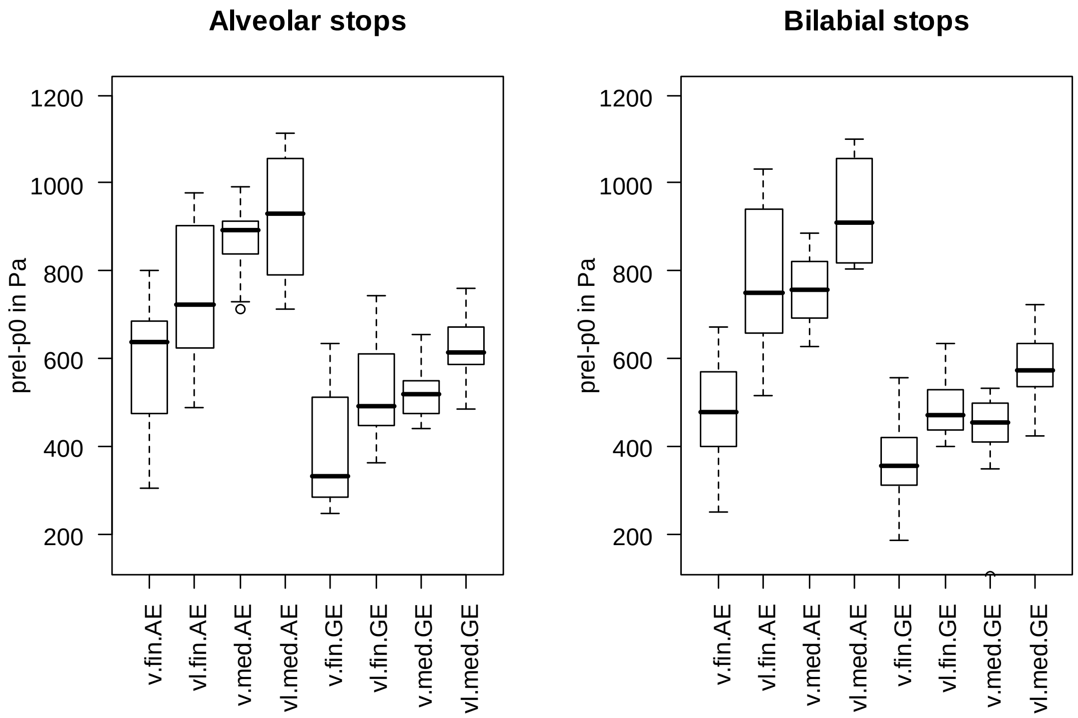


Figure 3: Release pressure relative to initial pressure.

Figure 3 shows the release pressure as compared to initial pressure. English speakers show consistently higher pressures for comparable stops. Figure 4 shows the maximum pressure acceleration. Interestingly, German speakers show higher values for this parameter, which is the opposite of what would be expected from the maximum pressure comparison.

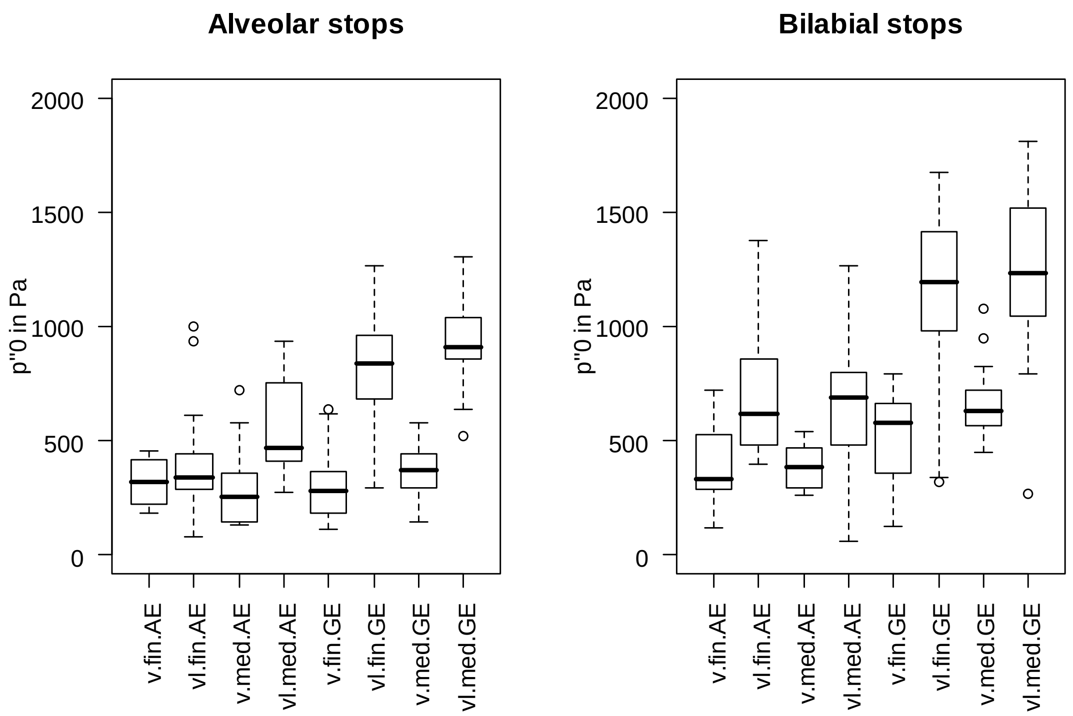


Figure 4: Pressure acceleration peak.

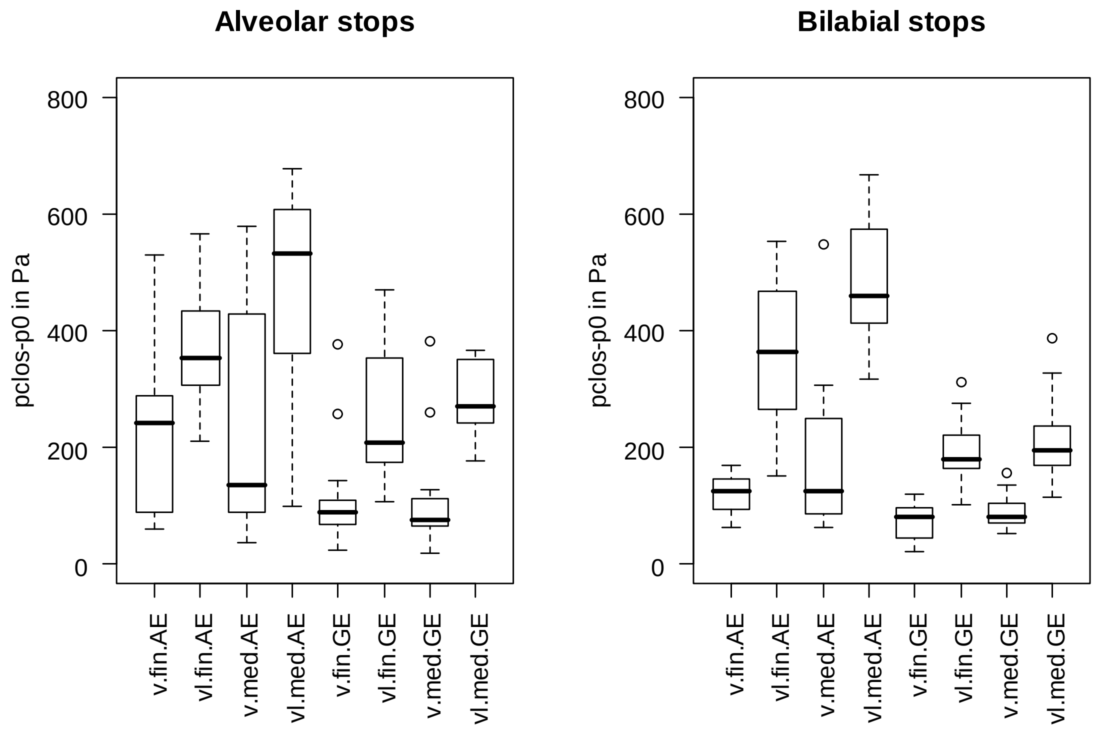


Figure 5: Close pressure relative to initial pressure.

Pressure at closure, as shown in Figure 5, shows that the English speakers are higher for this parameter, which is consistent with the lower pressure acceleration. Figure 6 shows the maximum pressure velocity. There are no clear trends between English and German. It does, however, show a clear difference between the VOICED and VOICELESS categories.

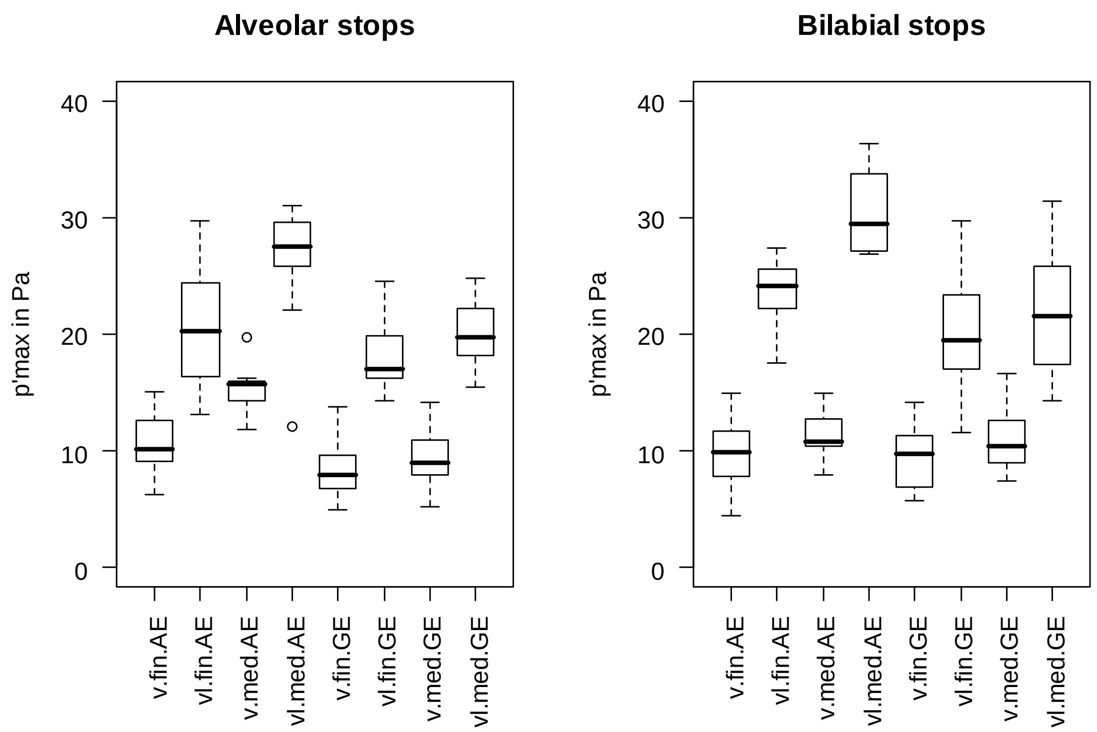


Figure 6: Maximum pressure velocity.

3.2 Final stops

There was a problem with the English carrier phrase for sentence internal final stops. The following fricative and its associated pressure increase completely changed the pressure characteristics. As a result, these tokens were not directly comparable to the other stops in the study, and so were excluded.

English final VOICED stops are characterized by significant voicing into closure and English final VOICELESS stops are characterized by termination of voicing often prior to pressure increase. Some tokens of final VOICELESS stops showed no intraoral pressure rise for the English subjects. Thus, for the latter group, a laryngeal mechanism for termination of voicing other than pressure increase is implicated. For many English final VOICELESS tokens, the voicing terminated prior to any pressure increase at all. This resulted in certain measures being undefined or aberrant for this category. These samples were excluded from analysis. For alveolar place of articulation, phonation terminated with no pressure increase for all tokens for speaker BR, and a majority of tokens for speaker SS. For bilabial place, phonation often terminated with no pressure increase. Speaker SS had 5 out of 8 tokens for final /p/ that were terminated with no pressure increase. Speaker BR had three tokens that seemed to terminate like the German final /p/, but the other five had much smaller pressure increases, and two of the five didn't actually terminate until after the pressure passed the peak. As a group, the English final /p/ tokens had much smaller pressure at the knee than the German tokens. The conclusion is that the English and German final VOICELESS stops are realized with different articulatory mechanisms, and possibly the two English speakers utilize different strategies as well depending on the adjacent vowel.

English final VOICED stops have pressure curves that are generally flat or convex during the rise, with rare instances of irregular or concave shapes. Pressure acceleration tends to hover around zero with several sign changes (Figure 7a,c).

German final stops are almost completely neutralized to a passive termination of voicing. The pressure rises due to occlusion, and voicing terminates when the pressure differential drops. German final stops, both VOICED and VOICELESS, are characterized by a concave rising pressure slope. This is confirmed by the pressure acceleration, which generally remains negative after closure. Voicing terminates near the peak of the pressure excursion. The statistical differences between VOICING categories are not significant (Figure 7b,d). Figure 8 shows the pressure acceleration peak at initiation of stop gesture. As with initial stops, German pressure acceleration values are consistently higher than American English.

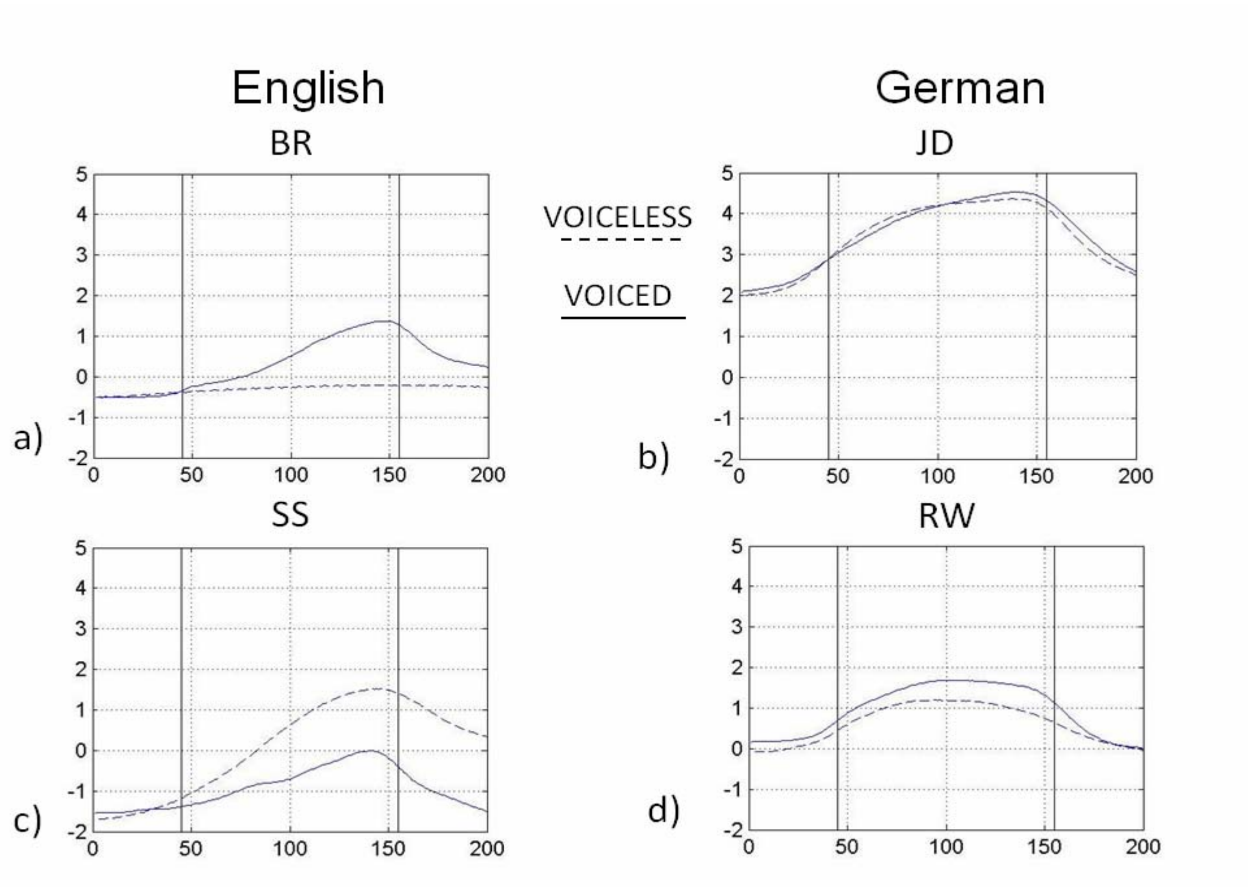


Figure 7: Average pressure for syllable final alveolar stops for individual speakers. Time alignment is at closure time and release time.

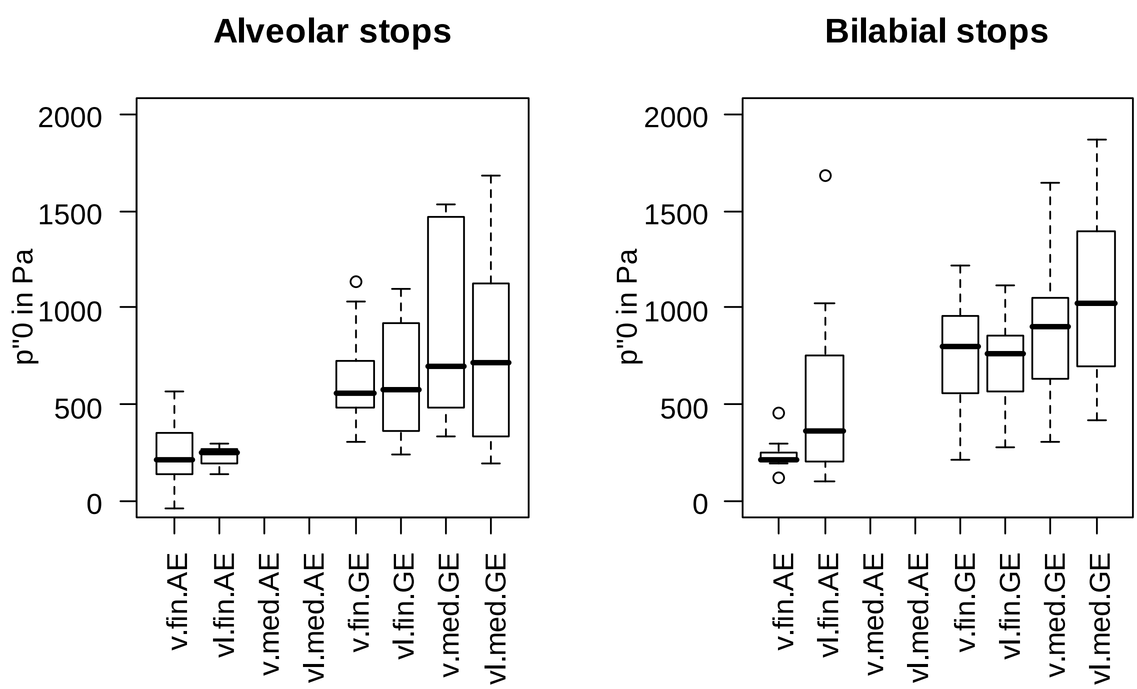


Figure 8: Maximum pressure acceleration for syllable final stops

4 Discussion

Müller and Brown (1980) performed simulations of pressure for intervocalic stops using a model developed by Rothenberg (1968). The model included vocal tract wall impedance, glottal area, supraglottal constriction area and volume of the supraglottal cavity. Running the simulation with fixed wall stiffness, fixed source pressure, laryngeal gestures that changed glottal area and supralaryngeal gestures that changed constriction area produced a pressure trajectory that was always convex (negative pressure acceleration) on the upslope. In order to get a concave or linear pressure trajectory, they found that it was necessary to include an active volume adjustment in the model. As the volume is increased, this will tend to lower the pressure acceleration, and in the return phase of the gesture to original state the pressure acceleration will be higher than otherwise.

Following this line of reasoning, all gestures for which the pressure curve has a linear or concave pressure rise will be assumed to have an associated volume expansion gesture. Absence of these characteristics, however, does not necessarily imply the absence of a volume expansion gesture.

4.1 Initial stops

The pressure curves for initial VOICED stops in English are generally flat or slightly concave. For the initial VOICED stops of the two German speakers, one showed a predominantly convex pressure curve, and one showed a predominantly concave pressure curve.

Using the zero crossing of pressure acceleration is presumed to be a reliable marker for closure except in the case of an active volume expansion gesture initiated prior to closure. It is unlikely that the difference in the pressure change from start of gesture to closure between VOICING categories is due to vocal tract wall compliance because the vocal tract is still open through this phase. The difference is presumably due to an active volume expansion during the closure phase, a difference in tongue shape (alveolar only), or a difference in the speed of articulation.

4.2 Final stops

Only the English VOICED final stops show a concave pressure curve, and therefore are assumed to have an active volume expansion. The English VOICELESS final stops appear to have two distinct pressure characteristic curves depending on preceding vowel, speaker, and place of articulation. One type shows little or no pressure increase at termination of phonation, and the other shows a small pressure increase similar to initial position. The neutralized

German final stops show a pressure characteristic that is consistent with passive termination of voicing. These stops are qualitatively different than either class of English final stop.

5 Conclusion

The pressure data allows a finer analysis of stops than acoustic data alone. There are broad differences in the small number of speakers considered between English and German. One mystery is the relationship between maximum pressure and maximum pressure acceleration. If the stops were produced in roughly the same manner, an increase of one of these parameters would seem to imply an increase in the other.

The data shows that English maximum pressure is higher, while German pressure acceleration is higher. It's not clear whether this implicates a different production mechanism or simply differences in articulator timing.

The pressure data shows that English final stops fall into at least two distinct categories. It also shows that German final stops, which are mostly neutralized, are qualitatively distinct from either VOICING category of English final stops. This raises the question of the phonological status of the final stops. From a phonetic standpoint, English VOICELESS final stops appear to have a gesture to terminate voicing, while VOICED stops appear to have a gesture to prolong voicing. The neutralized German stops seem to represent stop production with neither of these gestures.

One of the goals of this study was to identify a set of measures that allow automated comparison of pressure curves for stops. Pressure acceleration derived from smoothed pressure data allows marking of several time points including start of gesture (initial peak), closure (sign change), release (negative peak) and end of gesture (final peak).

References

Brockhaus, Wiebke (1995). *Final Devoicing in the Phonology of German*. Tübingen, Germany: Niemeyer.

Docherty, Gerard J. (1992). *The Timing of Voicing in British English Obstruents*. Berlin: Foris.

Fourakis, Marios and Iverson, Gregory K. (1984) On the incomplete neutralisation of German final obstruents. *Phonetica* 41: 140-149.

Fuchs, Susanne (2005). Articulatory correlates of the voicing contrast in alveolar obstruent production in German. *ZAS papers in Linguistics* 41.

Giegerich, Heinz J. (1989). *Syllable structure and lexical derivation in German*. Bloomington: Indiana University Linguistics Club.

Jessen. Michael (1998). Phonetics and Phonology of Tense and Lax Obstruents in German. (Series: Studies in Functional and Structural Linguistics 44, John Benjamins: Amsterdam.

Kent, Ray D., & Moll, Kenneth L. (1969). Vocal-tract characteristics of the stop cognates. *The Journal of the Acoustical Society of America,* 46(6B). 1549-1555.

Koenig, Laura L., & Lucero, Jorge C. (2008). Stop consonant voicing and intraoral pressure contours in women and children. *The Journal of the Acoustical Society of America,* 123(2), 1077-1088.

Kohler, Klaus J. (2001). Plosive-related glottalization phenomena in read and spontaneous speech. A stød in German? In Grønnum, N. & Rischel, J. (eds.), *To Honour Eli Fischer-Jørgensen. Festschrift on the Occasion of her 90th Birthday, February 11th, 2001* (Travaux du Cercle Linguistique de Copenhague 31), 174-211. Copenhagen: C. A. Reitzel.

Lombardi, Linda (1991). *Laryngeal Features and Laryngeal Neutralization.* Ph.D. dissertation, University of Massachusetts, Amherst. Published 1994 by Garland Publishing.

Lombardi, Linda (1995). Laryngeal neutralization and syllable wellformedness, *Natural Language and Linguistic Theory* 13: 39-74.

Lombardi, Linda (1999). Positional faithfulness and voicing assimilation in Optimality Theory, *NLLT* 17:276-302.

Mascaró, Joan (1987). A reduction and spreading theory of voicing and other sound effects. Barcelona, Universitat Autònoma. Unpublished manuscript. [Published 1995 in *Catalan Working Papers in Linguistics,* 4(2). 267-328.]

Milenkovic, Paul (2005). TF32 [Computer software]. Madison, WI: University of Wisconsin–Madison.

Müller, Eric M., & Brown, W. S. (1980). Variations in the supraglottal air pressure signal and their articulatory interpretation. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice, vol. 4*, Academic Press.

Perkell, Joseph S. (1969). *Physiology of speech production: Results and implications of a quantitative cineradiographic study*. Cambridge, MA: The MIT Press.

Piroth, Hans G. & Peter M. Janker (2004). Speaker-dependent differences in voicing and devoicing of German obstruents, *Journal of Phonetics* 32: 81-109.

Port, Robert & Crawford, Penny (1989). Incomplete neutralization and pragmatics in German, *Journal of Phonetics*, 17: 257- 282.

Port, Robert F. and O'Dell, Michael L. (1985). Neutralization of syllable-final voicing in German, *Journal of Phonetics* 13: 455-471.

R Development Core Team (2010). R: A language and environment for statistical computing. Vienna, Austria.

Rothenberg, Martin R. (1968). *The Breath-Stream Dynamics of Simple-Released-Plosive Production Vol. 6*. Basel: S. Karger.

Westbury, John R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *The Journal of the Acoustical Society of America,* 73(4), 1322-1336.

An Empirical View on Raising to Subject

Scott Grimm
Stanford University

This paper employs empirical methods to examine verbs such as *seem*, for which the traditional raising to subject analysis relates pairs of sentences which differ by taking an infinitival or sentential complement. A corpus-driven investigation of the verbs *seem* and *appear* demonstrates that information structure and evidentiality both play a determinate role in the choice between infinitival or sentential complementation. The second half of the paper builds upon the corpus results and examines the implications for the standard claims concerning these constructions. First, pairs of sentences related by the subject-to-subject raising analysis of verbs are often viewed as equivalent. New evidence from indefinite generic subjects shows that whether an indefinite generic subject occurs in the infinitival or sentential complement construction leads to truth-conditional differences. Further implications are explored for the claim that subjects of the infinitival variant may take narrow-scope: once various confounds are controlled for, the subject of the infinitival construction is shown to most naturally take wide-scope.

1 Introduction

The construction known as *subject-to-subject raising*, proposed for verbs such as *seem* or *appear*, relates pairs of sentences which differ by taking an infinitival or sentential complement. An example typical of those usually considered is shown in (1) (Davies and Dubinsky 2004, p. 3).

(1) a. Barnett seemed to understand the formula.
    b. It seemed that Barnett understood the formula.

Despite the prominent role this analysis has played in generative syntax, the empirical behavior of pairs of sentences related by the raising to subject analysis is relatively under-studied from an empirical standpoint, especially compared to the amount of recent work on the genitive and dative alternations in English (see representative studies Rosenbach (2002) and Bresnan et al. (2007), respectively). This paper focuses on establishing an empirical benchmark for the pairs of sentences at issue in the raising to subject analysis.

A guiding intuition has been that such pairs of sentences are truth-conditionally equivalent (Davies and Dubinsky 2004, p. 4), although this equivalence has recently become a subject of debate (see Lasnik 2003 for a denial of the equivalence and Boeckx 2001 for counter-arguments). In particular, it has been argued that the raising predicate in (1a), in parallel to (1b), has wide-scope over the situation designated by the proposition, which has served as motivation for analyses such as quantifier lowering (May 1977). Two other claims, which concern selectional behavior, are central to the raising analysis. First, a raising predicate uniformly selects for a proposition as its complement at the level of semantic selection, in the sense of, for instance, Grimshaw (1979), whether a raising predicate is followed by an infinitival or a clausal complement. Second, the raising predicate does not select for its subject, rather the subject is selected with respect to the proposition.

This paper will present empirical evidence that the behavior of verbs such as *seem* and *appear* is much more intricate than these standard claims allow for. While the pairs of sentences in constructed examples such as (1) often plausibly coincide in interpretation, data from actual language use manifest a rich set of differences in the use of one sentence type as opposed to the other, and at the same time demonstrate that none of the three claims above can be maintained unmodified. As the term "raising" presupposes a certain analysis of the constructions, I will avoid it in what follows, and refer to the two items as infinitival complementation (InfComp), as in (1a), and sentential complementation (SentComp), as in (1b), respectively.

The first three sections present results from a corpus study of the verbs *seem* and *appear*. In section 2, I demonstrate a distinction between the infinitival and sentential complement constructions in terms of information structure. The basic result is that the InfComp construction aligns with a topic-comment configuration, and the subjects of the InfComp are topics, while the SentComp is not so restricted, permitting structures not conforming to a topic-comment configuration. Section 3 examines a usage difference between the InfComp and SentComp in terms of evidentiality, demonstrating that the SentComp construction is strongly associated with indirect evidential readings, as opposed to the InfComp construction which is associated with direct evidential readings. Section 4 uses statistical modeling techniques to examine both information structure and evidentiality against other factors known to influence argument alternations. The latter half of the paper examines the implications of the empirical generalizations. Section 5 demonstrates that for sentences with indefinite generic subjects, whether the subject occurs in an InfComp or a SentComp construction leads to differences in truth conditions. Finally, I discuss the implications of the empirical results for the claim that *seem* always has wide-scope over a proposition, even within an InfComp construction, showing that once various confounds are controlled for, the subject of the InfComp most naturally takes wide-scope.

2 Information Structure Constraints

The information, or discourse, status of the referent of a noun, e.g. whether reference to an entity has been previously established or is somehow accessible to the hearer, has been held responsible for a large number of syntactic alternations, such as locative inversion or *there*-insertion (Ward et al. 2002). This section examines the influence of the information status of the subjects in both the InfComp and SentComp constructions.

2.1 Measuring Information Status

I tested the influence of information status on data gathered from the British National Corpus (BNC). I limited my study to two verbs: *seem* and *appear*. I extracted from the parsed BNC all occurrences of both *appear* and *seem* followed by *to* + *INF* or *that* + *CLAUSE*, from which I took a randomized sample of 200 occurrences for each pairing of verb and complement type, for a total data set of 800 instances. Later in the process, three instances of the InfComp and two of the SentComp were eventually discarded, as they were structurally inadequate, leaving a total of 795. For each of these instances, I automatically extracted the (embedded) subject, whose accuracy I verified later in the process.

I will first discuss the annotation scheme I used to measure the information status of the various subjects. After discussing the results of the corpus study, I will relate the findings to the notion of topicality in general, showing that these findings support considering the subject of the InfComp to be a topic.

2.1.1 The Annotation Scheme of Nissim et al. (2004)

For each sentence in my sample, I determined the status of the subject with respect to the discourse. I anchored this study in the annotation scheme implemented in Nissim et al. (2004). This scheme, however, was constructed to measure hearer-accessibility, which is restricted compared with the more general notion of topicality. Topicality covers not only accessibility of discourse elements from the hearer's perspective, but also from the speaker's perspective.

For instance, hearer-accessibility considers specific indefinites, e.g. *a certain X*, as discourse new, since they are unknown to the hearer; however, specific indefinites are capable of serving as topics and are clearly accessible to and identified as such by the speaker. I will use Nissim et al. (2004) as a basis for the corpus work, and return to examining the discrepancy between hearer- and speaker-accessibility at the end of this section.

The scheme of Nissim et al. (2004) subdivides *old*, *mediated* and *new* information into more fine-grained categories:[1]

Old: A nominal entity is considered discourse old if it has been previously mentioned, if it is a generic pronoun, or if it is a personal pronoun referring to the dialogue participants. Nissim et al. (2004) establish the subtypes: *identity* (co-reference with previously mentioned entity), *event*, *general* (dialogue participants), *generic*, *ident-generic* (co-referential with a generic entity), *relative* (relative pronouns).

New: An entity is new if it has not been previously mentioned and is not otherwise accessible to the hearer.

Mediated: "Mediated entities have not yet been directly introduced in the dialogue, but are inferable from previously mentioned ones, or generally known to the hearer" (Nissim et al. 2004, p. 1024). Nine subtypes are specified: *general* (culturally known entities, e.g. *the moon*), *bound* (bound pronouns), *part* (entities which stand in a part-whole relation to a previously mentioned entity), *situation* (entities which stand in a part-whole relation to, or are "evoked" by, a previously mentioned situation), *event* ("whenever an entity is related to a previously mentioned VP" (ibid. p. 1024)), *set* (when an entity is a sub- or superset of a previously mentioned entity), *poss* (entities which stand in a possessed-possessor relation to a previously mentioned entity), *function value* ("entity refers to a value of a previously mentioned function" (ibid. p. 1024)), *aggregation* (a conjoined entity which refers to a previously mentioned individual entity, or entities).

2.1.2 Results and Discussion

Both the InfComp and SentComp constructions admit instances of all the different categories, yet there are distributional asymmetries. The results of the corpus study showed a clear association between the subjects of the InfComp construction and elements made accessible in the preceding discourse, i.e. discourse-linked material, while subjects of the SentComp construction permitted material that was not linked to the preceding discourse.

[1] See Nissim et al. (2004) for details and heuristics for the annotation procedure.

This can be seen in table 1, which gives the totals for the occurrences of each construction by coarse-grained information status type. While both the InfComp and SentComp are relatively close in the number of mediated entities as subject, the SentComp has a greater proclivity toward discourse-new entities than the InfComp, and inversely concerning discourse-old entities. These distributional differences are statistically reliable (χ^2 : $p < .0001$).

Table 1: Coarse-Grained Information Status Distribution

	mediated	**new**	**old**	**expletive**	**Total**
SentComp	144	47	174	33	398
InfComp	124	13	234	26	397

Further analysis of the information status subtypes (see section 4) shows that genericity is an additional factor which correlates with the choice of complement. For both old-generic (generic pronouns/entities) and old-ident-generic (referring to previously mentioned generic entities) are primarily populated by instances of the SentComp construction.

2.2 Information Status and Topicality

The distribution of information status types across the SentComp and InfComp shows a compelling trend for elements linked to the preceding discourse to appear in the InfComp, with the exception of generics. Information status, however, belongs to the broader notion of topicality. In particular, since the information status notions used here are directed towards the information at the disposition of the hearer, this measurement was not able to capture ways in which the speaker may identify a nominal. In fact, most of the InfComps with discourse-new subjects, while hearer new, are clearly speaker-identified, and coincide with noun phrases which would still qualify as ‘aboutness topics’ (Reinhart 1981). For instance, (2) possesses a discourse new subject, *David Liddle*; however, this NP is the head of a relative clause, which would qualify it as specific and capable of being a topic (see discussion in Erteschik-Shir 2007, p. 48). Similarly, in (3), the *employer's combinations* does not appear in the preceding discourse, and it is difficult to relate it to anything else substantially enough to qualify for *mediated*, but the entity is prefaced by *certain*, typically indicative of a specific reading, and thus speaker-identified, which once again qualifies it as a topic.

(2) And indeed David Liddle, Director of Community Leisure in Avon, who is in charge of Avon County Libraries, seems to have taken this very line

at the same seminar.

(3) Certain kinds of employers' combinations seem to have been hindered by antitrust legislation and this may help to explain the generally low level of organisation among employers.

Similarly, there were nominal entities which, by the principles of the annotation scheme, counted as hearer-new, since they had not been previously mentioned nor could be said to be explicitly evoked by the previous situation in the discourse, yet, these nominals stood in contrast to the preceding topic, and so qualified as a contrastive topic. The following passage illustrates this situation, where the topic is the Zostera species, a type of sea grass, and the author brings up a previously unmentioned entity, green algae, as a point of contrast.

(4) Although present in Langstone Harbour there are now no beds of Zostera species in either Sussex estuary, although these certainly existed in the early years of this century, and Z. angustifolia was recorded near the Hayling shore of Chichester Harbour as recently as 1963. But **substantial beds of green algae** are present and appear to have spread in recent years; they are probably continuing to do so.

The above subject of an InfComp, then, qualifies as a specialized type of topic, a contrastive topic (see Erteschik-Shir 2007).

There was one legitimate instance of completely new discourse entity, yet its occurrence is probably to be taken with a grain of salt. The example in (5) is the opening line of a chapter of a novel and most likely serves a specific literary function.

(5) There was an explosion of blood and the nose seemed to burst.

As corpus results can often be biased by modality, e.g. written rather than spoken material, as well as particular annotators, I sought to independently verify the results of the annotation performed on the BNC. I examined the occurrences of *seem* in a version of the Switchboard Corpus (Godfrey et al. 1992) for which a portion was annotated for information status as part of the LINK project (based on Bresnan et al. 2002 and Zaenen et al. 2004). The annotation scheme used in the previous section is that which was developed for the LINK corpus, so the comparison was straightforward.

The number of occurrences of the SentComp construction was too few to draw any conclusions; however, the corpus yielded 45 occurrences of the InfComp construction which were marked for information status. The general picture established in section 2.1.2 emerged in the LINK corpus as well, the majority of occurrences were *old* (31), and the rest were *mediated* (14). Therefore, the findings from the LINK corpus, that subjects of the InfComp are linked to the preceding discourse and qualify as topics, concur with the findings of the corpus extracted from the BNC.

3 Evidentiality

One previously noted meaning difference between the InfComp and SentComp for raising verbs is that the "raised construction" appears to entail a perceptual experience of the "raised" element (Postal 1974, Langlacker 1995, and Asudeh & Toivonen 2007 for copy-raising). Postal (1974, p. 358), building on Cantrall (1970), gives examples such as the following:

(6) Julius Caesar strikes me as honest.
(7) It struck me that Julius Caesar was honest.

The verbs considered by Postal differ from those that are central to the investigation, namely *seem* and *appear*, but the generalization holds for these forms as well:

(8) Julius Caesar appeared (appears) to be honest.
(9) It appears that Julius Caesar is (was) honest.

In the examples from Postal, there is a discrepancy in tense; however, even normalizing that, it would be difficult to place (8) in a context where some sort of perceptual experience was not at issue.

This observation concerning perceptual experience would appear to be on the right track, yet it is not clear what type of perceptual experience counts as sufficient to guarantee the semantic well-formedness of a given sentence. In particular, subjects of InfComps need not be concrete, i.e. perceptible, objects and may be abstract nouns, as in (10).

(10) The argument appeared to make sense. (BNC G17)

Accordingly, a broader notion than the perceptual experience constraint is needed, and here I will reorient the perceptual constraint to connect with the domain of evidentiality, which supplies a contrast between direct and indirect evidentiality.

Direct evidential marking is used when "the speaker has some sort of sensory evidence for the action or event he/she is describing" (de Haan, 2004, p. 314). This can include visual, auditory or other means of directly witnessing the action, entity or event under consideration. Supplanting the perceptual experience constraint with a direct evidential one provides the generality needed to account for uses such as (10); further, this move relates the subject constraint to a clear function which is frequently grammaticalized cross-linguistically.

In contrast, indirect evidentials "are used when the speaker was not a witness to the event but when he/she learned of it after the fact" (ibid. p. 314). This includes two main subcategories of inference, "used when the speaker draws an inference on the basis of available, physical, evidence", and quotatives, "used when the speaker has been told about the action or event by another person", i.e. a hearsay function (ibid. p. 314).

Recasting the perceptual experience observation as a direct evidential function of the InfComp leads one to expect that this function of the InfComp would stand in contrast to a function of the SentComp, with which it alternates. It follows that the SentComp should possess a function of marking indirect evidentiality. Indeed, corpus work reveals that marking inferences, one of the subfunctions of indirect evidentiality, is highly visible in occurrences of SentComp constructs. I now turn to the results of a corpus examination.

3.1 Quantifying Evidential Marking

The hypothesis that the InfComp is associated with direct evidential uses while the SentComp is associated with indirect evidential uses leads to a straightforward corpus method for determining whether the hypothesis holds. Since indirect evidentials are the marked member of the direct/indirect contrast, one would expect to find the SentComp construction overtly displaying its inferential function by reporting the source of evidence. As direct evidentials are the unmarked form, it is difficult to code for direct evidentiality in an objective manner.

I tested the evidentiality hypothesis on the corpus described in section 2.1. The corpus revealed a strong tendency for the SentComp to explicitly mark an inference based on evidence, one extremely clear example of which is (11).

(11) On this evidence, then, it would appear that the UK is a service economy.

When a token contained an explicit indicator of evidential source and/or inference, I counted the token as positive for indirect evidentiality. The most frequent indicator was explicit mention of the evidence upon which the judgment was made in a *from*-phrase, as shown in (12). Other frequent indicators were *therefore*, *then* and *thus*. If the context made it exceptionally clear that a conclusion was being drawn, for instance as in scientific textual contexts as exemplified in (13), I marked such instances as positive.

(12) It also seemed, from the feathers on the kitchen floor, that one of the pigeons had come down for a warm and had got too close.
(13) Experiments suggest that it has a fluid consistency and that there is movement of the liquid molecules within the membrane. It would also seem that some proteins are free to move laterally within the membrane.

The source of evidence was signaled by a wide variety of grammatical means. Table 2 summarizes the different indicators of evidential source found in the SentComp sentences along with the number of tokens for each in the corpus. As over a third of the SentComp sentences had some marking for inference and/or evidential source, the association between SentComp and indirect evidentials appears to be quite strong.

Table 2: Distribution of Indirect Evidential Markings

as	2
because	2
consequent	10
evidential source	11
for phrase	5
from phrase	40
given that	2
legal conclusion	5
on phrase	5
scientific conclusion	15
then	12
therefore	19
thus	7
other	14
Total (out of 398 tokens)	147

3.2 Variability of Evidential Type for the InfComp

The direct evidential/perceptual constraint interpretation is consistent with the majority of the InfComp occurrences; however, some instances of the InfComp show inferential uses as well, as in (14).

(14) Some expert clauses have referred the issue not just to one expert but to

> two experts and an umpire. **From the cases, this** seems to have been common practice in the nineteenth century: see for instance Re Carus-Wilson & Greene (1886) 18 QBD 7, where there were two valuers and an umpire to value timber in a land sale. (BNC J6Y)

In the BNC corpus, I counted five instances of *appear* and one of *seem* as *indirect evidential*. In particular, when the complement is in the perfect, as in (14), the indirect evidential interpretation is found. Additionally, when the subject is not in the same location as the speaker, as in "Ed seems to be in his office", indirect evidential interpretation can arise.

There is a substantial range of types of evidentials between the extremes of direct visual evidence and abstract inference. For instance, inference based on first-hand evidence, which seems to be the usage in (14). Although evidentiality was operationalized above as a binary category, it is more properly viewed as a spectrum of multiple methods of observation and inference. Further work is needed to examine correlations between more specific types of evidential statements and complementation choice.

4 Controlling for Multiple Factors

At this point, two different strains of evidence have been presented that show a distributional asymmetry between the InfComp and SentComp constructions. Yet, it is increasingly understood that construction choice can be influenced by multiple factors, as argued, for instance, in the study of Bresnan et al. (2007) on the dative alternation. In order to examine whether other factors which have been known to have an influence on other constructions played a role here, I coded the data for two additional factors: subject length and nominal expression type, summarized below.

Subject Length Length has been long noted as an important factor in syntactic realization; for example, heavy NP shift places a longer constituent at the end of the clause (Behagel 1909, Wasow 2002). I measured the length of the subject, either of the matrix subject in the case of InfComps or the embedded subject in the case of SentComps, in number of words. In determining the final model, I also included the log of the number of words.

Nominal Expression Type The choice between the use of a pronoun and full noun phrase has been shown to affect syntactic realization, for instance in the case of the dative alternation (Green 1971, Bresnan et al. 2007). I divided the possible nominal expression types into the following categories PRONOUN (including definite, personal and reflexive pronouns, as well as demonstratives), DEFINITE, INDEFINITE (including phrases with the indefinite article as well as bare plurals), PROPER NAME, QUANTIFIER (such as *most*, *few*, *any*), RELATIVE PRONOUN, THERE, and VERBAL. As a factor in the eventual model, I additionally used a simplified metric coding for pronoun vs. non-pronoun.

Results and Discussion The influence of the above factors were examined using a generalized linear model, fitted to the data by using a stepwise model comparison, computed with the R statistical programming software. The following factors served as input to the model: length of the subject, the log of length of the subject, the coarse-grained information status types (old/mediated/new), the fine-grained information status subtypes, nominal expression type, (non-) pronoun and indirect evidential.

Four factors turned out to be significant, all of which increased the likelihood of the SentComp: INDIRECT EVIDENTIAL ($p < .0001$), the information status subtypes NEW ($p < 0.005$), OLD-GENERIC ($p < 0.05$), and OLD-IDENT-GENERIC ($p < 0.005$). The log of the subject length demonstrated a non-significant trend ($p = 0.067$), in the direction of the SentComp. Thus the information status categories and evidentiality are more predictive for these particular constructions than some of the better-known measures.

Having isolated the independent influence of topicality and evidentiality on the realization of the InfComp and the SentComp, it is incumbent upon a theory of verbs such as *seem* and *appear* to address why direct and indirect evidentiality functions should be associated with the InfComp and SentComp, respectively, rather than, for instance, the other way around. The previous sections, in fact, do suggest a motivation for this alignment. Direct evidential statements require an entity or event upon which the evidential statement is based, and this entity or event is presupposed with respect to the evidential statement. In the InfComp examples of *seem* and *appear*, the subject coincided with the entity that was the evidential source—for instance, in *Ed seems to be tired*, *Ed* serves as both the evidential source and the subject of the predication. If the evidential statement concerns a property of the entity which also serves as the evidential source, then this aligns with typical topic-comment structures—the entity or event is backgrounded (topic) and the property predicated of the entity is asserted (comment), and thus the InfComp construction is appropriate. In contrast, for statements based on abstract inference, there is no particular entity that is presupposed with respect to the evidential statement, i.e. no entity is necessarily perceived prior to being able to make an abstract inference. Therefore, there is no ready backgrounded candidate to align with a topic, and thus, the InfComp is less likely to be appropriate.

5 Indefinite Generics

Section 4 established that there was an asymmetry in the occurrence of generics with the two complementation types as generics occurred significantly more often with the SentComp. This section examines a further asymmetry in the acceptability of different types of generics in the InfComp construction. I will first differentiate the bare plural and the indefinite singular generics, and then examine how they behave within the context of the verbs and constructions at issue, showing that there are clear acceptability and truth-conditional differences dependent on whether an indefinite generic serves as the subject of *seem* or *appear* in the InfComp or the SentComp.

5.1 Bare Plural and the Indefinite Singular Generics

There is general agreement that one difference between the bare plural and indefinite singular generic constructions is that bare plurals may refer to a kind while indefinite singulars may not (see Krifka et al. 1995, Cohen 2001, and references therein). Sentences such as (15) are standard examples where something is predicated of a kind.

(15) Dinosaurs are extinct.

When the generic is expressed by an indefinite singular, rather than a bare plural, the statement is infelicitous on the kind reading, as in (16).

(16) ?A dinosaur is extinct.

This contrast is taken as evidence that bare plural, but not indefinite singular, generics may denote kinds (Krifka et al. 1995, Cohen 2001).

A second difference between bare plural and indefinite singular generics is the different interpretations they allow. The bare plural allows for an *inductivist* or a *normative* (or *definitional*) reading, while the indefinite singular only permits a normative reading, suited to describing conventions and definitions (Cohen 2001).[2]

An inductivist use of a generic statement is true "iff *sufficiently many relevant* individuals in the domain of the generic satisfy the predicated property. (Cohen 2001, p. 194)" For instance, (17) is true if there are enough (relevant) kings to satisfy the *sufficiently many* requirement.

[2] The discussion here follows Cohen (2001), which builds on Carlson (1995).

(17) Kings are generous.

Indefinite singular generics, if unmodified, do not permit the inductivist reading, as shown in (18).

(18) ?A king is generous.

If a statement such as (18) is modified so as to include quantification over situations, such statements can be rendered felicitous under the inductivist reading.

(19) A king is usually generous.

In contrast, the *normative* reading is most clearly in effect for statements of convention, such as (20) (Carlson 1995, p.225):

(20) Bishops move diagonally.

Normative generic sentences "do not get their truth or falsity as a consequence of properties of individual instances…instead, [they] are evaluated with regard to rules and regulations" (Cohen 2001, p. 194). If the generic sentence designates a rule which is *in effect*, then it is true, and otherwise false. Indefinite singular generics designate the normative reading, as shown in (21).

(21) A bishop moves diagonally.

On this account of generics, then, there is a truth conditional difference between the inductivist and normative readings. This truth conditional difference has been modeled in different ways. For instance, Papafragou (1996) in a modal treatment of generics distinguishes between bare plurals, which "require a realistic modal base", and indefinite generics, which may have modal bases which are not necessarily realistic. As an instance of the latter, Papafragou (1996) gives (22) as an example, which involves a deontic modal base which does not include the real world.

(22) A Christian is forgiving.

(22) clearly contrasts with (23), which on the preferred reading does say something about the real world.

(23) Christians are forgiving.

The contrast is summarized by Papafragou (1996, p. 17): "[(22)] does not make a statement about actual Christians but only about Christians in ideal (deontic) worlds; in fact, [(22)] does not even presuppose/imply the existence of Christians in the actual world, since the actual world does not belong to the modal base."

5.2 Generics and Construction Choice

Applying this distinction between bare plural and the indefinite singular generics to raising predicates reveals an asymmetry in acceptability of generics with the different complementation types. As shown in (24)-(25), indefinite singular generics with the InfComp, assuming neutral intonation and holding the generic interpretation constant, are unacceptable.

(24) a. A madrigal is polyphonic. (Cohen 2001)
b. ?A madrigal seems to be polyphonic.
(25) a. A pheasant lays speckled eggs. (Krifka et al. 1995)
b. ?A pheasant seems to lay speckled eggs.

Neither (24b) nor (25b) are interpretable while maintaining the definitional/normative reading. The more the generic statement hinges upon a convention, the worse these examples become:

(26) a. A bishop moves diagonally.
b. ?A bishop seems to move diagonally.

In contrast, when occurring in the SentComp, indefinite generics sentences maintain their definitional reading, as seen in (27)-(29).

(27) It seems that a madrigal is polyphonic.
(28) It seems that a pheasant lays speckled eggs.
(29) It seems that a bishop moves diagonally.

Indefinite singular generic subjects also appear freely as embedded subjects in naturally-occurring SentComp constructions, as witnessed in (30).

(30) It seems that a parent has a right to stalk their minor child, even if he is a

teenager. (from BOARDS.LP.FINDLAW.COM)

InfComp constructions with indefinite generics which give a deontic reading of the type discussed for (22), repeated here as (31a) are also infelicitous.[3] Once *seems* is inserted, the deontic force is no longer available.

(31) a. A Christian is forgiving.
 b. ?A Christian seems to be forgiving.
(32) a. A gentleman opens doors. (Burton-Roberts 1977)
 b. ?A gentleman seems to open doors.

The argumentation here should not be taken as saying that indefinite generics can under no circumstances be subjects of *seem* or *appear*, for of course they can, when their appearance is licensed by other factors. For instance, indefinite generics make excellent *contrastive topics*. Provided with the proper context, one can force acceptable readings of some of the above examples:

(33) A: Birds only lay pure white eggs.
 B: A *pheasant* seems to lay speckled eggs.

These occur naturally as shown in (34), where *health food shops* is the preceding topic:

(34) What always strikes me about health food shops are the rows and rows of bottles and tablets. **A greengrocer** seems to be a much better source of natural products than such collections of distilled essences and the like. (HTTP://WWW.FALLACYFILES.ORG/ADNATURE.HTML)

Additionally, when embedded in a larger structure, namely *when*- or *if*- clauses or other modalized environments, the acceptability of indefinite generic subjects may also be affected, as the example in (35) shows.

(35) We will ignore the fact that **a pronoun** seems to be a special kind of noun, a noun that refers to a previously understood antecedent. (WWW.LLRX.COM/COLUMNS/GRAMMAR7.HTM)

[3] Thanks to Cleo Condoravdi for pointing out this class of examples to me.

The claim here is that for the normative/definitional reading in neutral contexts, indefinite generic subjects occur without difficulty in the SentComp construction, but not in the InfComp construction.

Up until now, the focus has been on *seem* and *appear*, yet there are other verbs and adjectives analyzed as raising predicates, such as *tend* or *likely*. Parallel to *seem* and *appear*, the SentComp construction with *likely* does permit the normative reading. (*Tend* does not dispose of a SentComp construction.)

(36) It is likely that a bishop moves diagonally.

In contrast, indefinite singulars appear felicitously with both predicates:

(37) A bishop tends to move diagonally.
(38) A bishop is likely to move diagonally.

Yet, these predicates with the InfComp disallow the normative/definitional readings. To state that *X tends to V* implicates that this is not always the case—which stands in contradiction to the normative/definitional reading, and similarly for *likely*.

(37) and (38), however, are not as difficult to accommodate as *seem* and *appear* with indefinite singular generics in the InfComp construction, for they do permit the inductivist reading. As was seen with (19), repeated here as (39), if an additional modifier is present, such as *usually*, indefinite singular generic sentences become acceptable under the inductivist reading, for such adverbs permit quantification over situations. A similar reading is available for *tend*, as shown in (40).

(39) A king is usually generous.
(40) A king tends to be generous.

Tend would appear to provide implicit quantification over situations, as would *likely*, which normally implies previous observations in order to declare whether something is likely or not. Thus, the subjects of predicates such as *tend* and *likely* are evaluated under the inductivist reading, where reference is made to "relevant individuals in the domain of the generic satisfy the predicated property." These subjects, then, are not an instance of non-referential, non-topical subjects with the InfComp, but rather of induced inductivist readings.

The above data leads to a generalization concerning the purported uniformity of propositional content selected by raising verbs discussed in section 1. The InfComp construction disallows the normative reading—indefinite singular statements in the InfComp are either infelicitous or induce an inductivist reading. No such constraint holds for the SentComp construction. Since the inductivist and normative readings correspond to different truth conditions, the InfComp and SentComp cannot be truth conditionally equivalent in all cases.

5.3 Relating Generics and Evidentiality

At this point, the question turns to why such a contrast should be present. One avenue of explanation involves the interaction between indefinite generics and topicality. According to Cohen (2001), the different readings of generics are directly related to the notion of topicality. A central requirement of topics is that they must refer to an individual of some sort. Bare plural subjects may refer to kinds in statements of direct kind predication, as well as individual instances of kinds in the inductivist readings. Kinds are individual entities (Carlson 1977), and so in both cases, there is a specific entity, kind or individual, which serves as a topic. Indefinite singular generics, however, do not refer to kinds and do not refer to individuals, as they do not permit the inductivist reading. On the normative reading, the subject need not be anchored in any particular referent. Thus, the subjects of indefinite singular statements do not refer to any specific individual, and so do not meet the requirement on topics.

If topicality is associated with the subjects of the InfComp, this predicts that indefinite generic subjects should be quite marked as subjects of the InfComp. This is indeed the case, as manipulating examples of indefinite singular generics shows. While this is a plausible line of explanation, the initial assumption that indefinite generics are non-topics is not universally shared, as the inventory of grammatical elements which are able to be topics is widely debated. For instance, the information structure annotation scheme elaborated in Götze et al. (2007) explicitly codes indefinite generics as aboutness topics.

At the same time, the type of evidentiality also contributes to the issue. In fact, the same acceptability patterns observed with indefinite generics by altering construction type (InfComp vs. SentComp) can be obtained by altering evidential type, as show in (41).

(41) a. I saw that a king is generous. (only inductive)
b. I heard that a king is generous. (inductive or normative)

This contrast opens the possibility that the (un)acceptability of indefinite generics is not directly due to topicality, but to the modal base given by the evidential reading at issue. Visual evidence is naturally constrained to be from a realistic modal base, while indirect/hearsay evidence has no such constraint, a difference in modal base requirements corresponds exactly to the difference in modal bases required for the inductive reading (realistic) and the normative reading (non-realistic/ideal) as discussed by Papafragou (1996).

Much more remains to be explored concerning the connections between genericity and evidentiality. Yet, if the observations above hold, then the behavioral differences of the InfComp and the SentComp can be tied to a meaning difference grounded in the forms of evidentiality associated with each type of construction.

6 Scope

The strong association between the subject of the InfComp and topicality has broad implications for possible scopal relations with raising predicates. In particular, a central claim in the raising analysis of predicates such as *seem* is that these predicates are able to take wide-scope with respect to the subject of the sentence, whether its form is the InfComp or SentComp, and this has been achieved through various mechanisms, such as "reconstruction" or "quantifier lowering". This section demonstrates that the facts concerning scoping are more complicated.

The actual data and observations concerning scoping are consistent with the topicality of InfComp subjects. As topics are by definition backgrounded and therefore referential, one would expect the subjects of the InfComp construction to be backgrounded, and therefore have wide-scope over the material in the VP. Thus, establishing that the subjects of InfComp are topics leads to a prediction that is in direct opposition to the claim that *seem* is able to be interpreted as taking wide-scope over its subject, regardless of its syntactic structure. This section assesses this prediction.

6.1 Scopal Specificity

The verbs which have canonically been considered to be "raising" verbs, *seem* and *appear* as well as adjectives such as *likely* and *certain*, have also been noted to be part of the class of lexical items known as non-factives (Kiparsky and Kiparsky 1970). In contrast to factives, which presuppose the content of the embedded clause, as in (42), non-factives have no such presupposition, as in (43).

(42) It is odd that it is raining. (factive)
presupposes that it is raining
(43) It is likely that it is raining (non-factive)

does not presuppose that it is raining.

As non-factives permit an irrealis reading of their complement, they qualify as intensional verbs. One of the hallmarks of intensional verbs is that they permit both specific and non-specific readings of indefinites within their scope. The parade example is shown with the verb *want* in (44).

(44) I want a book.
⇒ There is a particular book such that I want it
(indefinite has wide-scope; the interpretation of the indefinite depends on the speaker's intended referent)
⇒ I want a book and, if there is one, any book will do
(indefinite has narrow-scope; the interpretation of the indefinite is bound by the intensional context introduced by *want*)

Since *seem* and *appear* are non-factive and induce irrealis contexts, both readings for indefinites should be available when within the scope of *seem* or *appear*. This is clearly true for the SentComp construction. In (45), the indefinite subject is non-specific with respect to the scope of *appear*; it is neither required to refer to a specific gun, nor to refer at all beyond the context established by *appear*.

(45) It appeared that a German S.P. gun had joined the snipers and was lobbing the occasional shell into the vicinity of the orchard. (BNC A61)
does not presuppose that there existed a particular German S.P. gun

The phrase could be felicitously continued by *But, in fact, these shells came from a tank on the other side of ravine*, and thus force the narrow-scope reading where *a German S.P. gun* does not refer outside of the context of *appear*.

Given the above observations, along with the interpretation traditionally claimed by the raising analysis, viz. that raising verbs always have wide-scope over the proposition, a testable prediction emerges: if raising verbs always have wide-scope, then indefinites should have both specific and non-specific irrealis readings available for both the InfComp and SentComp constructions. After discussing the notion of specificity in more depth, I will assess this claim.

6.2 Specifying Specificity

As specificity is a term used in a variety of senses, I begin with a few words about the sort of specificity intended here. The specific/non-specific distinction evoked by intensional predicates and non-factives has been termed *scopal* *specificity* (Farkas 1994, 2002). The difference between the two interpretations of the indefinite is determined by whether the indefinite is linked to the general context or to that context evoked by the intensional predicate: "…when scopally specific, the value of the indefinite is chosen from the domain of *w*, the world with respect to which the main clause is evaluated, while the value of scopally non-specific indefinites is to be chosen from the domain of worlds introduced by the predicate" (Farkas 1994, p. 4). In essence, the scopally non-specific reading is modal, determined by the set of possible worlds introduced by the predicate, while the scopally specific reading is actual, determined by prior discourse.

This sense of specificity must be kept distinct from others found in the literature, namely *epistemic specificity* and *partitive specificity*, exemplified in (46) from Fodor and Sag (1982) and (47) from Diesing (1992), respectively.

(46) a. A student in Syntax 1 cheated on the exam.
b. His name is John. (epistemic specific reading)
c. We are all trying to figure out who it was. (epistemic non-specific reading)

(47) a. There are some ghosts in this house. (partitive specific reading)
b. Some ghosts live in the pantry; others live in the kitchen. (partitive non-specific reading)

Epistemic specificity is distinct from scopal specificity—in (46a), no one doubts that there *is* a student who cheated, but there are just doubts about that student's identity, as the continuations of (46) show. As for partitive specificity, it too can be shown to be distinct from scopal specificity. Generally, a partitive use "denotes a member or subset of a familiar discourse group" (Farkas 1994, p. 8). Partitive non-specificity arises as to which member of the familiar set is being referred to. As in the case of epistemic specificity, there is no question that these members are in the world assumed by the speaker, in contrast to scopally non-specific readings. In the following, epistemic and partitive non-specifics are not at issue, it is the behavior of scopal (non-)specificity which will provide insight into whether InfComps and SentComps pattern similarly with respect to scope.

6.3 Assessing Scopal Specificity

One foundational fact about topics is that they are backgrounded, and therefore are referential with respect to the discourse. Section 2 established an association between the subject of the InfComp and topichood. Based on this finding, one would expect that the subject of the InfComp would align with the scopally specific reading and be unlikely to support the scopally non-specific reading. Naturally occurring examples provide evidence that the scopally specific/non-specific distinction is employed in a manner consistent with these expectations, as demonstrated by a pair of discourses concerning governmental proceedings. The first in (48) shows the InfComp associated with the scopally specific reading while the second in (49) shows the SentComp construction used to indicate a scopally non-specific reading.

(48) The Senate adopted an extension of the provision with little attention to the issue, and the House opposed the measure, although on the only vote taken on it, **a majority appeared** to support continuing the measure. (The vote scorecard is accessible here [link].) However, the issue was obscured by the fact that some who voted against the effort to kill 245(i) insisted they simply did not want to tie the hands of the conference committee members who would decide the measure's fate.
(HTTP://WWW.FAIRUS.ORG/SITE/PAGESERVER?PAGENAME=IIC_IMMIGRATIONISSUECENTERS326F)

In (48), the majority is actual, and can be verified by looking at the voting scorecard. However, what is undetermined is whether the majority intended to *support* the measure, or had an ulterior motive for voting as they did. Here, the subject of the InfComp is actual, i.e. scopally specific, while the complement of *appears*—the support for the measure—is merely potential. This differs from the SentComp structure in (49):

(49) A private member's bill to effect this change was brought forward in late February, 1925. Although there was some division of Conservative opinion, **it appeared that a majority** both of the Cabinet and of the backbenchers favoured either the acceptance of the bill or a Government measure doing roughly the same job in its place. Baldwin treated the matter with the utmost seriousness. He made a lot of soundings, appointed a special Cabinet committee to go into the subject, and held a special Cabinet to receive its report. (BNC EFN)

The context of (49) makes it is clear that the majority need not be actual—the votes have not taken place, and the main actor of the passage is busy sounding out the eventual voters.

Data such as the above would indicate that the subjects of the InfComp tend to be associated with the specific reading of indefinites, while subjects of the SentComp permit the scopally non-specific reading. Altering the SentComp example (45), repeated here as (50a), to (50b) shows that this distinction is at play.

(50) a. It appeared that a German S.P. gun had joined the snipers.
b. A German S.P. gun appeared to have joined the snipers.

The most natural reading, with neutral intonation, for (50b) is that there is a specific German S.P. gun.

Similar contrasts are visible with other predicates traditionally analyzed as raising predicates, as shown for *certain* in (51). (51a) is a SentComp corpus example which permits a continuation explicitly denying the subject. Modulating the form to the InfComp as in (51b) once again makes such a continuation inaccessible.

(51) a. Nobody knows exactly who built South Luffenham, but it is almost certain that an architect called John Sturges supplied the drawings. (BNC AB4)
…although some deny the architect's existence.
b. Nobody knows exactly who built South Luffenham, but an architect called John Sturges is almost certain to have supplied the drawings.
…?although some deny the architect's existence.

6.4 Lexical Semantic Confounds

The previous section has added empirical support to the hypothesis that the SentComp and InfComp are distinguished in scopal properties just as one would expect if the subjects of the InfComp construction are topics, i.e. these subjects scope wide. Yet, one claim that is repeated throughout the literature, at least since Montague (1973), is that raising constructions permit scopally non-specific (i.e. narrow-scope) readings of the subject. The standard examples are of the following type:

(52) A cat seems to be in the garden. [embedded verb of existence]
(53) A train seems to be approaching. [embedded verb of appearance]

The intuition that these examples purportedly prompt is that no particular cat is actually in existence, and therefore, these sentences qualify as scopally non-specific, and the entire phrase has a de dicto reading, equivalent to *It seems that a cat is in the garden*. While this intuition is relatively clear in the above examples, this could occur for independent reasons. The verbs used in such examples are from a restricted set—verbs of existence and appearance. When the phrases are minimally altered in the choice of the verb, the scopally non-specific reading becomes far less accessible, if not impossible, as witnessed by (54) and (55), where the most natural readings involve a specific cat and a specific train.

(54) A cat seems to be sleeping in the garden.

(55) A train seems to be leaving.

The contrast between (52) and (54) and between (53) and (55) indicates that perhaps verbs of existence and appearance have a peculiar behavior which explains what has been claimed as the apparently exceptional behavior of the raising verbs. Indeed, these two classes of verbs share two peculiarities. First, most all transitive verbs entail that the subject exists prior to the event designated by the verb. For example, in *John hit Bill*, to perform the act of hitting, John must exist before the onset of the event. Verbs of existence and appearance, however, do not entail prior existence. Verbs of existence assert existence, and therefore do not entail prior existence. Verbs of appearance only entail that the argument exists at the location at the *end* of the event, but do not entail existence prior to the end of the event. Since *seem* modifies a VP, for statements such as *seems to be arriving*, *seems* indicates that the occurrence of the event is in question. As the existence of the subject is dependent on the occurrence of the event, it is unsurprising that the actuality of the subject is in question as well.

A similar confound is visible in examples containing predicates of *negative existence*, as exemplified in (56).

(56) Someone seems to be absent from class. (Boeckx 2001, p. 512)

While these examples do seem to have a preference for a narrow-scope reading, this has previously been shown to be independent of the 'raising' predicate—these verbs result in intensional object readings in the absence of any raising verb:

(57) A screw is missing from this TV set. (E. Bach reported in Dowty 1985)

In summary, while the intuition that the subjects of *seem* in the InfComp may take narrow-scope is indeed plausible for select sentences, these sentences may prefer such a reading for independent reasons. Given the strong association between the subjects of the InfComp and topicality for these verbs, it is unlikely that such readings are often generally available, as the evidence from corpus examples shows.

7 Conclusion

In section 1, some of the standard assumptions about the raising analysis were sketched, namely (i) a raising predicate uniformly selects for a proposition at the level of semantic selection, (ii) the InfComp and SentComp are taken to be interpretationally equivalent—in particular, in both the raising predicate has wide-scope over the situation designated by the proposition, and (iii) the raising predicate does not select for its subject, rather, the subject is selected with respect to the proposition. In light of the generalization concerning the information structure properties associated with the InfComp and SentComp structures, it is improbable that *seem* or other such predicates select for propositions in a simple or unified manner. Not just any proposition is felicitous in the InfComp formulation, but only those conforming to a topic-comment structure, in contrast to the less restrictive requirements of the SentComp. The second claim, that the InfComp and SentComp are meaning equivalent and that the raising predicate always has wide-scope, were brought into question in sections 5, which pointed to instances of non-equivalent truth conditions between the two structures, and 6, which demonstrated a consistent association between the subject taking wide-scope and the InfComp construction, as well as between the subject of *seem* or *appear* taking narrow-scope and the SentComp construction. The third claim, that the raising verbs in the InfComp do not select for their subject, must also be attenuated. While the subjects of the InfComp do not discriminate in terms of thematic content, it appears that such verbs, at minimum, do select for topics, and for *seems* and *appear*, the subjects of the InfComp are most often constrained to be direct evidential sources.

The results connect to higher-level theoretical points as well. The raising to subject analysis goes counter to the assumption of *economy*, common in linguistic theory: if the raised and non-raised constructions are equivalent, it is puzzling why this particular optionality of expression should exist. Why should a language bother to enforce two manners of saying the same thing? From the perspective of the data in this paper, this particular puzzle never occurs, for the two constructions simply differ in their associated meanings and uses.

The new empirical generalizations presented in this paper demonstrate that verbs such as *seem* and *appear* have a rich host of associations that can be used by in sophisticated ways. Incorporating these associations within the theoretical analysis of these verbs remains a challenge for future work.

Acknowledgements

This work was presented at the 2009 Annual Meeting of the Linguistic Society of America, the Workshop on Linking of Sentential Arguments at the 2009 Deutschen Gesellschaft für Sprachwissenschaft, and at the Zentrum für Allgemeine Sprachwissenschaft in March, 2009, and I would like to thank the audiences of these conferences. This paper also greatly benefited from discussions with Eve Clark, Cleo Condoravdi, Hans-Martin Gärtner, Beth Levin, Asya Pereltsvaig, Ivan Sag, Kerstin Schwabe, Barbara Stiebels, and Tom Wasow. I would also like to thank in particular Beth Levin for acute comments on various drafts.

References

Asudeh, Ash & Ida Toivonen (2007). Copy raising and perception. Unpublished ms.

Behagel, Otto (1909). Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen*, 25/110.

Boeckx, Cedric (2001). Scope reconstruction and A-movement. *Natural Language and Linguistic Theory*, 19:503–548.

Bresnan, Joan, Jean Carletta, Richard Crouch, Malvina Nissim, Mark Steedman, Thomas Wasow & Annie Zaenen (2002). Paraphrase analysis for improved generation. LINK project, HRCR Edinburgh-CSLI Stanford.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & Harald Baayen (2007). Predicting the dative alternation. In: Gerlof Bouma, Irene Krämer & Joost Zwarts (eds.). *Cognitive Foundations of Interpretation*. Amsterdam: Royal Netherlands Academy of Science, 69–94.

Burton-Roberts, Noel (1977). Generic sentences and analyticity. *Studies in Language*, 1:155–196.

Cantrell, William (1970). Why do trains remind me of throwing up? Unpublished paper.

Carlson, Gregory (1977). *Reference to Kinds in English*. Ph.D. thesis, University of Massachusetts.

Carlson, Gregory (1995). Truth-conditions of generic sentences. In: Gregory Carlson & Francis Pelletier (eds.). *The Generic Book*. Chicago: The University of Chicago Press, 224–237.

Cohen, Ariel. (2001). On the generic use of indefinite singulars. *Journal of Semantics*, 18:183–209.

Davies, William & Stanley Dubinsky (2004). *The Grammar of Raising and Control: A course in syntactic argumentation*. Oxford: Blackwell.

de Haan, Ferdinand (2004). Semantic distinctions of evidentiality. In: Matthew Dryer, Martin Haspelmath, David Gil & Bernard Comrie (eds.). *World Atlas of Linguistic Structures*. Oxford: Oxford University Press, 314–321.

Diesing, Molly (1992). *Indefinites*. Cambridge, Mass.: MIT Press.

Dowty, David (1985). On Recent Analyses of the Semantics of Control. *Linguistics and Philosophy*, 8:291–331.

Erteschik-Shir, Nomi (2007). *Information Structure: The Syntax-Discourse Interface*. Oxford: Oxford University Press.

Farkas, Donka (1994). Specificity and scope. In: Léa Nash & George Tsoulas (eds.). *Langues et Grammaires 1*. Paris.

Farkas, Donka (2002). Specificity distinctions. *Journal of Semantics*, 19:1–31.

Fodor, Janet & Ivan Sag (1982). Referential and quantificational indefinites. *Linguistics and Philosophy*, 5:355–398.

Godfrey, John, Edward Holliman, & Jane McDaniel (1992). Switchboard: Telephone speech corpus for research and development. In: *ICASSP-92*.

Götze, Michael, Thomas Weskott, Cornelia Endriss, Ines Fiedler, Stefan Hinterwimmer, Svetlana Petrova, Anne Schwarz., Stavros Skopeteas, & Ruben Stoel (2007). Information structure. In: Stefanie Dipper, Michael Götze, & Stavors Skopeteas (eds.). *Working Papers of the SFB632*, *Interdisciplinary Studies on Information Structure* (ISIS). Potsdam: Universitätsverlag Potsdam, 147–187.

Green, Georgia (1971). Some implications of an interaction among constraints. In: *Papers from the Seventh Regional Meeting*, *Chicago Linguistic Society*. Chicago: The University of Chicago, 85–100.

Grimshaw, Jane (1979). Complement selection and the lexicon. *Linguistic Inquiry*, 10:279–326.

Kiparsky, Paul & Carol Kiparsky (1971). Fact. In: Danny Steinberg & Leon Jakobovits (eds.). *Semantics*. Cambridge: Cambridge University Press, 143–173.

Krifka, Manfred, Francis Pelletier, Gregory Carlson, Alice ter Meulen, Gennaro Chierchia & Godehard Link (1995). Genericity: An introduction. In: Gregory Carlson & Francis Pelletier (eds.). *The Generic Book*. Chicago: The University of Chicago Press, 1–124.

Langacker, Ronald (1995). Raising and transparency. *Language*, 71:1–62.

Lasnik, Howard (2003). *Minimalist Investigations in Linguistic Theory*. New York: Routledge Ltd.

May, Robert (1977). *The Grammar of Quantification*. Ph.D. thesis, MIT.

Montague, Richard (1973). The proper treatment of quantification in ordinary English. In: Jaakko Hintikka, Julius M.E. Moravcsik, & Patrick Suppes (eds.). *Approaches to Natural Language. Dordrecht*: D. Reidel Publishing Co, 221–242.

Nissim, Malvina, Shipra Dingare, Jean Carletta, & Mark Steedman (2004). An annotation scheme for information status in dialogue. In: *Proceedings of the 4th Conference on Language Resources and Evaluation (LREC2004)*.

Papafragou, Anna. (1996). On generics. *UCL Working Papers*.

Postal, Paul (1974). *On Raising*. Cambridge, Mass.: MIT Press.

Reinhart, Tanya (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, 27:53–94.

Rosenbach, Anette (2002). *Genitive Variation in English. Conceptual Factors in Synchronic and Diachronic Studies*. Mouton de Gruyter.

Ward, Gregory, Betty Birner & Rodney Huddleston (2002). Information packaging. In: Rodney Huddleston & Geoffrey Pullum (eds.). *Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press, 1363–1447.

Wasow, Thomas (2002). *Postverbal Behavior*. Stanford: CSLI.

Zaenen, Annie, Jean Carletta, Gregory Garretson, Joan Bresnan, Andrew Koontz-Garboden, Tatiana Nikitina, M. Catherine O'Connor & Thomas Wasow (2004). Animacy encoding in English: why and how. In: Donna Byron & Bonnie Webber (eds.). *The 2004 ACL Workshop on Discourse Annotation,* 118–125.

Typology of Consonantal Insertions

Marzena Żygis
Zentrum für Allgemeine Sprachwissenschaft, Berlin

The present study, based on a typological survey of ca. 70 languages, offers a systematization of consonantal insertions by classifying them into three main types: grammatical, phonetic, and prosodic insertions. The three epenthesis types essentially differ from each other in terms of preferred sounds, domains of application, the role of segmental context, their occurrence cross-linguistically, the extent of variation and phonetic explication.

The present investigation is significantly different from other analyses of consonantal epentheses in the sense that it neither invokes markedness nor diachronic state of the processes under discussion. Instead, it considers the different nature of the epenthetic segments by referring to the representational levels and/or domains which are relevant for their appearance.

1 Introduction

The dynamics of language manifests itself inter alia in the lack of mapping between underlying and surface representations. Underlying segments are sometimes not realized, and vice versa, sounds which are not present underlyingly are articulated and perceived as such. Phonologically, in the former case we are faced with deletion processes, and in the latter with insertions. Different reasons are responsible for both types of processes, including prosodic constraints, stress conditions, segmental neighbourhood, phonotactic requirements, and others.

The present paper focuses on *consonantal* insertions, i.e. processes in which consonants which are not present underlyingly appear on the phonetic surface. The inserted segments have a different status depending on the language: they can be phonemes, allophones or even sounds which do not occur in a given language. The only criterion adopted for the purposes of the present analysis is that they are perceived as segments. Thus, for example, inserted articulatory gestures or acoustic traces which are not perceived by listeners as segments will not be treated as insertions.

Examples of consonantal insertions are provided in (2).

(2) Insertions

wesen+lich	wesen[t]lich	‘considerably’	German
sytuacja	sytu[w]acja	‘situation’	Polish
aapa inni	[?]aapa inni	‘What is this?	Selayarese

Although, at first sight, the examples in (2) illustrate the same kind of process, i.e. epenthesis, each process is in fact different in its nature and motivation as will be analyzed in detail below. The present study proposes a classification of all epentheses into three main categories, i.e. (i) grammatical, (ii) phonetic, and (iii) prosodic. In the remaining part of the article, all insertion types will be discussed in detail and illustrated with examples.

The article is organized as follows. In section 2 a classification of various types of insertion processes based on a typological study is proposed. Section 3 discusses selected accounts of consonantal insertion processes. Section 4 concludes.

2 Typology of insertions

For the purposes of the present study, ca. 70 typologically different languages have been investigated (Austronesian, Romance, Slavic, Germanic, and Semitic among others). The survey leads to a classification of the insertions into three main types:

(3) Insertion types
i) grammatical
ii) phonetic
iii) prosodic

Grammatical insertions comprise all types of insertions which are conditioned morphologically, syntactically and morpho-syntactically, among others. By contrast, phonetic insertions are found on the surface representations and are explicable on articulatory, acoustic, aerodynamic or perceptual bases. Finally, prosodically conditioned insertions include epenthetical processes which refer to prosodic boundaries/domains.

Besides the main insertion types listed in (3), there are also cases attested the classification of which is not straightforward. This is a natural consequence of the interaction of different components, as e.g. syntax and prosody. Such insertions will also be considered in subsequent sections.

In the following, all insertion types will be discussed in detail and illustrated with examples. The discussion comprises the preferred sounds of a given type, their insertion domains, and the role of segmental context as well as phonetics in the explanation of the processes. In addition, it will be analyzed to what extent the processes are subject to inter-speaker variation.

2.1 Grammatical insertions

Several insertions already take place in the grammar of a given language. They are idiosyncratic, characteristic of a particular language. Insertions of this type are subject to phonological, morphological or syntactic regularities and are determined strictly by grammatical categories or constituents. Preferably, in this type of insertions coronal sounds are dominating, albeit labial, velar or glottal sounds are found as well. In a selected number of cases the processes depend on segmental context as well, as will be illustrated below. Finally, the insertions are not subject to inter- or intra-speaker variation and are rather resistant to other factors, such as e.g. speech rate.

For reasons of simplification, the following discussion is limited to morphologically and syntactically conditioned epentheses. It is assumed that morphological insertions take place within words, i.e. mainly between prefixes and stems or stems and suffixes and between constituents of compounds, while syntactic insertions occur between words, i.e. at a syntactic level.

A classic, often cited example is [t]-insertion in Axinica Campa which is morphologically restricted: it takes place only in suffixation processes. Consider the examples in (4).

(4) Axinica Campa

/i-N-koma-i/	[iŋkomati]	'he will paddle'
/i-N-koma-aa-i/	[iŋkomataati]	'he will paddle again'

(Payne 1981:108, Lombardi 2002:239)

Lombardi (2002), when discussing the examples in (4), draws attention to the fact that another potential candidate for insertion, i.e. the glottal stop [ʔ] is not possible here since Axinica Campa does not allow [ʔ] at all, cf. discussion in 3.2.

[t]-insertion is also found e.g. in Odawa, where it takes place at a personal prefix and stem boundary. If, however, the same prefix is attached to different classes of nouns (of inalienable possession), the hiatus is resolved by deletion of a vowel. Consider examples in (5a) and (b).

(5) Odawa

a)	/ki-akat-i/	[kitakat͡ʃi]	'you are shy'

	/ni-ompass/	[nito:mpass]	'you (pl) oversleep'
b)	/ni-o:ss/	[no:ss]	'my bus'
	(Pigott 1980)		

Frequent epenthetical processes include glide insertions which on the one hand take place in a well-defined morphological context and on the other hand might be governed phonologically, see discussion below. For example, in Sinhala the underlying vowel hiatus at root-suffix boundaries is always resolved, whereby the type of repaired strategy depends on the lexical category of the root. For instance, the hiatus in nouns is always resolved by glide insertion. Examples in (6) illustrate this point. (Verbs prefer to resolve hiatus through deletion of an input vowel or, in case of both monosyllabic morphemes, glide epenthesis applies as a last-resort strategy).

(6) Sinhala

/ræ+a/	[ræjə]	'night, sg.def.'
/toppi+a/	[toppijə]	'hat, sg.def.'
/ašu+a/	[ašuwə]	'attic, sg.def.'
/maaligaa+a/	[maaligaawə]	'palace, sg.def.'
(Smith 2001:63)		

Similarly, in Shona [j] is inserted in verb inflections when the second vowel in the sequence is a coronal vowel [e] or [i]. It is also epenthesized in roots when the second vowel is [a] and the first vowel is either [e] or [i], or when the second vowel is a coronal vowel. This is exemplified in (7).

(7) Shona

ta-end-a	[taje^{n}da]	2 subj.pl past-go-fv	'we went'
chi-it-o	[t͡ʃijito]	cl7-do-fv	'an act'
ta-i-p-a	[ta-i-p-a]	1 subj.past-3obj.c19-give-fv	'we gave it'
(Mudzingwa 2007:10)			

In Shona [w]-insertion also takes place: the glide is inserted in roots and verb inflections when the second vowel in the sequence is a labial vowel [o] or [u]; it is also epenthesized in roots when the second vowel is [a] and the first vowel is a labial vowel. This is shown in (8).

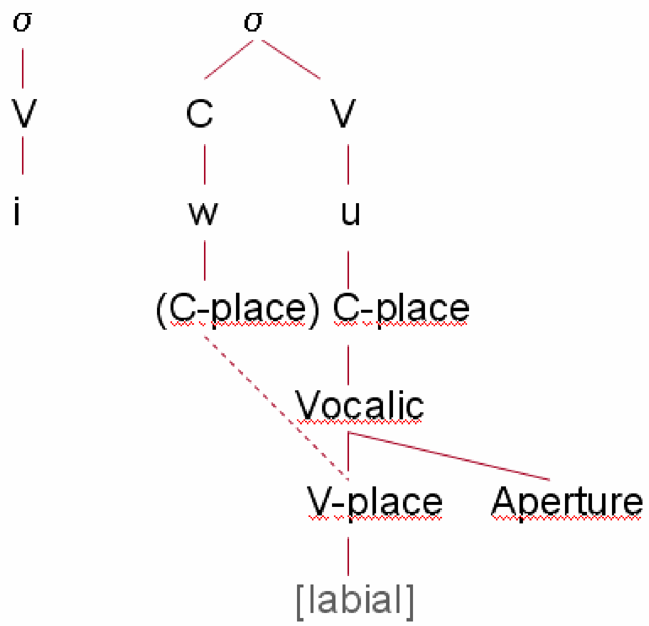
(8) Shona

a-or-a	a[w]ora	3subj.cl6pat-rot-fv	'they rot'
duo	du[w]o	cl5.fish trap	'fish trap'
ama-úta	ama[w]úta	cl1a-deaf person	'deaf person'

(Mudzingwa 2007:10**)**

[w]- and [j]-insertions are found in several other languages (see also examples discussed in section 2.2.1). Phonologically, the insertions in most cases serve to resolve hiatus by inserting a consonantal onset. Hiatus resolution is frequently analyzed as either the so-called default insertion or as feature(s) spreading. In the former case, the whole segment is inserted and it does not share features with neighbouring sounds. Such cases are also not explicable phonetically. In the latter case, a feature or even a bundle of features spreads from the existing vowel, creating a new segment. An example of [w]-insertion is presented in (9), cf. Mudzingwa (2007).

(9) Spreading of features in [w]-insertion



The representation in (9) shows that the inserted sound is not accidental but rather depends on the feature content of the neighbouring sound. It receives vocalic-place features including [labial] via spreading from the following [u]. This conclusion also links phonology with phonetics, where the emergence of glides is due to a percept of formant transition between the adjacent vowels. This is discussed in detail in section 2.2.1.

Grammatical insertions also comprise the so-called linking elements (or linking morphemes) which are found mainly in compounds. Wiese (2000) lists several of such morphemes for Standard German (< –t, –s, –es, –er, –e, –es, –en, –n >) the appearance of which is generally not predictable. Wiese argues that [t] and [s] are in fact not morphemes (e.g. plural markers) but inserted consonants: [s]-insertion applies in the case of certain nouns in the non-head position of compounds, cf. (10a), whereas [t]-insertion is found in morphological formations in which the first compound part ends in [n], cf. (10b). Obviously, such insertions are not explicable in phonological terms as was the case for the Shona examples presented above.

(10)	Standard German	
a)	Schwingung+s+zahl	'frequency'
	Lösung+s+vorschlag	'proposal for solution'
b)	orden+t+lich	'orderly'
	mein+et+wegen	'I don't mind'
	Wiese (2000:145, 146)	

In several other Germanic languages compounds are linked with inserted sounds. For example, in Dutch the nominal compounds may be linked together with [s] or [ə]/[ən] written <s>, and <en> or <e> respectively. The linking segments are originally genitive markers (de Schutter 1994:453); examples are provided in (11).

(11)	Dutch	
	arbeider+s+dochter	'workman's daughter'
	binn+en+pretje	'private joke'
	alleman+s+vriend	'everybody's friend'
	(de Schutter 1994:453)	

Similarly, in Swedish the linking segment /s/ and a vowel are sometimes attached to the first element, as shown in (12).

(12)	Swedish	
	land+s+ting	'county council'
	läs+e+bok	'textbook'
	dag+s+inkomst	'daily income'
	kvinn-o-arbete	'women's work'
	(Andersson 1994:277)	

In Icelandic the so-called genitive compound appears with [s] (Thráinsson 1994:165). The linking element [s] is often required when the first part of a compound is itself a compound, e.g. borðplata 'table top'(stem compound) vs. skrifborð[s]plata 'writing desk top' (genitive compound). As noted by Thráinsson (1994) it is difficult to formulate the rules which underlie the formation of compounds and the appearance of [s] in particular.

In several of the cases presented above it is [s] which connects the compound parts. Although the insertions are not motivated phonetically, i.e. they can not be predicted on the base of the phonetic characteristics of the neighbouring sounds, it is still worth considering that [s] is an extremely salient sound from a perceptual point of view due to its high-frequency noisy characteristics.

There are also other coronals which are inserted, namely rhotics and nasals. The so-called intrusive [r] is found in Bavarian (cf. Bayer & Brander 2008, Bayer & Brander submitted, Ortmann 1998) or in Middle Frankish spoken in Nürnberg in Northern Bavaria (Kabak, & Schiering 2006) where it appears intervocalically, cf. (13).

(13) Middle Frankish

a) [t͡sʊ(-r-)ənɐʃu:l] 'to a school'
zu [r] einer Schule

b) [vɔʊ(-r-)ɪ bin] 'where I am'
wo [r] ich bin
(Kabak & Schiering 2006:69)

The [r]-insertions in Middle Frankish apply in sequences of two function words, i.e. in preposition-determiner, complementizer-pronoun and complementizer-determiner combinations. It should be also noted that the [r]-intrusion is not found within words or compounds. Kabak & Schiering (2006) argue that the function words create a foot which is attached to a phonological word, a fact which is crucial for their analysis. This example shows the interaction of syntactic and prosodic constituents.

In Northern Bavarian or in East Frankonian consonantal insertions may appear with lexical verbs as shown in (14). Note that in the former case an [r] and in the latter case an [n] is epenthesized.

(14) Northern Bavarian

a) [zi:a-r-i] 'I see'
sehe ich

b) East Frankonian

[gi-n-i]   ‘I go’
gehe ich
(Kabak & Schiering 2006:71)

Kabak & Schiering (2006) notice that the common denominator of lexical words in (14) and complementizers in structures as in (13b) is that they occupy the same syntactic position, i.e. the COMP (C^0) slot. A similar conclusion is also drawn by Ortmann (1998) with respect to High Alemannic, where [n] is epenthesized, cf. the examples in (15).

(15) High Alemannic
a) so, wie-n-er gsait hätt   ‘right as he said’
so as-EP-he said has
b) däs Stuck, wo-n-ere id gfalle hätt   ‘the play she didn’t like’
the play REL-EP-her not pleased has
Ortmann (1998:59)

Another well-known example of grammatical insertions is [ɹ]-insertion at certain morphological and syntactic structures found in non-rhotic dialects spoken in several parts of England, Wales and Australia as well as in several non-rhotic dialects of the United States. This phenomenon has been a subject of extensive theoretical and empirical studies (cf. e.g. McCarthy 1991, 1993, Gutch 1992, Gick 1999, Foulkes & Docherty 2000, McMahon 2000, Ito & Mester 2009).

The [ɹ]-insertion appears after the non-high vowels [ə, ɔ:, ɑ:], cf. (16a), as pointed out by Wells (1982). This fully productive process is found in acronyms, loan words and even in the pronunciation of foreign languages, cf. the third example from German in (16b). (Intrusive [ɹ] arose historically as hypocorrection of linking [ɹ] in non-rhotic dialects.)

(16) English ɹ-insertion
a) the idea-[ɹ]-is
law-[ɹ]-and order
b) as far as NAFTA-ɹ-is concerned
schwa-[ɹ]-insertion
Ich bin ja-[ɹ]-auch fertig   ‘I’m also done’
Ito & Mester (2009:167)

The intrusive [ɹ] is also found word-internally, as for instance in *draw*[ɹ]*ing* or *withdraw*[ɹ]*al.*

As pointed out by Kahn (1976) and McCarthy (1993), this fully productive process is restricted insofar as the intrusive [ɹ] does not appear after function words, e.g. 'the apples' is pronounced as [ði æplz̩] and not *[ði ɹ æplz].

The appearance of intrusive [r] is also prosodically conditioned. It has been argued that the insertion of [r] is a manifestation of the requirement for maximal prosodic words to have an onset (cf. e.g. Ito & Mester to appear, cf. also McCarthy 1993).

Even if the process is determined by the word type and prosodic requirements, it has been argued that it is at least partly motivated phonetically, cf. McMahon, Foulkes & Tollfree (1994), Gick (1999), Gick, Kang & Whalen (2002), and Gick (2002). McMahon et al. (1994) propose that if the anterior raising gesture of [ɹ] were removed, the remaining tongue configuration would resemble the articulation of schwa. This hypothesis is further investigated by Gick (1999) who argues that all final schwas in lexical words (in dialects with intrusive [ɹ]) are allophones of /ɹ/. Both studies also support the prediction that an [ɹ]-like pharyngeal constriction should be found in schwa. Furthermore, the analysis of midsagittal MRIs of the vocal tracts of several vowels as well as [ɹ] and [l] by Gick, Kang & Whalen (2002) reveals that at least in some dialects of American English, a single postoral gesture is shared between [l] and [ɔ] and between [ɹ] and schwa. Finally, in Gick's study (2002) on schwas, one subject even showed a bimodal pattern in schwa, which may indicate that this subject has distinct schwas in lexical vs. functional words, a property that has also been observed with respect to [ɹ] in ɹ-vocalizing dialects. In summary, this example illustrates an interaction of grammar, prosody and phonetics.

Another example of a grammatical insertion is [g]-insertion taking place in Mongolian. If the base ends in a long vowel or a diphthong, [g] is inserted before suffixes with a long initial vowel. Consider examples in (17).

(17)	Mongolian		
a)	ablative /-AAs/	dalai 'sea'	dalai[g]aas
		dülii 'deaf'	dülii[g]ees
b)	genitive /-IIn/	xii 'air'	xii[g]iin
		debee 'swampland'	debee[g]iin
	Rialland & Djamouri (1984)		

In summary, grammatical insertions as illustrated by several examples are dominated by coronal sounds (stops, glides, rhotics). They are idiosyncratic, hence their presence is not predictable, albeit at least in selected cases phonetic motivation cannot be excluded. These insertions are not subject to variation.

2.2 Phonetic insertions

Phonetic insertions considerably differ from grammatical ones as the former emerge on the surface representation as a natural consequence of the interaction of phonetic factors. Therefore, in contrast to grammatical insertions, they are explicable solely in phonetic terms. This type of insertion comprises a wide spectrum of possible sounds the occurrence of which depends exclusively on the context, i.e. neighbouring sounds. As shown below, the most frequent insertions are stops produced at different places of articulation as well as glides found in various vocalic contexts. Since the insertions are phonetically grounded, they are expected to occur in typologically different languages, a hypothesis confirmed by the examples provided below. Phonetic insertions are subject to variation, a point which is discussed below.

Phonetic insertions are often gradual processes evolving e.g. from overlapping gestures which at the beginning may not be perceived as categorical sound. Therefore, it is not obvious to predict the final (by)product of a given process and, moreover, to classify it as an insertion. For the purposes of the present study, only phonetic instances in terms of articulatory gestures, acoustic traces, etc. which are perceived as categorical sounds are treated as phonetic insertions. As several studies show, the phonetic insertion can also be incorporated into underlying representation and orthography being a final product of an insertion process. Before the insertion enters the orthographic convention of a given language, it sporadically happens to appear in orthographical representation and is generally treated as a mistake. For example, the Polish word <sytuacja> is permanently written by children as <sytułacja>, with <ł> corresponding to [w] in IPA terms.

In the following, the most frequent phonetic insertions along with their phonetic explanations will be provided.

Among the most common outputs of phonetic insertions are the glides [w] and [j]. The examples in (18a) and (18b) illustrate glide insertions in Japanese and Polish.

(18) [w] and [j]-insertions

a) Japanese

guai	gu[w]ai	'condition'
siawase	si[j]awase	'happiness'

Kawahara (2002)

b) Polish

sytuacja	sytu[w]acja	'situation'
trio	tri[j]o	'trio'
Rubach (1984)		

The context of [w] insertions in (18) always implies the presence of a following or preceding /u/, whereas the glide [j] is inserted in the context of the preceding /i/. Thus, the outputs are not accidental but depend on the neighbouring sounds as has already been mentioned in section 2.1. In contrast to grammatical insertions of this type, the glide insertion is not determined by morphological or syntactic constituents/rules.

Phonetically, it is probably the (lengthened) formant transition of neighbouring sounds which gives the perceptual impression of a new emerging sound (in line with Ohala's 1981 interpretations of sound change). Although such observations have been made intuitively (Blevins 2007), the processes have not been investigated in detail. It is not clear what exact requirements should be met for a formant transition to be perceived as a glide.

In order to gain more insight into the topic, Zygis (2009) conducted a perceptual study which involved the manipulation of transition length in the item [ia]. The transition between [i] and [a] was lengthened from 50 to 230 ms in 30 ms steps so that a 7–step continuum was created. Both the total duration of the item as well as the duration of [a] was constant. The files were played twice from a laptop via headphones to ten native speakers of German, three native speakers of Polish and three of English (in all languages, /j/ is present in the phonemic inventory). The informants were asked to write down what they hear. In all answers [j] was heard, albeit at various transition lengths, cf. Table 1. The preliminary results imply that there is no categorical boundary common for the listeners since (i) different results were obtained for the same stimulus independent of the native language of the informant and (ii) variation was found within informant groups with the same native language background, cf. results obtained for five German native speakers presented in Table 2.

Table 1: Responses of a Catalan, German and Polish speaker to different stimuli files.

Stimulus	English	German	Polish
S_1 50 ms	ja	ija	ija
S_2 80 ms	ja	ija	ija
S_3 110 ms	ija	ija	ija
S_4 140 ms	ija	ija	ija
S_5 170 ms	ija	ija	ija
S_6 200 ms	ija	ijar	ija
S_7 230 ms	ija	i:ja	ja

Table 2: Responses of five German speakers to different stimuli files.

Stimulus	German 1	German 2	German 3	German 4	German 5
S_1 50 ms	ija	ija	ija	ija	i:a
S_2 80 ms	ija	ija	ija	ija	i:a
S_3 110 ms	ija	ija	(i)ja	ija	i:a
S_4 140 ms	ija	ija	ja	ija	i:a
S_5 170 ms	ija	ja	ja	ija	i:a
S_6 200 ms	ija	ja	ia	ija	i:a
S_7 230 ms	ja	ja	ea	ja	ja

Thus, the preliminary results strongly suggest that the obtained differences might be explicable not only in terms of the language-specific phonological background, but depend on the individual perceptual (and probably production) characteristics as well.

Other frequent insertions comprise stops articulated at different places of articulation. Well-known examples refer to coronal stop [t], the so-called intrusive stop which appears in sonorant-fricative clusters, as shown in (19).

(19) American English

[tɛnts] [tɛnsiti] ‘tense, tensity’

[tɛndz] [tɛn] ‘tens, ten’

[fɑlts] [fɑlsiti] ‘false, falsity’

(Fourakis & Port 1986:1999)

Different phonological and phonetic explanations of the processes in (19) have been proposed; cf. e.g. Zwicky (1972), Donegan & Stampe (1979) Dinnsen (1980), Wetzels (1985), Clements (1987).

Zwicky (1972) and Dinnsen (1980) argue for a phonological rule describing the epenthesis Ø → [t]/ S_F (S=sonorant, F=fricative). Ohala (1974) and Donegan & Stampe (1979) state that the emergence of [t] and [d] in processes shown in (19) does not follow from the rule application but from mistiming effects. Clements (1987) accounts for the process in terms of feature spreading. Ohala (1974) claims that closing the velum before the release of the nasal closure produces a configuration of articulators similar to that of a homorganic stop.

Fourakis & Port (1986) examined the production of sonorant-fricative and sonorant-stop-fricative clusters by two groups of speakers: South African dialect speakers and American mid-western dialect speakers. Results of an acoustic analysis show a clear difference between the informants: while the South African speakers maintained a contrast in the sequences sonorant-fricative and sonorant-stop-fricative, American speakers neutralized the contrast in the sense that they always inserted stops after the sonorant if the fricative was voiceless. If the fricative was voiced, the speakers either omitted the stops in clusters like /ldz/ or /ndz/ or sometimes inserted a stop in clusters such as /nz/ and /lz/. Fourakis & Port (1986) also show that the inserted stop in American was significantly shorter than a corresponding underlying stop, and that it also affected the length of the preceding nasal. On the basis of their results, Fourakis & Port (1986) argue that neither of the explanations proposed in the literature (see Zwicky 1972 and Dinnsen 1980 vs. Ohala 1974 and Donegan & Stampe 1979) is tenable. The insertion rule does not consider the fact that the inserted sounds are different from the underlying. The phonetic explanation as proposed by Ohala (1974) is insofar problematic – according to Fourakis & Port – as it is supposed to be universally applicable, but as the South African data show it does not appear universally. Instead, the authors propose the application of language-specific rules, the so-called *phase* rules which are probably confined within one or two syllabic cycles and are partially controlled by phonological, contextual features. The rules are learned so that they may vary in the details of the articulatory output from speaker to speaker. Finally, the rules may change very rapidly as they are sensitive to pragmatic communicative needs, word frequency, usage frequency, etc.

Although the proposal made by Fourakis & Port gained much attention in the literature, it is still undeniable that a phonetic explanation given by Ohala (1974) can be still maintained as Ohala does not state that the mistiming of articulators takes place universally; otherwise the epenthesis would take place in every language. The mistiming probably appears due to the inability to finally control the articulators in every statement pronounced by every speaker. It may occur that some realizations are less accurate articulatorily, leading to an articulatory configuration which is perceived in a different way than the planned underlying representation. In fact, such 'misperceptions' trigger a sound change.

It is also worth noting that coronal stop insertions which are similar to the ones presented above are found cross-linguistically. Examples in (20) illustrate [d]-insertion in Spanish.

(20) Spanish

ven(i)ra > ven[d]rá	'he will come'
sal(i)ra > sal[d]rá	'he will leave'

(Wetzels 1985:287)

In (20) the coronal stop [d] is inserted in two contexts: a) between the nasal /n/ and the rhotic /r/, and b) between the liquid /l/ and the rhotic /r/. The same insertion contexts are found in Old French. Consider the examples in (21) where besides /n/_/r/ and /l/_/r/, the context /s/_/r/ also triggers the emergence of not only [d] but also [t]. All examples imply deletions of vowels.

(21) Old French

	menour > man[d]re	‘smaller’
	failir > fal[d]ra	‘fail, lack’
	valier > val[d]ra	‘be worth’
	cousons > cos[d]re	‘sew’
	(Pigott & Singh 1985:419)	

Vincent (1988) reports that in Italian dialects, [t] can be heard in words in which it is not present underlyingly. His examples refer to contexts such as /l/_/s/ and /n/_/s/. The appearing [t] optionally undergoes merging with the following /s/, which eventually leads to the emergence of an affricate. Thus, e.g. *falso* ‘false’ or *senso* ‘sense’ can be pronounced either with an affricate as [falt͡so] and [sent͡so], or with a fricative preceded by a ‘less perceptible’ plosive, i.e. [fal^{t}so] or [sen^{t}so] (Vincent 1988:291). For a restrictive number of southern Italian speakers the process has spread to labials, i.e. the affected words could be optionally pronounced as affricates or stop+fricative sequences, e.g. tonfo ‘thud’ is pronounced as [toɱ^{p}fo] or [tomp͡fo] and inverno ‘winter’ as [iɱ^{b}vɛrno] or [imb͡vɛrno] Vincent (1988:291).

Next to coronal stops, several insertions found cross-linguistically include labial stops. Examples in (22) show labial insertions from Old and Modern English. (Phonetic explanations of these processes are given below).

(22)	Old English	>	Middle English	
	sceamol	>	scham[b]el	‘stool’
	nemnan	>	nem[p]ne	‘to name’
	æmtig	>	em[p]ti	‘empty’
	(Pigott & Singh 1985:418)			

Similar processes are found in Modern English and German, where [p] can emerge between a labial nasal [m] and a coronal sound. Examples in (23) illustrate this variable process.

(23) Modern English

	warmth	warm[p]th
	Thomson	Thom[p]son
	(Wetzels 1985:288)	

Standard German		
Wams	Wam[p]s	'doublet'
Ramsch	Ram[p]sch	'junk'
Wiese (2000:233)		

The emergence of bilabial stops in sequences [mbn] was also found in Old Spanish, as illustrated by the examples in (24).

(24) Old Spanish

hominem > hom[b]re	'man'
nomine > nom[b]re	'woman'
(Millardet 1923:293ff)	

In Ancient Greek, the labial [b] appeared in a well-defined context, namely between the nasal /m/ and the following sonorant /r, l/. This is shown in (25).

(25) Ancient Greek

gam-ros > gam[b]ros	'married'
a-mrotos > am[b]rotos	'immortal'
me-mlōka > mem[b]lōka	'he walked'
(Wetzels 1985:287)	

Ohala (1995:161) also lists several examples from Latin, Landais dialects of French, Old Swedish, and Old Provençal in which the labial stops are inserted between /m_n/. The epenthesis emerges in a phonetically natural way, as claimed by Ohala. He shows that it is a temporal overlap of the /m/ and /n/ closures which is responsible for [p]- and [b]-epenthesis. The simultaneous labial closure of /m/ and apical closure of /n/ create a pocket of air between them. At the time when the labial closure is released, the compressed air undergoes a momentary rarefaction of pressure and is released with a click-like burst. Listeners, as argued by Ohala, are likely to interpret the stop burst auditorily as [p], which creates the basis for their own pronunciation. It is also probable that listeners will interpret the emerging sound as a voiced [b].

Phonologically, examples presented in (19), (21), (24), (25) can be explained by applying the syllable contact law according to which 'a syllable contact A$B is the more preferred, the less the Consonantal Strength of the offset A and the greater the Consonantal Strength of the onset B' (Vennemann 1988). In other words, contacts like e.g. [n.d], [l.d], [m.p], [m.b] are favoured over [n.r], [l.r], [m.n], [m.r] since the differences in Consonantal Strength (sonority) are higher in the former than in the latter sequences.

In summary, a crucial difference between grammatical insertions and phonetic insertions is that the former are determined by a specific morphological/ syntactic context, whereas the latter are the result of articulatory timing relationships of neighbouring sounds which motivates their potential appearance cross-linguistically.

2.3 Prosodic insertions

Another type of insertions is primarily prosodically conditioned, i.e. their appearance is determined by prosodic constituents such as syllable, prosodic foot, prosodic word, phonological phrase, intonational phrase, and phonological utterance, cf. prosodic hierarchies postulated by Selkirk (1980 a, b), Booij (1983), Nespor & Vogel (1988). The insertions are found either at the boundaries of the prosodic constituents or they require a given prosodic constituent as a domain of their application. Since prosodic constituents are often domains of stress assignment, prosodic insertions can also interact with stress.

The most frequent prosodic insertions are glottal stops and glottalizations followed by coronal sounds (see below). They often serve as boundary markers/signals and are subject to inter- and intra-speaker variation. Their occurrence depends among other factors on speech rate, speaker's gender, dialect, register, phrasal position, stress conditions, and others.

It seems that the variation found in prosodic insertions is greater at higher than at lower prosodic boundaries. This could be caused by the fact that lower prosodic constituents are created in the lexicon, cf. e.g. the discussion in Zec (2005); they do not incur phrasal stress or intonation and therefore behave differently from higher ones. This point undoubtedly requires further investigation.

Glottal stops and glottalization of prosodic insertions are perceptually distinct, albeit only to some extent; they are 'merely' boundary markers facilitating prosodic parsing and do not contribute to the content of a given word/phrase. Therefore, sounds such as for example sibilants which are perceptually extremely salient are not inserted as boundary markers. In addition, sounds which are phonemes in a given language are less optimal candidates for appearing at prosodic boundaries as they could potentially lead to confusion and misinterpretation of existing words.

An important question concerning prosodic insertions is whether their appearance is also related to neighbouring sounds, e.g. sounds appearing in the initial position of a prosodic phrase. Several studies have shown that glottal stops/glottalization are found if the prosodic constituent starts with a vowel (cf. examples discussed below). From a phonological point of view such insertions serve to create the constituent onset, mostly the syllable onset needed for the creation of an optimal CV syllable. The fact that vowels and glottal stops do not have supraglottal constrictions and share glottal constrictions (see below) favours their co-occurrence. Although it is still not entirely clear whether certain vowels facilitate the appearance of prosodic insertions more than others, a few studies have suggested that glottal stops and glottalization are favoured in the context of following low vowels, which is probably due to their similar larynx configuration (cf. studies discussed below).

The most striking characteristics of glottal stops and glottalization when they are inserted seem to be their huge variability found not only among speakers of a given language but also in the pronunciations of individual speakers. This inter- and intra-speaker variability has been observed in several languages and is argued to be dependent on several parameters such as phrasal position, stressed vs. unstressed syllable, speech rate, segmental context, dialect, speaker's gender, and others (see e.g. *American English*: Umeda 1978, Pierrehumbert 1994, Pierrehumbert & Talkin 1991, Redi & Shattuck-Hufnagel 2001; *Chitwan Tharu*: Leal 1972, *Danish*: Haberland 1994; *Garo*: Burling 1992, *Nootka*: Shank & Wilson 2000; *Tümpisa (Panamint) Shoshone*: Dayley 1989).

For instance, Umeda (1978) analyzed the occurrence of glottal stops in American English in dependence on several factors such as speaker's reading style, difficulty of the material, the segmental context, stress conditions, type of words (functional vs. lexical), and frequency of occurrence of words. Umeda's results show that speaker's reading style and difficulty of material influence the insertions of glottal stops more strongly than phonological and grammatical factors. Furthermore, slow speech rate and grammatical breaks (e.g. after adverbs as 'however' or 'instead') also induce a higher percentage of glottal stops than high speech rate and fluent speech without breaks. Finally, the study shows that rare words are more frequently marked with glottal stops than common words.

Redi & Shattuck-Hufnagel (2001), who investigated glottalization at phrase boundaries which are medial or final in an utterance in American English, also stress the great range in the rate of glottalization in individual speakers' pronunciations. The study reports that this rate is higher for words at the ends of utterances than for words at the ends of utterance-medial intonational phrases, and it is higher at the boundaries of full intonational phrases than at those of intermediate intonational phrases.

Kohler's (1994) study, which focuses on glottal stops and glottalization in word-initial and word-medial position in German read speech of a North German variety, also highlights the variation issue. Kohler reports on the occurrence of glottal stops and glottalization as boundary markers on the one hand and as reduction phenomena of supraglottal stop articulations on the other hand. The results indicate that after pauses/silence, the presence of a glottal stop is more frequent than its absence. Stressed vowels also substantially favour the presence of glottal stops/glottalization than unstressed ones after a segmental context. At word-intial boundaries the highest proportion of glottal stops is found before stressed vowels and after stops (72%), (other segments include vowels, sonorants, and fricatives). The glottal closure may also completely replace the supraglottal constriction, e.g. Freita[k a]bend is pronounced as Freita[ʔ a]bend (Kohler 1994:45).

Studies on prosodic insertions cited below show that glottal stops/glottalizations are found at boundaries of several prosodic units such as: syllable, foot, phonological word, clitic group, phonological phrase, intonational phrase and phonological utterance (cf. prosodic hierarchy proposed by Nespor &Vogel 1988 [2007:16]).

In the Bisu language, in onsetless syllables, the vowel which constitutes the nucleus usually carries a preceding glottal stop.

(26) Bisu

[ʔa ^{31}]	'don't'
[ʔe ^{55}]	'go'
[ʔup ^{31}]	'say'
(Shixuan 2001:22)	

In German, glottal stops/glottalizations occur optionally at the beginning of a vowel-initial foot, i.e. as onsets of stressed syllables, cf. examples in (27) (Hall 1992, also Wiese 2000).

(27) German

arm	[ʔaʀm] or [aʀm]	'poor'
oft	[ʔɔft] or [ɔft]	'often'
Theater	[te.ʔá:.tɐ] or [te.á:.tɐ]	'theater'
(Hall 1992:58)		

In Selayarese (an Austronesian language), a glottal stop is inserted initially in the 'intonation unit' cf. examples in (28) (Mithun & Basri 1986, cited in Lombardi 2002:266).

(28) Selayarese

ʔinnĩ

ʔaapa innĩ

(Mithun & Basri 1986)

In Anejom, a Vanuatu language, all vowels are preceded by a non-phonemic glottal stop in utterance-initial position, cf. examples in (29) ('denotes an accent).

(29) Anejom

/aek/	['ʔaek]	'you sg.'
/et apam aen/	[ʔɛd̥ abam' aen]	'he came'

(Lynch 2000:17)

Other frequent insertions found at edges of prosodic constituents are glottal fricatives. It seems that these are found especially in final position rather than in initial positions. Blust (2009) found the insertion of [h] at the prosodic-word final position in several Austronesian languages like e.g. Aklanon (and other Bisayan dialects of the C. Philippines), Tagabili, Taosung, many northern and central Sarawak languages, including: Miri, Narum, Kiput, Berawan, Western Penan, Long Wat Kenya, Sebop, Kelabit, Dalat, Matu, Serike Malanau. Blust also provides several examples of glottal stop insertion at the prosodic-word final position.

Note that the insertion of a glottal fricative is in opposition to a requirement for syllables not to have codas (expressed by the constraint NoCoda in Optimality Theory).

The presence of [h] in final positions, especially in larger prosodic units such as phonological phrases or utterances might be attributed to aerodynamics. In particular, towards the end of a given constituent (phrase, utterance) speakers begin spreading the vocal folds in anticipation of non-speech breathing (Lisker et al. 1969). This effect is accompanied by the decline in subglottal pressure over the course of an utterance (e.g. Westbury & Keating 1986). The outcoming air flow might be eventually perceived as a glottal fricative by listeners. This hypothesis requires, however, a detailed aerodynamic study.

Besides glottal segments, coronal segments are inserted at prosodic boundaries. The following examples illustrate the insertion of [r] and [n].

(30) Anejom

inpeke Anejom > inpeke [r] Anejom	island Aneityum	'the island of Aneityum'

(Lynch 2000: 29)

In Anejom̃, trilled [r] is inserted between words in the same phrase when a vowel-final word is followed by a vowel-initial one, cf. (30). As stressed by Lynch (2000) the latter process appears to be a variable or infrequent rule. (The trilled [r] is also inserted in compounds when a morpheme ending in a back vowel is followed by a vowel-initial morpheme, e.g. awo-upni > awo[r]upni 'do well'.)

In Tunica, phrasal-final words end in a consonant. There are two strategies to meet this requirement: deletion of a word-final vowel, which applies in a limited number of words, and epenthesis of a phrase-final consonant, a strategy followed in most cases. The consonant which is epenthesized is the nasal [n]. Examples are provided in (31).

(31) Tunica

hatika	hatika[n]	'again'
sahku	sahku[n]	'one'

Haas (1940) cited after Lombardi (2002:233)

In summary, glottal stops and fricatives are preferable sounds cross-linguistically as far as prosodic insertions are concerned. Prosodic insertions take reference to prosodic boundaries/domains. It remains to be seen to what extent they are explicable phonetically. This type of insertion is variable, especially when higher prosodic units are concerned, depending on factors such as speech rate, phrasal position, degree of prosodic prominence, word type, segmental context and others.

3 Previous approaches

Although it seems that consonantal insertions in comparison to vocalic insertions have been investigated considerably less frequently, there are several approaches dealing with consonantal epentheses. These, however, are mostly limited to selected languages, cf. e.g. Ortmann (1998), Alber (2001), Kawahara (2002), Ito & Mester (2009).

There are also a few studies such as Rubach (2000), Lombardi (2002), Uffmann (2007), de Lacy (2006) and Blevins (2007) which analyze consonantal epentheses from a broader, cross-linguistic perspective and offer a unified account of them. Whereas the first four studies approach the epentheses from the markedness point of view, the latter one takes a different route: it analyzes naturalness and unnaturalness of the processes by focussing on phonetic properties of the inserted sounds and their contexts as well as the historical background. In the following sections, selected approaches will be briefly sketched and commented.

3.1 Lombardi (2002)

Lombardi's approach (2002) treats insertions from the point of view of markedness. She argues, running counter to Prince & Smolensky (1993) assumptions, that the Place markedness hierarchy should be revised to include Pharyngeal as the least marked place, cf. hierarchies provided in (32a) and (b).

Pharyngeals – in Lombardi's account comprising only [ʔ h] – occur indeed as epenthetic consonants, but as Lombardi shows, only in specific situations. The best candidates for epenthesis are glottal consonants, which are frequently epenthetic. This is because they are the least marked in the Place markedness hierarchy and more specifically, they lack a place node.

(32) Place markedness hierarchy

a. *Dors, *Lab >> *Cor (Prince & Smolensky 1993)

b. *Dors, *Lab>> *Cor>> *Phar (Lombardi 2002)

Lombardi (2002) offers an explanation of insertions by proposing the place markedness scale in (32b) where *Pharyngeals ([ʔ h]) are the least marked segments and therefore the most optimal candidates for epenthesis.

Several examples provided by Lombardi are aimed at showing that [ʔ h], being the best candidate for insertions, are sometimes blocked by other restrictions (constraints). In such a case, the next candidate, i.e. a coronal, is epenthesized. The analyses show that whereas a glottal stop is a phonologically driven epenthetic consonant, e.g. inserted in order to satisfy the Onset constraint, coronals only occur in specific, mostly morphologically restricted cases.

A major objection to Lombardi's proposal is that insertions are not treated with respect to context, which implies that every context potentially allows (glottal) insertions. As several examples in 2.1. show, this claim is not evidenced by cross-linguistic data as they are motivated not only by the context, but also by morphological restrictions. Following the line of Lombardi's reasoning it is difficult to explain why e.g. certain languages epenthesize [s] instead of [t]. It seems that the hierarchy in (32b) is indifferent as far as the choice of a particular consonant within a given class is concerned, which leaves a broad spectrum of choices even within a natural class. However, as far as e.g. coronals are concerned, the choice is not random, but as the examples show, [t]-epenthesis takes place relatively often while others (see [s̨]) seem not to occur at all. This conclusion leads us to a crucial point, namely, that the analysis by Lombardi does not differentiate between different types of insertions. German [ʔ]-insertion, an example provided by Lombardi, is clearly an optional, prosodically motivated process, cf. (27). In Lombardi's analysis it serves to support the unmarkedness of glottal stops and therefore the scale in (32b), whereas the same scale is meant to account for morphologically conditioned processes. In other words, it seems that different levels of representation (including phonological, morphological and prosodic/phonetic) are treated as one homogenous representation.

3.2 Uffmann (2007)

Although the study by Uffmann (2007) focuses on [ɹ]-epenthesis in English, it also offers a unified account of consonantal insertions. Similarly to Lombardi's (2002) approach, Uffmann's proposal is based on markedness scale(s), but in contrast to Lombardi, the account takes into consideration both segments and the position in which they are frequently found. His approach refers to scales which align prominent positions with prominent segments, cf. (33), according to which a vowel is highly preferable in a peak syllable position (least marked) and least preferable in a syllable margin (highly marked). A laryngeal segment is least marked in a syllable margin position and highest marked in a syllable peak position, cf. (33a,b).

(33) Markedness scales: segments in margin and peak position

(a) *Margin/V >> *Margin/r >> *Margin/l >> *Margin/nas >> *Margin/obs >> *Margin/lar
(b) *Peak/lar >> *Peak/obs >> *Peak/nasal >> *Peak/l >> *Peak/r >> *Peak/V
Prince & Smolensky (1993)

Uffmann makes use of the scale in (33a) when he proposes an account of glottal stop epenthesis which frequently appears in a margin position. In Uffmann's example, the glottal stop is inserted in the word-initial position of a one-syllabic word which is also an onset position. If a word with more syllables were evaluated, an additional constraint would have to be proposed in order to place the glottal stop at the word-initial position. Clearly, if there is a hiatus, most languages resolve it by inserting a glide, and not a glottal stop. In order to ensure an epenthesis of a glide intervocalically, Uffmann (2007) proposes another markedness scale with special reference to the intervocalic context. The scale is shown in (34).

(34) Markedness scale: segments in the intervocalic context
*V_V/lar >> *V_V/obs >> *V_V/nas >> *V_V/l >> *V_V/r >> *V_V/V
Uffmann (2007)

The scale in (34) treats a laryngeal segment as highly marked in the intervocalic position and a vowel (or a glide) as the most preferred segment, i.e. the least unmarked. According to this proposal, glottal stops are inserted to maximize the contrast with the following vowel, and thus the perceptual salience of the epenthetic segment, and glides are inserted to minimize the contrast to the following or preceding vowel, and thus the perceptual salience of the epenthetic element (Uffmann 2007:458). Furthermore, the contrastiveness statements are – according to Uffmann (2007) – enhanced by the degree of phonetic realisation which may vary, i.e. glide insertion is optional in many languages, and the degree of gliding may vary, whereas 'glottal stop epenthesis hardly ever shows this amount of variability' (Uffmann 2007:458). However, several experimental studies on glottal stops and glottalization cited in section 2.3.1. point to the opposite conclusion, namely, glottal stops show an almost chameleon-like behaviour in terms of inter- and intra-speaker variation, cf. also Pompino-Marschall & Zygis (2010).

It also remains unclear why the inserted glide should be similar to a vowel if one considers the fact that the contrast between the onset and the peak should be maximized in an optimal CV syllable. Moreover, some languages show free variation between glides and glottal stops found in intervocalic position, and it would be difficult to account for this with markedness constraints based on the scale in (34). Selected examples are given in (35).

(35) German

Hi.[ʔ]á.tus	or	Hi.[j]á.tus	'hiatus'
Lin.gu.[ʔ]ís.tik	or	Lin.g[ʋ]ís.tik	'linguistics'

(Alber 2001)

Finally, in light of the diversity of cross-linguistic data partly presented in the previous sections and the scale in (34), other questions arise, as for instance why some languages prefer the epenthetic [n], as e.g. Alemannic, instead of [r], as is the case in Bavarian or Ilocano. More importantly, it is not clear why the languages do not employ the least unmarked epenthetic candidates, which would be in line with OT assumptions.

3.3 Blevins (2007)

A drastically different view on insertions from the ones presented above is taken by Blevins (2007) who argues against markedness as the driving force of epenthesis. Blevin's arguments refer to (i) the position of the inserted segments as governed by constraint ONSET demanding syllables to have onsets as well as (ii) the quality of the inserted segment interpreted in terms of segmental markedness constraints. By providing several examples Blevins claims that accounts of epenthesis in terms of fulfilling the constraint ONSET encounter severe problems if cross-linguistic data are taken into consideration. For example, if an epenthetic consonant occurs at the beginning of a prosodic domain but not intervocalically within the same domain, then the ONSET-filling approach is difficult to maintain. Arguing against the segmental markedness constraints, Blevins claims that markedness constraints are not able to account for the fact that in cases where the epenthetical sounds are not explicable in phonetic terms and the insertion processes can be reconstructed, epenthetic sounds are those for which earlier consonant loss is evidenced. There are two other arguments brought forward by Blevins: it is not possible to account for highly marked epenthetic consonants by the segmental markedness constraints, and finally, it appears that in some languages the epenthetic consonant is not a contrastive segment and is therefore unlikely to be a direct consequence of the interplay of segmental markedness constraints (but see the discussion in 3.2 on this point).

Blevins (2007) analyzes insertions against their historical background. She differentiates between natural and unnatural histories for patterns of consonant insertions by making no reference to syllable onset or segmental markedness. In the domain of natural histories, glide and laryngeal epenthesis are found, while consonants which were/are subject to coda weakening and evolved into epenthesis processes are counted as unnatural histories. Finally, a mix of natural and unnatural history in some epenthetical processes is also found.

Although Blevins's approach differs from the present one which refers to representational levels and dispenses with diachronic data as an explanation source, there are some important points which both approaches share. First, the concept of markedness is definitely rejected as an explaining tool for consonantal epentheses as it does not differentiate between different types of processes. Second, both approaches take into consideration phonetic facts which in consequence lead to a clear differentiation between various epenthesis types.

4 Summary

The present study offers a systematization of consonantal insertions by classifying them into three main types; (i) grammatical, (ii) phonetic, and (iii) prosodic insertions. In addition, it has been shown that in some cases the epenthesized segment depends not only on morphology/syntax but is also conditioned prosodically and/or phonetically.

The epenthesis types significantly differ from each other with respect to some parameters, as e.g. preferred sounds, domains of application, the role of segmental context, their appearance (or absence) cross-linguistically, or the extent of variation and phonetic explication.

While grammatical insertions take reference to morphological or syntactic constituents/rules, prosodic insertions are bound to prosodic boundaries/constituents. By contrast, phonetic insertions depend on their neighbouring sounds. The insertions significantly differ with respect to preferred sounds. Whereas grammatical insertions seem to prefer coronal sounds, prosodic insertions are mostly restricted to glottal stops and fricatives, and phonetic insertions to glides and stops (coronal, labial, and velar). As far as the role of the segmental context is concerned, it conditions phonetic insertions, but less so prosodic ones, and seems to have no influence on grammatical insertions. This property is related to the fact that grammatical insertions are idiosyncratic, characteristic for one particular language, whereas both phonetic and prosodic insertions of the same type are found cross-linguistically. In the same vein, grammatical insertions are (mostly) not explicable phonetically, while phonetic insertions are output of the interaction of phonetic principles. It remains to be seen to what extent prosodic insertions can be explained by phonetic notions, especially in terms of glottal and supraglottal coordination. Finally, a clear difference between grammatical epentheses on the one hand and phonetic and prosodic epentheses on the other hand regards their phonetic realization in terms of inter- or intra-speaker variation: whereas the former are not subject to variation, the latter vary to a great extent.

Table 3 provides an overview of the most important characteristics of different types of consonantal insertions.

Table 3: An overview of consonantal insertions.

Type of insertion	**Preferred sounds**	**Domains**	**The role of segmental context**	**Appearance**	**Phonetically explicable**	**Variation**
Gramma-tical	coronals (stops, glides, rhotics)	morphological/ syntactic	context-independent	idiosyncratic	no	no
Phonetic	glides: j, w, stops: coronal, labial & velar	no reference to domains	context-dependent	cross-linguistically	yes	yes
Prosodic	glottal stops and fricatives	prosodic domains	context-independent (to some extent)	cross-linguistically	partly (?)	great

The present study considerably differs from other analyses of consonantal epentheses (a selection of which is presented in section 3) in the sense that it neither invokes markedness, as this appears problematic in light of cross-linguistic data, nor the diachronic state of processes. Instead, it considers the different nature of the epenthetic segments by referring to the representational levels or domains which are relevant for their appearance. The fact that epentheses are different in their nature can be additionally underpinned by phonetic evidence partly available in the literature, i.e. phonetic insertions show different phonetic characteristics from their underlying counterparts, e.g. 'intrusive' [t] is shorter than the underlying [t] (Fourakis & Port 1986). The hypothesis that phonetic insertions also differ from grammatical ones in terms of phonetic parameters I leave open for further studies.

Acknowledgments

I would like to thank T.A. Hall, Stefanie Jannedy, Jaye Padgett, Melanie Weirich and audiences at the Workshop on Form and Function of Insertions and Deletions in Speech (Osnabrück' 09) for valuable comments on this paper. The usual disclaimers apply. This research has been supported by Federal Ministry of Education and Research (01UG0711).

References

Alber, Birgit (2001). Regional variation and edges: Glottal stop epenthesis and dissimilation in standard and southern varieties of German. *Zeitschrift für Sprachwissenschaft* 20, 3-41.

Andersson, Erik (1994). Swedish. In: König, Ekkehard & *Johan* van der Auwera (eds.). *The Germanic Languages*. London: Routledge, 271-312.

Blevins, Juliette. (2007). Consonant epenthesis: natural and unnatural histories. In: Good, Jeff (ed.) *Language universals and language change*. Oxford: OUP, 79-107.

Bayer, Josef & Ellen Brander (2008). On Wh-Head-Movement and the Doubly-Filles-Comp Filter. *Proceedings of the 26th West Coast Conference on Formal Linguistics* 87-95.

Bayer, Josef & Ellen Brander (submitted). Wie oberflächlich ist die syntaktische Variation zwischen Dialekten? Doubly-filled COMP revisited. Ms. Universität Konstanz.

Booij, Geert (1983). Principles and parameters in phonology. *Linguistics* 21, 249-280.

Burling, Robbing (1992). Garo is a minimal tone language. *Linguistics of the Tibeto-Burman Area*, 15.5: 33-51.

Blust, Robert (2009). *The Austronesian Languages*. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University

Clements, G. Nick (1987). *Phonological feature representation and the description of intrusive stops.* Parassesion on Autosegmental and Metrical Phonology. Chicago: Chicago Linguistic Society, 29-50.

Dayley, Jon P. (1989). *Tümpisa (Panamint) Shoshone Grammar.* University of California. Publications in Linguistics. 115.

Dinnsen, Daniel A. (1980). Phonological rules and phonetic explanation. *Journal of Linguistics* 16, 171-191.

Donegan, Patricia & David Stampe (1979). The study of natural phonology. In: Dennisen, Daniel (ed.). *Current Approaches to Phonological Theory*. Bloomington: Indiana University Press, 126-173.

Foulkes, Paul & Gerard J. Docherty (2000). Another chapter in the story of /r/: labiodental variants in British English. *Journal of Sociolinguistic* 4, 30-59.

Fourakis, Marios & Robert Port (1986). Stop epenthesis in English. *Journal of Phonetics* 14, 197-221.

Gick, Bryan (1999). A gesture-based account of intrusive consonants in English. *Phonology* 16, 29-54.

Gick, Bryan (2002). An X-ray investigation of pharyngeal constriction in American English schwa. *Phonetica* 59, 38-48.

Gick, Bryan, Kang, Min A. & Douglas H. Whalen (2002). MRI evidence for commonality in the post-oral articulations of English vowels and liquids. *Journal of Phonetics* 30, 357-371.

Gutch, Donald (1992). Linking and intrusive r in English and Bavarian. In: Blank, Claudia (ed.) *Language and Civilization: A Concerted Profusion of Esseys and Studies in Honor of Otto Hietsch*. Frankfurt/Main: Lang, 555-611.

Haberland, Hartmut (1994) Danish. In: König, Ekkehard & *Johan* van der Auwera (eds.). *The Germanic Languages*. London: Routledge, 313-348.

Hall, Tracy A. (1992). *Syllable Structure and Syllable-Related Processes in German*. Tübingen: Max Niemeyer Verlag.

Haas, Mary R. (1940). Tunica. In: *Handbook of American Indian languages*. Vol.4. New York: J.J. Augustin, 1-143.

Ito, Junko & Armin Mester (2009). The extended prosodic word. In: Grijzenhout, Janet & Bariş Kabak (eds.) *Phonological Domains. Universals and Deviations*. Berlin: Mouton de Gruyter. 135-194.

Ito, Junko & Armin Mester (to appear). The onset of the prosodic word. In: Parker, Steve (ed.) *Phonological Argumentation: Essays on Evidence and Motivation.* London: Equinox.

Kabak, Baris & René Schiering (2006). The phonology and morphology of function word contractions in German. *Journal of Comparative Germanic Linguistics* 9, 53-99.

Kahn, Daniel (1976). *Syllable-Based Generalizations in English Phonology*. Ph.D. diss., MIT. Published [1980]. Garland Press: New York.

Kawahara, Shigeto (2002). *Feature recycling: Hiatus resolution by way of glide formation in Japanese*. Talk presented at International Christian University, Tokyo.

Kohler, Klaus J. (1994). Glottal stops and glottalization in German. Data and theory of connected speech processes. *Phonetica* 51, 38-51.

Lacy, de Paul (2006). *Markedness: Reduction and Preservation in Phonology*. Cambridge: Cambridge University Press.

Leal, Dorothy (1972). *Chitwan Tharu Phonemic Summary*. Kirtipur: Summer Institute of Linguistics and Institute of Nepal Studies, Tribhuvan University.

Lisker, Leigh, Abramson, Arthur S., Cooper, Franklin S. and Malcolm H. Schvey (1969). Transillumination of the larynx in running speech. *Journal of the Acoustical Society of America* 45, 1544-1546.

Lombardi, Linda (2002). Coronal epenthesis and unmarkedness. *Phonology* 19, 2, 219-251.

Lynch, John (2000). *A Grammar of Anejom*. Pacific Linguistics 507. Canberra: The Australian National University.

McCarthy, John (1993). A case of surface constraint violation. *Canadian Journal of Linguistics* 3, 127-153.

McCarthy, John (1991). Synchronic rule inversion. In: Sutton, Laurel, Christopher Johnson, & Ruth Shields (eds.). *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*. 192-297.

McCarthy, John (1993). A case of surface constraint variation. *Canadian Journal of Linguistics* 38, 127-153.

McMahon, April M.S. (2000). *Change, Chance, and Optimality*. Oxford: Oxford University Press.

McMahon, April, Foulkes, Paul & Laura Tollfree (1994). Gestural representation and lexical phonology. *Phonology* 11, 277-316.

Millardet, Georges (1923). *Linguistique et Dialectologie Romanes. Problèmes et Methods*. Paris: Champion.

Mithun, Marianne & Hasan Basri (1986). The phonology of Selayarese. *Oceanic Linguistics* 25 (1/2), 210–254.

Mudzingwa, Calisto (2007). *Hiatus resolution in Shona.* University of British Columbia. Ms.

Nespor, Marina & Irene Vogel (1988 [2007]). *Prosodic Phonology*. Berlin: Mouton de Gruyter.

Ohala, John J. (1974). Experimental historical phonology. In: Anderson, John M. & Charles Jones (eds.) *Historical Linguistics II: Theory and Description in Phonology*. Amsterdam: North-Holland Publishing, 353-389.

Ohala, John J. (1981). The listener as a source of sound change. In: Masek, Carrie S., Roberta A. Hendrick, & Mary F. Miller (eds.). *Papers from the Parasession on Language and Behavior*. Chicago: CLS, 178-203.

Ohala, John J. (1995). A probable case of clicks influencing the sound patterns of some European languages. *Phonetica* 52, 160-170.

Ortmann, Albert (1998). Consonant epenthesis: its distribution and phonological specification. In: Kehrein, Wolfgang and Richard Wiese (eds.). *The Phonology and Morphology of Germanic Languages*. Niemeyer: Tübingen, 51-76.

Payne, David L. (1981). *The phonology and morphology of Axinica Campa*. Arlington: Summer Institute of Linguistics & University of Texas at Arlington.

Pierrehumbert, Janet (1994). Prosodic effects on glottal allophones. In: Fujimura, Osamu and Minoru Hirano (eds). *Vocal Fold Physiology 8: Voice Quality Control*. Singular Press, San Diego, 39-60.

Pierrehumbert, Janet & David Talkin (1991). Lenition of /h/ and glottal stop. In: Kingston, John & Mary E. Beckman (eds.) *Papers in Laboratory Phonology II*. Cambridge: University Press, 90-117.

Pigott, Glyne L. (1980). *Aspects of Odawa Morphophonemics*. New York: Garland.

Pigott, Glyne L. & Rajendra Singh (1985). The phonology of epenthetic segments. *The Canadian Journal of Linguistics* 30, 247-405.

Pompino-Marschall, Bernd & Marzena Zygis (2010). Glottal marking of vowel-initial words in German. *ZAS Papers in Linguistics* (this volume).

Prince, Alan & Paul Smolensky (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms. Rutgers University & University of Colorado, Boulder.

Redi, Laura & Stefanie Shattuck-Hufnagel (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29, 407-429.

Rialland, Anne & Redouane Djamouri (1984). *Harmonie Vocalique, Consonantique et Structures de Dependance dans le Mot en Mongol Khalkha*. Bulletin de la Societe de Linguistique de Paris 79.

Rubach, Jerzy (1984). *Cyclic and Lexical Phonology. The Structure of Polish*. Dordrecht: Foris.

Rubach, Jerzy (2000).Glide and glottal stop insertion in Slavic languages: a DOT analysis. *Linguistic Inquiry* 31, 271-317.

Schutter de, Georges (1994). Dutch. In: König, Ekkehard & *Johan* van der Auwera (eds.). *The Germanic Languages*. London: Routledge. 439-477.

Selkirk, Elisabeth O. (1980a). The role of prosodic categories in English word stress. *Linguistic Inquiry* 11, 563-605.

Selkirk, Elisabeth O. (1980b). Prosodic domains in phonology: Sanskrit revised. In: Arnoff, Mark & Kean Mary-Louise (eds.). *Juncture. A collection of Original Papers*. Saratoga: Anma Libri. 107-129.

Shank, Scott & Ian Wilson (2000). An acoustic analysis of vowel formants in pharyngeal and glottal contexts in Nuu-chah-nulth. *University of Washington Working Papers in Linguistics* 19, 75-84.

Shixuan, Xu (2001). *The Bisu Language*. München: Lincom.

Smith, Jennifer L. (2001). Lexical category and phonological contrast. In: Kirchner, Robert, Pater, Joe & Wolf Wikely (eds.) *Proceedings of the Workshop on the Lexicon in Phonetics and Phonology*. Edmonton: University of Alberta, 61-72.

Thráinsson, Höskuldur (1994). Icelandic. In: König, Ekkehard & *Johan* van der Auwera (eds.). *The Germanic Languages*. London: Routledge. 142-189.

Uffmann, Christian (2007). Intrusive [r] and optimal epenthetic consonants. In: Carr, Philip & Patrick Honeybone. *Special volume of Languages Science* 29, 451-476.

Umeda, Noriko (1978). Occurrence of glottal stops in fluent speech. *JASA* 64, 81-94.

Vennemann, T (1988). *Preference Laws for Syllable Structure and the Explanation of Sound Change*. Berlin: Mouton de Gruyter.

Vincent, Nigel (1988). Italian. In: Harris, Martin & Vincent, Nigel (eds.). *The Romance Languages*. London: Routledge, 279-313.

Wells, John, C. (1982). *Accents of English*. Cambridge, U.K.: Cambridge University Press.

Westbury, John R. & Patricia A. *Keating* (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics 22*, 145-166.

Wetzels, Leo W. (1985). The historical phonology of intrusive stops. A nonlinear description. *Canadian Journal of Linguistics* 30, 285-333.

Wiese, Richard (2000). *The Phonology of German*. Oxford: Oxford University Press.

Zwicky, Arnold M. (1972). Note on a phonological hierarchy in English. In: Stockwell, Robert P. & Roland K. S. Macaulay (eds.). *Linguistics Change and Generative Theory*. Bloomington: Indiana University Press, 275-301.

Zec, Draga. (2005). Prosodic Differences among Function Words. *Phonology* 22, 77-112.

Zygis, Marzena (2009). A perceptual study on glide insertions. ZAS, Berlin. Ms.

Some Properties of Prosodic Phrasing in Thompson Salish

Karsten A. Koch
Zentrum für Allgemeine Sprachwissenschaft, Berlin

In Nłeʔkepmxcin, consonant-heavy inventories, lengthy obstruent clusters and widespread glottalization can make potential F0 cues to prosodic phrase boundaries (e.g. boundary tones or declination reset) difficult to observe phonetically. In this paper, I explore a test that exploits one behaviour of phrase-final consonant clusters to test for prosodic phrasing in Nłeʔkepmxcin clauses. Final /t/ of the 1pl marker *kt* is aspirated when phrase-final, but not phrase-internally. Use of this test suggests that Thompson Salish speakers parse verbs, arguments and adjuncts into separate phonological phrases. However, complex verbal predicates and complex noun phrases are parsed as single phonological phrases. Implications are discussed, especially in regards to findings that (absence of) pitch accent is not employed to signal the informational categories of Focus and Givenness, even though Nłeʔkepmxcin is a stress language.

1 Introduction

The Salish languages of the Pacific Northwest of North America are well known for their rich consonantal inventories, widespread glottalization, and lengthy obstruent clusters (e.g. Kinkade 1992, Bagemihl 1991, Shaw 2002). Because obstruents are well known to affect the pitch of adjacent resonants (e.g. Brown and Thompson 2006 on Upriver Halkomelem Salish), it can be very difficult to measure potential F0 cues to prosodic phrasing, such as boundary tones and declination reset, in Salish languages. In this paper, I explore an alternative phonetic cue to prosodic phrasing in Nłeʔkepmxcin (Thompson River Salish), one that in fact takes advantage of the widespread distribution of obstruents. Specifically, I will show that the final /t/ of the 1st person plural marker *kt* is aspirated in phrase-final position, but not phrase-internally.

Application of this test in different positions in the Thompson Salish clause will lead to the following conclusions about prosodic phrasing: (i) verb, arguments and adjuncts are parsed into separate prosodic phrases (unlike for example, English, where verb and object are often parsed into a single prosodic phrase), and (ii) complex predicates (auxiliary-verb constructions) and complex Noun Phrases are parsed into single prosodic phrases. Results are discussed in light of evidence that Salish languages, despite being stress languages, do not mark informational prominence through pitch accent.

In this paper, I will be referring to phonological phrases (p-phrase) and intonational phrases (i-phrases) in the prosodic hierarchy of Nespor and Vogel (1986, also Hayes 1989). I will be primarily interested in determining what syntactic constituents map into phonological phrases, currently a well-studied issue in the interface of syntax and phonology (e.g. Truckenbrodt 1995, Legate 2003, Selkirk and Kratzer 2007, An 2007, Kandybowicz 2009, etc.).

2 Background

Nłeʔkepmxcin is one of 23 Salish languages (Thompson and Thompson 1992, 1996; Kinkade 1992, Czaykowska-Higgins and Kinkade 1998, Kroeber 1999, for some general overviews of Salishan). It is spoken in the southwest of British Columbia, and is severely endangered, with no more than a few hundred elderly speakers remaining. The phonemic inventory is given in Table 1.

Table 1: Phonemic inventory (adapted from Thompson and Thompson 1992)

CONSONANTS	labial	alveolar	alveo-palatal	velar	uvular	pharyngeal	glottal
Stops	p	t		k kʷ	q qʷ		ʔ
Ejectives	p̓	t’		k̓ k̓ʷ	q̓ q̓ʷ		
Lateral Eject.		ƛ̓					
Nasal	m	n					
Glottalized	m̓	n̓					
Affricates		c̣ [ts]	c [tʃ]				
Ejective		ċ [ts’]					
Fricatives		ṣ [s]	s [ʃ]	x xʷ	x̣ x̣ʷ		h
Lateral		ł					
Approximant	(w)	z	y [j]	w		ʕ ʕʷ	
Lateral		l					
Glottalized	(w̓)	z̓	y̓	w̓		ʕ’ ʕ’ʷ	
Glott. Lateral		l’					

VOWELS	front		central		back
high	i	ị			u
mid	e		ə	ə̣	o
low			a		

Like all Salish languages, Thompson Salish is predicate-initial. The typical order is Verb-Subject-Object-Adjunct, though post-predicative verb order is in practice quite flexible. Predicates are obligatorily inflected for transitivity and subject/object agreement markers (see Thompson and Thompson 1992). Second position clitics (2CL) follow the first prosodic word. DPs are obligatorily marked with determiners. A typical sentence is shown in (1).[1]

(1)	Verb	2CL	Subject	Object
	kən-t-Ø-és	=xeʔ	e=skíxzeʔ-kt	e=sínciʔ-kt.
	help-TR-3O-3S	=DEM	DET=mother-our	DET=brother-our
	'Our mother helped our brother.'			

There has been little previous research on properties of prosodic phrasing in the language: the grammar mentions a few general pitch cues (Thompson and Thompson 1992:24), while Egesdal (1984) details some general rhythmic properties of narratives, again only impressionistically. Koch (2008) is the first work to examine potential phonetic cues to prosodic phrasing, and the current paper follows up on this work.

Looking across the Salish language family more generally, there again has been much work on prosodic categories below the level of phrases (e.g. Shaw 2002, Czaykowska-Higgins 1993, 1998, Thompson and Thompson 1992, etc.), but comparatively little at the phrasal level. A notable exception, Beck (1996, 1999) identifies the following indicators of p-phrase status in Lushootseed Salish (see also Beck and Bennett 2007):

(2) Characteristics of phonological phrases in Lushootseed Salish (Beck 1999)
   a. set off by 50-100 ms pause in careful speech
   b. lack phonological interaction (i.e. assimilation, etc.) across p-phrase boundaries
   c. contain a single phonological word with an amplitude peak plus clitics and affixes

In addition, Beck (1999) notes that intonational phrases in Lushootseed are characterized by a steady fall in F0, with a declination reset at the start of each i-phrase. In Okanagan Salish, prosodic boundaries are also marked by pauses, F0 fall, and reset or partial reset of declination across phrasal boundaries (Barthmaier 2004). Finally, recent work by Caldecott (2009) shows that

[1] See the appendix for a key to orthography and glosses.

prosodic phrases are right-headed in St'át'imcets Salish; Koch (2008) finds that Thompson Salish, too, has rightmost nuclear stress and right-headed phonological-phrases.

In the remainder of this paper, I explore a (lack of) assimilation effect in the spirit of (2b): phrase-final aspiration of /t/ in the 1st person plural marker *kt*.

3 Phrase-final aspiration of *kt*: A test to distinguish p-phrase boundaries

The enclitic or suffix *kt* indicates 1st person plural (1pl) subjects in both indicative and nominalized intransitive clauses, as well as 1st person plural possessors. In this section, I present phonetic evidence that the /t/ of the 1st person plural marker *kt* is aspirated phrase-finally, but not phrase internally. I start by showing this in simple verb phrases (3.1). Next, I show that the aspiration test indicates that, verb, subject and object are phrased separately, as are verbs and adjuncts (3.2). Finally, I show that the language does not simply parse each prosodic word (PWd) into a phrase (contra Beck 1999, who suggests this may happen in Lushootseed Salish – 2c; see also Hellmuth 2006 who argues that in Cairene Arabic each PWd is pitch accented, which is not the case here). First, I show that complex verbal predicates consisting of one or more auxiliaries and a verb are parsed as one p-phrase (3.3). I will close by showing that complex Noun Phrases consisting of an NP and a modifier are similarly parsed as a single p-phrase (3.4).

All data in this paper come from my own data corpus collected during fieldwork with two speakers of the ƛ̓q̓emcín (Lytton) dialect of Nłeʔkepmxcin. Speakers were recorded on separate channels using a digital audio recorder and individual microphones. The utterances examined in this paper all stem from a single breath group (single intonational phrase in the prosodic hierarchy).

3.1 Aspiration of *kt* in simple clauses: restriction to phrase-final position

In (3a), the 1pl indicative enclitic *kt* occurs sentence-finally, while in (3b) it is followed by the evidential marker *nukʷ*.

(3)	a.	téyt = kt. hungry=1PL.INCL 'We are hungry.'	b.	téyt = kt = nke. hungry=1PL.INCL=EVID 'We are hungry.'

When the 1pl marker *kt* occurs in a clearly phrase-final position – the end of a sentence — it is strongly aspirated (in itself a noteworthy property of Thompson Salish). This is shown below in the sentence *téyt kt* 'We are hungry.' In fact, all three of the final stops are aspirated; for our purposes we are concerned with phrase-final /t/ of the 1pl *kt* marker.

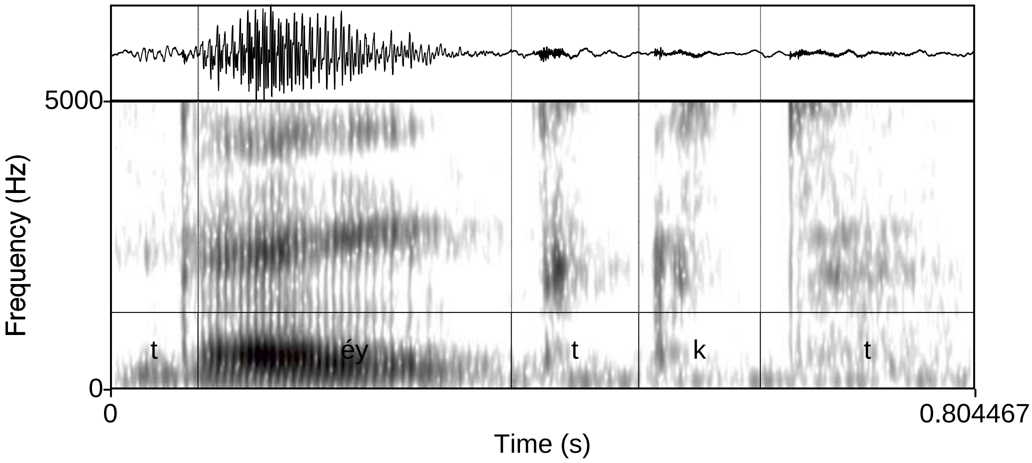


Figure 1: Final aspiration of /t/ in *kt* (3a)

When the 1pl marker *kt* is followed by another enclitic, the evidential *nke,* the final /t/ is no longer aspirated. I take this to show that the clitic string *=kt=nke* has undergone phrase-internal assimilation. Note that within the 1pl clitic *kt*, the /k/ is still strongly aspirated (as it is in all the examples we will see). Thus, the final aspiration of *kt* really is a boundary effect.

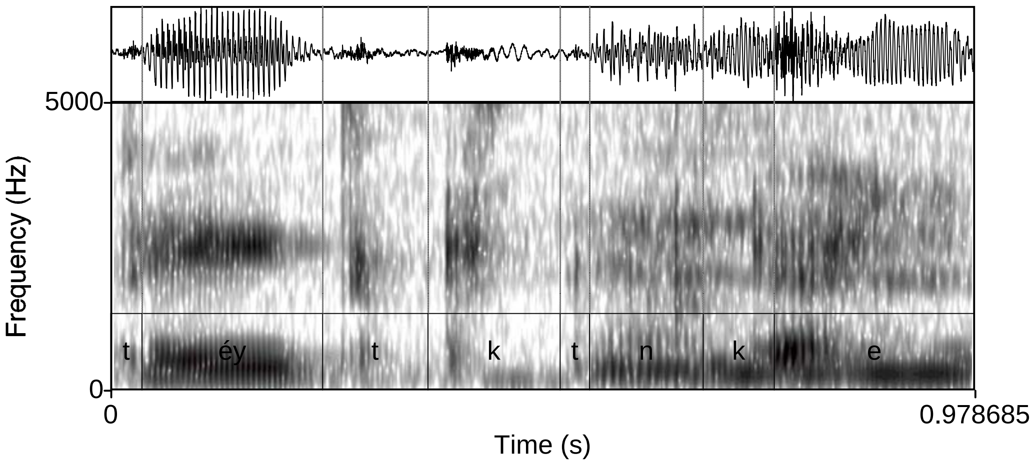


Figure 2: No aspiration of /t/ in *kt* when phrase-internal (3b)

3.2 Aspiration of *kt* between constituents of more complex clauses

In this section, I use the *kt* aspiration test to probe for prosodic phrase boundaries between major constituents within a clause. Following Nespor and Vogel (1986), the entire clause is parsed into a single i-phrase. Thus, this test will probe for p-phrase boundaries (though nothing hinges on the model used).

In (4), both subject and object are suffixed with *kt* to mark 1pl possession.

(4) kən-t-Ø-és =xeʔ e=skíxzeʔ-kt e=sínciʔ-kt.
help-TR-3O-3S =DEM DET=mother-our DET=brother-our
'Our mother helped our brother.'

As expected, the sentence-final [t] is aspirated (figure 3). In addition, however, we see that the 1pl marker *kt* after the subject *skixzeʔ* 'mother' is also aspirated. This indicates a phrase boundary between subject and object. Figure 3 shows just the two final Noun Phrases *skixzeʔ-kt e sinciʔ-kt,* and the two occurrences of *kt* are marked. At the same time, we see that the sentence-final aspiration is longer, as we would expect if the entire clause is parsed in a right-headed intonational phrase with increased final lengthening on the dominant constituent.

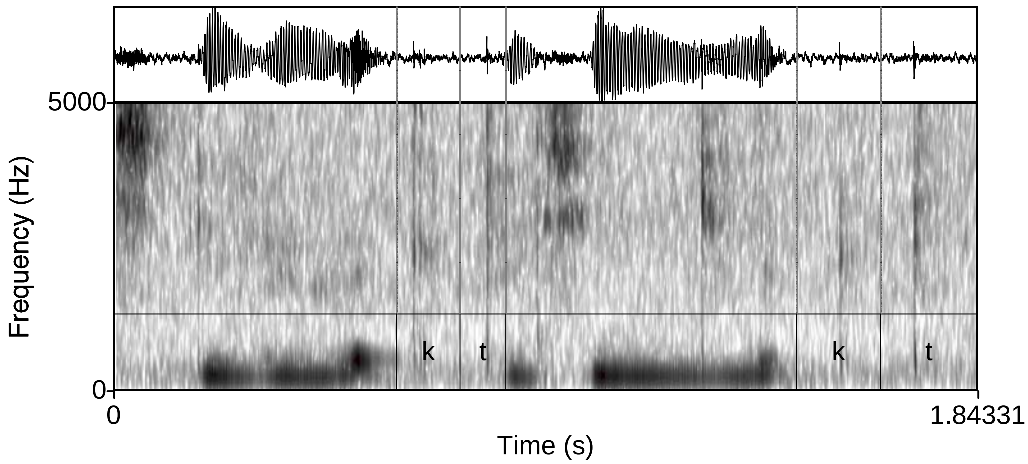


Figure 3: Aspiration of /t/ in *kt* between subject and object of (4)

In the next example, the verb *nʕʷoy̓t* 'sleep,' bearing the 1pl possessive subject enclitic *kt,* is followed by an adjunct, the temporal adverb *ɬ sitist* 'last night.'

(5) y̓é e=s=n-ʕʷóy̓t=kt ɬ=sítist.
good COMP=NOM=LOC-sleep=1PL.POCL DET=night
'We slept really good last night.'

In Figure 4, we see aspiration of the /t/ in *kt* before the fricatives [ɬ ʃ] of the temporal adjunct phrase. The *kt* aspiration test shows that verbs and adjuncts are parsed into separate prosodic phrases.

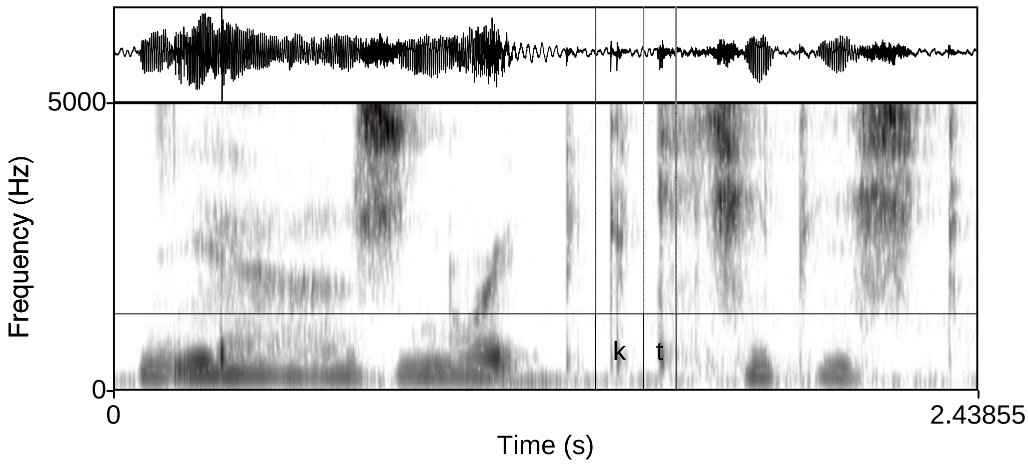


Figure 4: Aspiration of /t/ in *kt* between verb and temporal adjunct

Finally, in (6), the verb *ṗeṅt* 'return' bears the 1pl possessive subject enclitic *kt*, and is followed by the Preposition Phrase *w e* ṣ*kul* 'to school.'

(6) ʔe s = ṗéṅt = kt w = e = ṣkúl.
and NOM=return=1PL.POCL to=DET=school
'And we went back to school.'

In figure 5 we see aspiration of the final [t] of *kt*, suggesting that the verb is parsed into a separate phonological phrase from the PP adjunct.

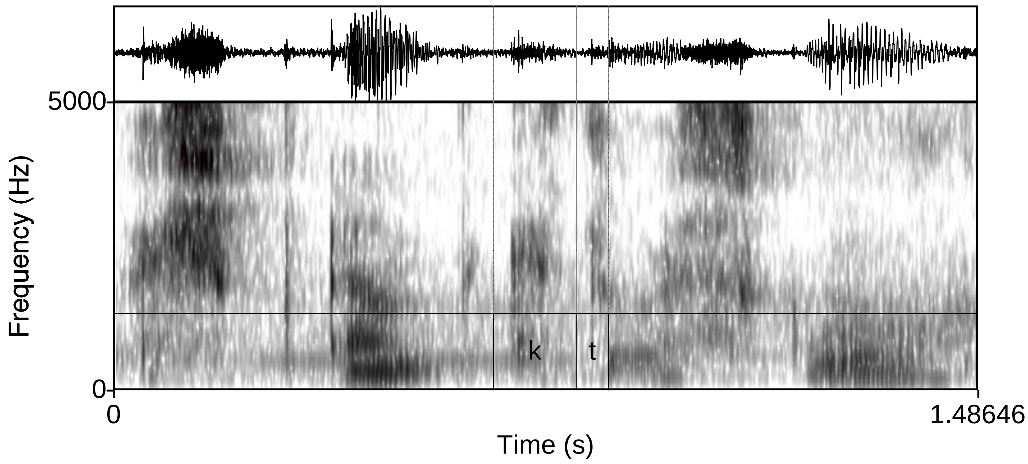


Figure 5: Aspiration of /t/ in *kt* between verb and PP adjunct

In this section, I showed that the *kt* aspiration test suggests that verbs, arguments and adjuncts are parsed into separate phonological phrases.

3.3 Complex verbal predicates are parsed as one p-phrase

The verb may co-occur with one or more auxiliaries at the left edge of the Thompson Salish clause. When this happens, the first auxiliary attracts the second position clitics. This shows us that auxiliaries count as prosodic words. In the previous section, all examples were consistent with a phonological system in which each prosodic word is parsed into its own p-phrase, bearing its own pitch accent (see Beck 1999, 2c, on Lushootseed Salish, Hellmuth 2006 on Cairene Arabic). In this section, I show that this cannot be right for Nłeʔkepmxcin, since auxiliaries and verbs are parsed into a single prosodic unit, by the *kt* aspiration test, even though both count as prosodic words.

In (7), the 1pl subject marker *kt* follows the future auxiliary *xʷuỷ*, and precedes a second auxiliary *nes* and the verb *tewcnme* ‘shop for groceries.’

(7) xʷúỷ=kt nés téw-cn-me.
FUT=1PL.INCL go buy-mouth-INTRANS
‘We’re going to go grocery shopping.’

Figure 6 illustrates that the [t] of *kt* is completely unaspirated, assimilating with the following [n]. By hypothesis, *kt* is not followed by a phrasal boundary.

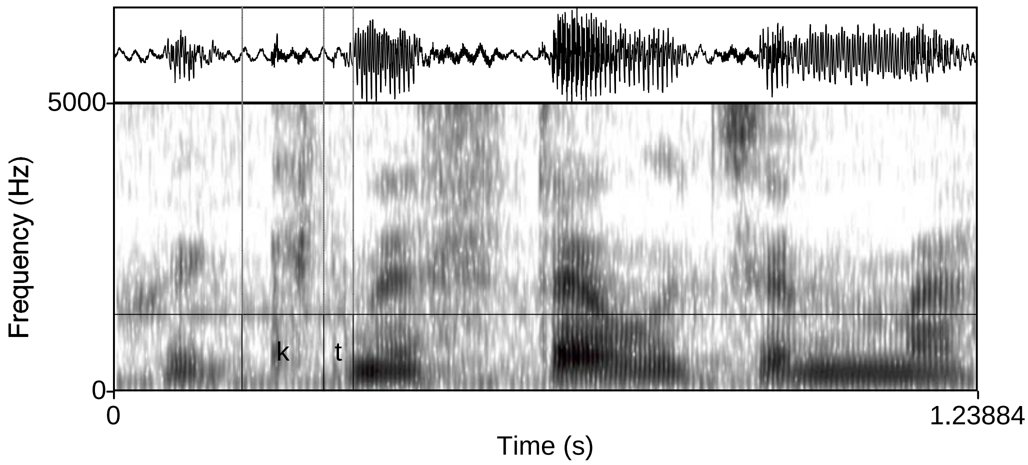


Figure 6: Non-aspiration of /t/ in *kt* within verbal complex in (7)

Another case is shown in (8) and figure 7. The imperfective auxiliary *wʔex* carries the 1pl possessive enclitic *kt*, and is followed by the verb *tans* 'dance.' Figure 7 shows that there is no release of the /t/ of *kt* at all, which has instead assimilated with the onset [t] of the verb *tans*.

(8) ʔe s=wʔéx=kt táns.
and NOM=IMPF=1PL.POCL dance
'And so we danced.'

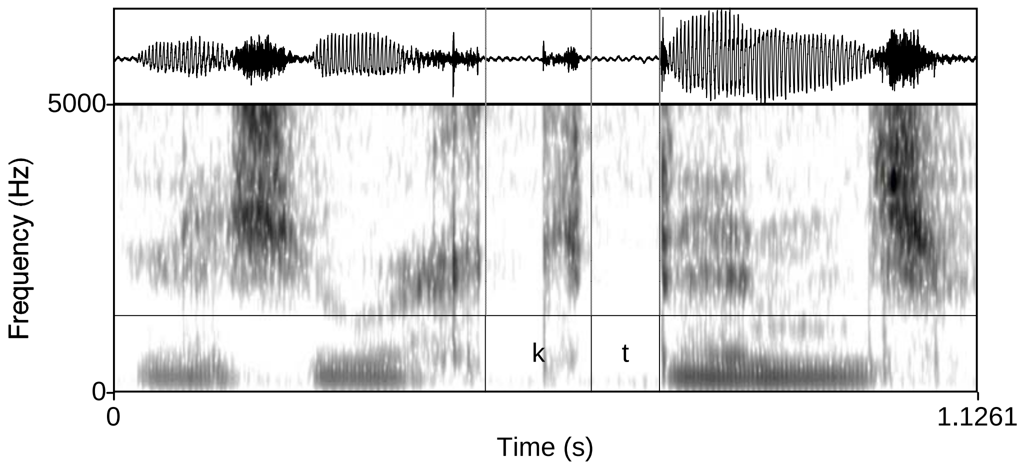


Figure 7: Non-aspiration of /t/ in *kt* within verbal complex in (8)

3.4 Complex Noun Phrases are parsed as one p-phrase

In the last section, I showed that the *kt* aspiration test indicates that more than one word can be parsed into a single phonological phrase: auxiliaries and verbs. Now I show that the *kt* aspiration test gives the same result for Noun Phrases consisting of nouns and modifiers.

In (9), 'our son' is expressed as the noun *skʷuzeʔ* 'offspring' modified by *sqayxʷ* 'man' (or 'male'). The 1pl possessor *kt* intervenes. Figure 8 shows that the final [t] is not aspirated, again assimilating with the onset [t] of the 'link' particle *te* (this marks predicate modification between nouns and modifiers).

(9) e = skʷúzeʔ-kt te = sqáyxʷ
DET=offspring-1PL.POSS LINK=man
'our son' (more literally 'our male offspring')

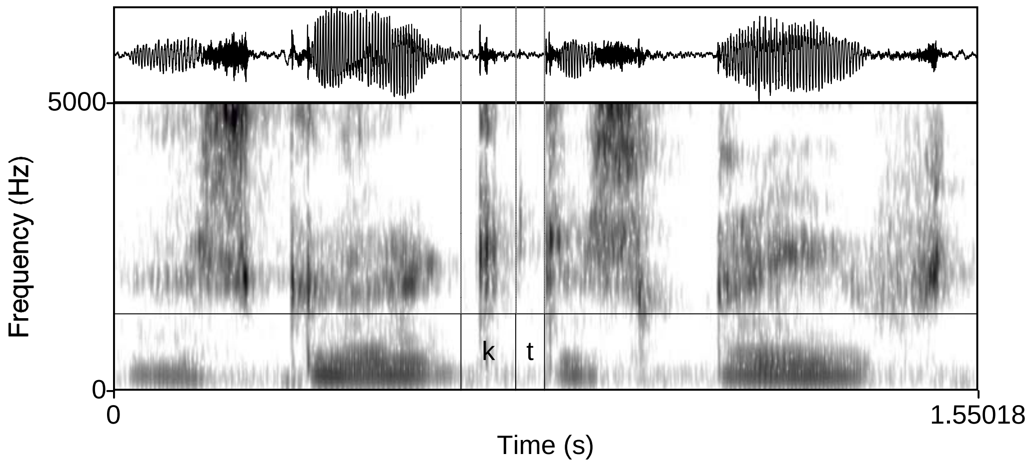


Figure 8: Non-aspiration of /t/ in *kt* within the complex Noun Phrase in (9)

The noun *smiyc* 'meat' is modified by a head-initial relative clause *skʷukʷ kt* 'that we cooked' in (10). Like in the previous example, the final [t] of *kt* is not aspirated, again assimilating with the onset [t] of the link particle *te*.

(10) ... e = s = kʷúkʷ = kt te = smíyc.
... DET=NOM=cook=1PL.POCL LINK=meat
'... the meat that we cooked.'

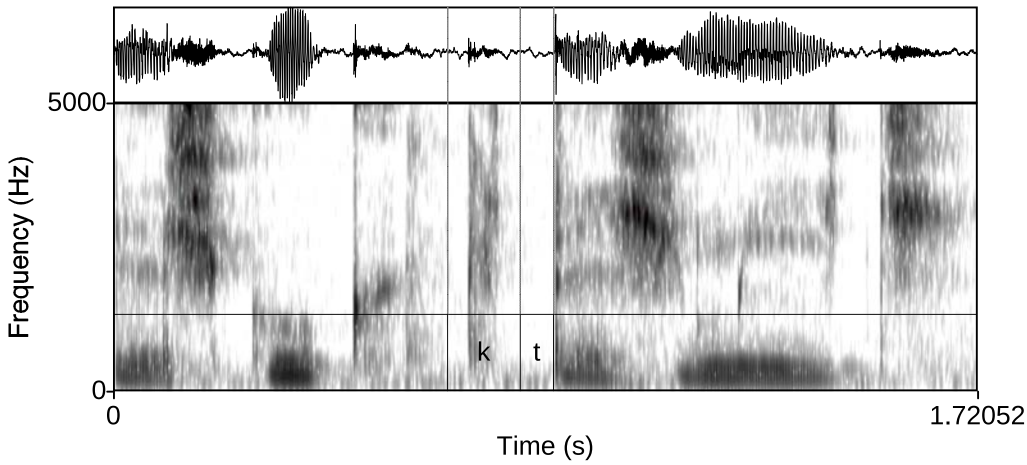


Figure 9: Non-aspiration of /t/ in *kt* within the complex Noun Phrase in (10)

In the final example shown here, the noun 'cat' in (11a) bears the 1pl possessive suffix *kt*, and is followed by the modifier *nmímł* 'our [emphatic].' The /t/ of *kt* is not aspirated, assimilating with the following [n]. Emphatic modifiers count as Prosodic Words, since when clefted, they attract second position clitics like any other Prosodic Word (shown for the 1sg emphatic *ncewe*ʔ in 11b; see Koch 2008 for further examples). Thus, this is another case where two Prosodic Words are parsed into one larger prosodic unit, the phonological phrase.

(11) a. e = púṣ-kt nmímł
DET=cat-1PL.POSS 1PL.EMPHATIC
'our cat'

b. ncéweʔ = us = meł k = ʔém̓c-t-Ø-mus e = púṣ.
1SG.EMPH=3CNCL=indeed COMP=feed-TRANS-3O-SUBJ.GAP DET=cat
'Let it be me that feeds the cat.'

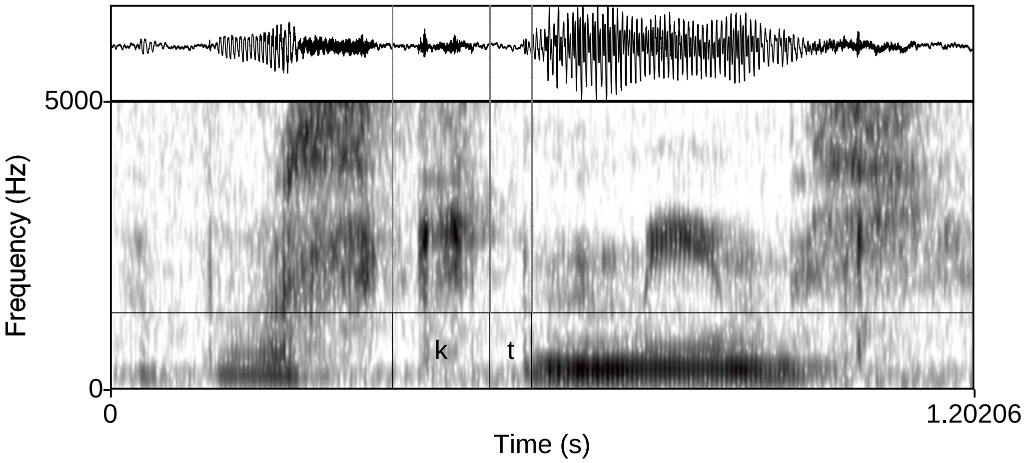


Figure 10: Non-aspiration of /t/ in *kt*
within the complex Noun Phrase in (11a)

In this section, I showed that the *kt* aspiration test suggested that complex noun phrases are parsed as single phonological phrases.

4 Implications

The *kt* aspiration test probes for phrasal boundaries within the Thompson Salish clause. In section 3, the results of the test suggest that syntactic categories are phonologically phrased in the following ways.

First, verbs and arguments are parsed into separate phonological phrases. Cross-linguistically, this is not uncommon. Beck (1999) and Barthmaier (2004) make similar claims for Lushootseed Salish and Okanagan Salish, respectively. Outside the Salish language family, Hayes and Lahiri (1991, on Bengali), Schafer and Jun (2002, on Korean), and Nespor and Sandler (1999, on Israeli Sign Language), also argue for parsing of verb and arguments into individual p-phrases (see also Ishihara 2007: 147-148, ex. 17b, for such parses of some Japanese sentences). This parsing is not typical of English, where verb and object are typically parsed into one phonological phrase, while the subject is realized in a separate p-phrase (Chomsky 1971, Jackendoff 1972, Gussenhoven 1983, Selkirk 1995, Kahnemuyipour 2004, Selkirk and Kratzer 2007). It is possible that the parsing pattern observed in Thompson Salish is correlated with a surface word order of Verb-Subject-Object (VSO), where the subject intervenes between verb and object, though this is a matter for further typological research.

Secondly, complex predicates (auxiliaries plus verb) and complex Noun Phrases (noun plus modifier) are parsed into a single phonological phrase. This is significant because it suggests that the language does not employ a strategy where each prosodic word is pitch accented and parsed into a phonological phrase independently of its greater syntactic structure. Rather, an intermediate category, the phonological phrase, exists between the word and i-phrase levels. This category maps onto syntactic categories (DP, and the extended VP), consistent with the idea that syntactic and phonological categories interface at the level of the p-phrase (e.g. Truckenbrodt 1995, 1999, Selkirk and Kratzer 2007, and many others).

(12) XP-to-P Mapping Condition (Truckenbrodt 1999:221)
Mapping constraints relate [syntactic] XPs to phonological phrases,
but do not relate XPs to other prosodic entities.

Phonological phrases are right-headed (Koch 2008; Caldecott 2009 on St'át'imcets Salish), and in stress languages, this is the category where focus is made prosodically prominent: focused items are heads of p-phrases. Previous findings indicate that, although it is a stress language (Thompson and Thompson 1992, Egesdal 1984), speakers of Nłeʔkepmxcin do not manipulate pitch accent cues to mark the informational categories of focus and givenness. That is, there are no "Stress-Focus" or "Destress-Given" effects (Koch 2008). One possible reason would have been that pitch accents are assigned at the level of the Prosodic Word (Hellmuth 2006 on Cairene Arabic), and thus there would be no opportunity for manipulating headedness at a higher level. However, the current study suggests this is not the case: the language does have p-phrases that culminate, but they are simply not exploited to mark information structure. In the terms of Selkirk and Kratzer (2007), for example, the constraints STRESS-FOCUS and DESTRESS-GIVEN (e.g. Féry and Samek-Lodovici 2006) are not part of the syntax-phonology interface in the grammar of Nłeʔkepmxcin, at least not in the way that they are commonly defined. This is a significant finding, given the widespread assumption that stress languages employ stress to mark focus; it may be that this is far less widespread once we stray outside the European realm (see also Rialland and Robert 2001 on Wolof, and Lindström and Remijsen 2005 on Kuot).

On the other hand, it has been observed that a general strategy for marking the focus in many (perhaps all) Salish languages is to make the focus part of the predicate (e.g. Kroeber 1997, 1999 for overviews of clefting strategies, Benner 2006 on Sencóthen, Davis 2007 on St'át'imcets, Koch 2008 on Thompson). If, as the current findings suggest, a complex predicate is a single phonological phrase, then this apparent syntactic focus-marking strategy may have a prosodic purpose as well: the focus is restructured into a single p-phrase — the initial p-phrase in the clause. While such a strategy has not been widely reported for stress languages, the manipulation of prosodic phrasing to mark focus is well-known from work on many tone languages in particular (e.g. Truckenbrodt 1999, Downing 2003, Ladd 1996 more generally on the role of phrasing).

Restructuring the focus into the initial p-phrase is consistent with a strategy that makes informationally prominent categories quickly recoverable from the speech stream for listeners. This is a strategy in line with psycholinguistic work that suggests that intonational parsing happens more rapidly than syntactic parsing, and is used to identify syntactic phrasing (Kjelgaard and Speer 1999; Jun 2003, and references on p. 220; Fodor's 1998 Implicit Prosody Hypothesis on silent reading; Callan et al. 2004 on listeners internally simulating the speech act of speakers). Kjelgaard and Speer suggest that prosodic parsing is more straighforward because it is easier to identify p-phrases and i-phrases than syntactic information. P-phrases and i-phrases have only edges and heads, and are parsed directly into each other; moreover, there are only two categories to identify. Syntactic parsing is much more complex, involving the identification of many syntactic categories, movement and traces. Moreover, signal information that demarcates phrase edges and heads can be recovered not just from the acoustic signal, but also from the visual signal (eg. Vatikiotis-Bateson 1988): acoustic parameters like F0 (Yehia et al. 2002), duration (Vatikiotis-Bateson 1988, Fletcher and Bateson 1989), and amplitude (Vatikiotis-Bateson 1988, Vatikiotis-Bateson and Kelso 1993) have visual reflexes in facial and head movement. In addition, neurolinguistic processing research provides some support for the view that p-phrase and i-phrase processing is different: evidence suggests that linguistic prosody over small domains (words or less) may be controlled by the left hemisphere; but processing of larger units (eg. p-phrases and i-phrases), appears to span both hemispheres (Baum and Pell 1999).

If the absence of stress-focus effects in Thompson Salish is not accounted for by the lack of phonological phrases, there may be other functional explanations. One possible reason is that F0 perturbations are important cues to obstruents and glottalization, and are thus not manipulated for information structure purposes. Given the widespread glottalization in the phonemic inventory (table 1), and the lengthy obstruent clusters in surface strings in the language, this seems a possible explanation worthy of further research.

Thus, the implications of the findings are widespread for evaluation of the grammar of the language, in particular the syntax-phonology interfaec and the system of focus marking. This points to the importance of finding further phonetic cues that will also help to identify prosodic phrases in Thompson Salish.

5 Conclusion

In this paper, I have used a consonant-oriented test for prosodic phrasing cues in Nłeʔkepmxcin (Thompson River Salish). Phrase-final aspiration of the /t/ in *kt* (1pl) and its non-aspiration when not at a p-phrase boundary was exploited to probe the phrasing of Thompson clauses. While verbs and arguments are parsed into different p-phrases, complex predicates and complex Noun Phrases are parsed into a single p-phrase. This finding has implications for how the syntax-phonology interface operates in Salish, both for the mapping of syntactic XPs onto phonological categories, and for the mapping of information structure into phonological categories in the absence of a stress-focus effect.

Acknowledgements

I am indebted to consultants Flora Ehrhardt and Patricia McKay, without whom this research would not be possible. This research has been supported by Jacobs and Kinkade Research Grants from the Whatcom Museum Foundation, Bellingham, WA; and by NSERC, SSHRC, and DAAD research fellowships.

References

An, Duk-Ho (2007). Clauses in non-canonical positions at the syntax-phonology interface. *Syntax* 10(1), 38-79.

Bagemihl, Bruce (1991). Syllable structure in Bella Coola. *Linguistic Inquiry* 22, 589–646.

Barthmeier, Paul (2004). Intonation units in Okanagan. In Donna B. Gerdts and Lisa Matthewson, eds. *Studies in Salish linguistics in honor of M. Dale Kinkade.* Missoula: UMOPL 17. 30-42.

Baum, Shari R., and Marc D. Pell (1999). The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology* 8, 581-608.

Beck, David (1996). Some notes on phonological phrasing in Lushootseed. *Toronto Working Papers in Linguistics* 15(2), 37-60.

———— (1999). Words and prosodic phrasing in Lushootseed. In T. Alan Hall & Ursula Kleinhenz, eds. *Studies on the Phonological Word*. Amsterdam: Benjamins. 23-46.

————, and David Bennett (2007). Extending the prosodic hierarchy: evidence from Lushootseed narrative. *Northwest Journal of Linguistics* 1(1), 1 – 34.

Benner, Allison (2006). *The prosody of Senchóthen.* Paper given at The 41st International Conference on Salish and Neighbouring Languages, University of Victoria.

Brown, Jason, and James Thompson (2006). Criteria for tonality: The case of Upriver Halkomelem. In Shannon T. Bischoff, Lynnika Butler, Peter Norquest, Daniel Siddiqi eds. *MIT Working Papers on Endangered and Lesser Known Languages: Studies in Salishan*. 113-126.

Caldecott, Marion (2009). *Non-exhaustive parsing: phonetic and phonological evidence from St'át'imcets.* Ph.D. dissertation, University of British Columbia.

Callan, Daniel E., Jeffery A. Jones, Kevin Munhall, Christian Kroos, Akiko M. Callan, and Eric Vatikiotis-Bateson (2004). Neural processes underlying perceptual enhancement by visual speech gestures. *NeuroReport* 14, 2213-2218.

Chomsky, Noam (1971). Deep structure, surface structure, and semantic interpretation. In D.D.Steinberg and L.A. Jakobovits, eds. *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology.* Cambridge: Cambridge University Press. 183-216.

Czaykowska-Higgins, Ewa (1993). Cyclicity and stress in Moses-Columbia Salish (Nxa'amxcín). *Natural Language & Linguistic Theory* 11, 197-278.

——— (1998). The morphological and phonological constituent structure of words in Moses-Columbia Salish (Nxa'amxcín). In E. Czaykowska-Higgins and M.D. Kinkade, (eds.) *Salish Languages and Linguistics*. Berlin: Mouton de Gruyter. 153–196.

———, and M. Dale Kinkade (eds.) (1998). *Salish Languages and Linguistics: Theoretical and Descriptive Perspectives. Trends in Linguistics: Studies and Monographs 107*. Berlin: Mouton de Gruyter.

Davis, Henry (2007). *Prosody-focus dissociation and its consequences: The case of Salish.* Paper presented November 10, 2007, Nagoya, Japan.

Downing, Laura J. (2003). Stress, tone and focus in Chichewa and Xhosa. In Rose-Juliet Anyanwu, ed. *Stress and Tone: The African Experience*. *Franfurter Afrikanisitische Blätter* 15. Köln: Rüdiger Köppe Verlag. 59-81.

Egesdal, Steven M. (1984). *Stylized Characters' Speech in Thompson Salish Narrative.* Ph.D. dissertation, University of Hawaii.

Féry, Caroline, and Vieri Samek-Lodovici (2006). Focus projection and prosodic prominence in nested foci. *Language* 82(1), 131-150.

Fletcher, Janet, and Eric Vatikiotis-Bateson (1989). The kinematics of final lengthening in French. *Journal of the Acoustical Society of America* 86, 114-115.

Fodor, J.D. (1998). Learning to parse. *Journal of Psycholinguistic Research* 27, 285–319.

Gussenhoven, Carlos (1983). Testing the reality of focus domains. *Language and Speech* 26, 61-80.

Hayes, Bruce (1989). The prosodic hierarchy in meter. In P. Kiparsky and G. Youmans (eds.) *Phonetics and Phonology 1: Rhythm and Meter.* San Diego: Academic Press. 201-260.

———, and Aditi Lahiri (1991). Bengali intonational phonology. *Natural Language & Linguistic Theory* 9, 47-96.

Hellmuth, Sam (2006). *Intonational pitch accent distribution in Egyptian Arabic.* PhD dissertation, SOAS, University of London.

Ishihara, Shinichiro (2007). Major phrase, focus intonation, multiple spell-out. *The Linguistic Review* (2-3), 137-167.

Jackendoff, Ray (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jun, Sun-Ah (2003). Prosodic phrasing and attachment preferences. *Journal of Psycholinguistic Research* 32(2), 219-249.

Kahnemuyipour, Arsalan (2004). *The Syntax of Sentential Stress.* Ph.D. dissertation. University of Toronto.

Kandybowicz, Jason (2009). Embracing edges: syntactic and phono-syntactic edge sensitivity in Nupe. *Natural Language and Linguistic Theory* 27, 305-344.

Kinkade, M. Dale (1992). Salishan languages. In: *International Encyclopedia of Linguistics* (Bright, William, ed.). New york: Oxford University Press. 359-362.

Kjelgaard, M.M., and S.R. Speer (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language* 40(2), 153-194.

Koch, Karsten. 2008. *Intonation and Focus in Nłeʔkepmxcin (Thompson River Salish).* Ph.D. dissertation, University of British Columbia, July 2008. https://circle.ubc.ca/dspace/handle/2429/2848

Kroeber, Paul (1997). Relativization in Thompson Salish. *Anthropological Linguistics* 39(3), 376-422.

——— (1999). *The Salish Language Family: Reconstructing Syntax.* Lincoln: University of Nebraska Press.

Ladd, D. Robert (1996). *Intonational Phonology*. Cambridge: Cambridge UP.

Legate, Julie Ann (2003). Some interface properties of the phase. *Linguistic Inquiry* 34, 506-516.

Lindström, Eva, and Bert Remijsen (2005). Aspects of the prosody of Kuot, a language where intonation ignores stress. *Linguistics* 43(4), 839-870.

Nespor, Marina, and Wendy Sandler (1999). Prosody in Israeli Sign Language. *Language and Speech* 42, 143-176.

———, and Irene Vogel. 1986. *Prosodic Phonology.* Dordrecht: Foris.

Rialland, Annie, and Stéphane Robert (2001). The intonational system of Wolof. *Linguistics* 39(5), 893-939.

Schafer, A.J., and Sun-Ah Jun (2002). Effects of accentual phrasing on adjective interpretation in Korean. In M. Nakayama (ed.) *East Asian Language Processing*. Stanford: CSLI. 223-255.

Selkirk, Elizabeth. 1995. Sentence prosody: Intonation, stress and phrasing. In J. Goldsmith, ed. *The Handbook of Phonological Theory.* Cambridge, MA: Blackwell. 550-569.

———, and Angelika Kratzer (2007). Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24(2-3), 93-135.

Shaw, Patricia (2002). On the edge: obstruent clusters in Salish. In L. Bar-el, L. Watt, and I. Wilson, eds. *Proceedings of WSCLA 7.* Vanvouver: UBCWPL 10. 119–136.

Thompson, Laurence C., and M. Terry Thompson (1992). *The Thompson Language*. Missoula: University of Montana Occasional Papers in Linguistics 8.

——— (1996). *Thompson River Salish Dictionary*. Missoula: University of Montana Occasional Papers in Linguistics 12.

Truckenbrodt, Hubert (1995). Phonological phrases: their relation to syntax, focus and prominence. Ph.D. dissertation. MIT.

——— (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30, 219-255.

Vatikiotis-Bateson, Eric (1988). *Linguistic Structure and Articulatory Dynamics*. Ph.D. Dissertation. Bloomington, IN: Indiana University Linguistics Club.

Vatikiotis-Bateson, Eric, and J. A. S. Kelso (1993). Rhythm type and articulatory dynamics in English, French, and Japanese. *Journal of Phonetics* 21. 231-265.

Yehia, Hani C., Takaaki Kuratate, and Eric Vatikiotis-Bateson (2002). Linking facial animation, head motion and speech acoustics. *Journal of Phonetics* 30(3), 555-568.

Appendix

Data are presented in the orthography developed in Thompson and Thompson (1992, 1996). Acute accent ´ indicates word-level stress. Symbols not listed are the standard IPA forms. Surface realization of vowels varies depending on context (see Thompson and Thompson 1992).

c = [tʃ]
ç̣ = [ts]
ċ = [ts']
e = [æ, a, ə, ɛ, e]
ə̣ = [ʌ]
s = [ʃ]
ṣ = [s]
x̣ = [χ]
y = [i, j]

Abbreviations in the glosses are based on Thompson and Thompson 1992, 1996, Kroeber 1997:

'-' = affix
'=' = clitic
COMP = complementizer
CNCL = conjunctive subject clitic
DEM = demonstrative
DET = determiner
EMPH = emphatic (independent pronoun)
EVID = evidential
FUT = future
IMPF = imperfective
INCL = indicative subject clitic
INTRANS = intransitive
LINK = predicate modification
LOC = locative
NOM = nominalizer
O, OBJ = object
PL = plural
POSS = possessive (affix)
POCL = possessive subject clitic
S, SUBJ = subject
SG = singular
SUBJ.GAP = subject gap suffix
TRANS, TR = transitivizer

Opacity is a Matter of Representation: Shimakonde Vowel Harmony and Vowel Reduction

Laura J. Downing
Zentrum für Allgemeine Sprachwissenschaft, Berlin

As work like McCarthy (2002: 128) notes, pre-Optimality Theory (OT) phonology was primarily concerned with representations and theories of subsegmental structure. In contrast, the role of representations and choice of structural models has received little attention in OT. Some central representational issues of the pre-OT era have, in fact, become moot in OT (McCarthy 2002: 128). Further, as work like Baković (2007) notes, even for assimilatory processes where representation played a central role in the pre-OT era, constraint interaction now carries the main explanatory burden. Indeed, relatively few studies in OT (e.g., Rose 2000; Hargus & Beavert 2006; Huffmann 2005, 2007; Morén 2006) have argued for the importance of phonological representations. This paper intends to contribute to this work by reanalyzing a set of processes related to vowel harmony in Shimakonde, a Bantu language spoken in Mozambique and Tanzania. These processes are of particular interest, as Liphola's (2001) study argues that they are derivationally opaque and so not amenable to an OT analysis. I show that the opacity disappears given the proper choice of representations for vowel features and a metrical harmony domain.

1 Introduction

Phonological generalizations lead to derivational opacity, as defined by McCarthy (1999), when they are either not surface true or not surface apparent. A process is not surface-true when it underapplies: its context is met, yet the process fails to apply. A process is not surface-apparent when it overapplies: its context is not met, yet the process applies. (See Baković 2007a for a recent survey of opacity types.) Constraint ranking and interaction account for many cases of opacity which fit this definition. However, there are also problematic cases which seem to require reference to an intermediate level of representation besides the input and output representations allowed in standard Optimality Theory (OT). In the OT literature, one typical response to these cases of opacity

ZAS Papers in Linguistics 52, 2010: 159 – 191

is to admit limited serialism into OT, allowing for a constrained set of intermediate representations (see, e.g., Ettlinger 2008; Kiparsky 2000; Rubach 2003; Bermudez-Otero, forthcoming; and papers in Vaux & Nevins (2008) for detailed recent discussion of arguments in favor of serialism in phonology). Another is to propose new types of Correspondence which formalize intermediate representations in a way said to be compatible with non-derivational OT: e.g., sympathy theory, comparative markedness or turbid constraints (McCarthy 1999, 2003; Ito & Mester 1997, 2003a; Goldrick 2000).

Another response is to show that opacity is a feature of a particular analysis of a particular problem; it is not inherent to a data set. Adopting a different, equally well-motivated set of generalizations or representations can 'clarify' processes characterized as opaque by other OT practitioners. Examples of this approach include Harris's (1997) and Downing's (2006) reanalyses of German "Spitznamen" (opaque in Ito & Mester 1997), Downing's (2005) reanalysis of pre-NC compensatory lengthening in Bantu languages (opaque in Goldrick 2000), Green's (2007) reanalysis of Tiberian Hebrew epenthesis (opaque in McCarthy 1999), Krämer's (2008) reanalysis of English *r*-loss or intrusion (opaque in McCarthy 1993 and Orgun 2001), Mielke et al.'s (2003) reanalysis of Sea Dayak (opaque in McCarthy 2003). (See van der Hulst & Ritter (1999b) for other examples.)

In this paper, I adopt the second approach to reanalyze a set of processes related to vowel harmony in Shimakonde that have been argued to be opaque by Goldrick (2000), Ettlinger (2008) and Liphola (2001). I show that opacity vanishes under an alternative theory of representation. The goal of the paper is to show, echoing Uffmann (2007), that the choice of representations is as crucial to an OT analysis as the choice of constraints and constraint interactions.

The paper is organized as follows. In section 2, I present the data to be analyzed, namely, Shimakonde vowel height harmony (VHH) and the processes of vowel coalescence and vowel reduction that can make the output of harmony opaque. In section 3, I develop an analysis of VHH, crucially adopting the element theory of vowel representations, and a markedness licensing approach to harmony. In section 4, I extend this analysis to account for vowel reduction and coalescence, and show that it eliminates the apparent opacity in the interaction of these processes with VHH.

2 Vowel height harmony (VHH), vowel reduction and coalescence

Shimakonde is a Bantu language (P20) spoken in Mozambique and Tanzania. The source of the data and generalizations discussed is Liphola's (2001) study of the Mozambican dialect, his native language. (Kraal (2005) discusses similar patterns from a Tanzanian dialect.) Shimakonde has what Hyman (1999: 238) calls ‘canonical Bantu Vowel Height Harmony’ (VHH), with the following characteristic properties. First, it is subject to morphological conditioning. It is Root-controlled and motivates alternations in suffixes following the verb root, except that it does not apply to the final vowel morpheme (FV). It also does not apply to prefixes. The bolded domain in (1) summarizes these generalizations:[1]

(1) [[Prefixes] [[**Root+Derivational Suffixes**] FV]

Canonical VHH is also subject to phonological conditioning. As shown in (2c, d), below, only mid Root vowels [e, o] trigger harmony and are the output of harmony. Peripheral vowels – [i, u, a] – do not trigger harmony. As shown in (2a,b, e), they are all followed by high vowels. The low vowel [a] is also opaque. It does not undergo harmony and blocks the spread of harmony. Finally, non-initial back vowels often harmonize only to the back mid vowel, [o], as shown in (2f) vs (2g). Note in (2), that the Shimakonde VHH patterns are essentially identical to those discussed in Beckman’s (1997) OT analysis of Shona VHH:

(2)	Shona (Beckman 1997: 1)			Shimakonde (Liphola 2001: 147)		
(a)	-bvis-a	‘remove’	-bvis-ik-a	-píínd-a	‘bend’	-pind-íík-a
(b)	-bvum-a	‘agree’	-bvum-is-a	-púút-a	‘wash’	-put-ííl-a
(c)	**-tond-a**	**‘face’**	**-tond-es-a**	**-tóót-a**	**‘sew’**	**-tot-éék-a**
(d)	**-per-a**	**‘end’**	**-per-er-a**	**-péét-a**	**‘sift’**	**-pet-éél-a**
(e)	-shamb-a	‘wash’	-shamb-is-a	-páát-a	‘get’	-pat-ííl-a
(f)		**‘jump’**	**-svetuk-a**	**-tééng-a**	**‘set a fire’**	**-téng-úúl-a**
(g)		**‘uproot’**	**-gobor-a**		**‘cough’**	**-kólómóól-a**

As Liphola (2001) shows, derivational opacity arises when vowel coalescence and vowel reduction interact with VHH. In Shimakonde, as in many Bantu languages (see e.g. Casali 1998; Downing 1995), vowel hiatus across the prefix+[Root boundary is resolved by coalescence, accompanied by lengthening: a+[i,e > ee ; a+[u,o > oo. Since the output of coalescence in this context is a Root-initial mid vowel, we would expect vowels following it in the domain to harmonize to mid. However, only some coalesced mid vowels (3a, b) are followed by mid vowels, while others (3c, d) are followed by [+high] vowels.

[1] The domain of VHH bolded in (1) is variously termed the verbal base (Harris 1987; Harris & Lindsey 1995), the prosodic trough (Hyman 1998), or the prosodic stem (Downing 1999; Mutaka 1994). It is beyond the scope of this paper to discuss the morphological domain of VHH. See the work cited for further discussion of phonological processes like VHH which motivate the domain.

(3) Shimakonde vowel coalescence and vowel harmony (Liphola 2001: 101, 178)

		Input	Coalesced	Gloss
(a)	/a+[e/	/vandá-[ep-íl-a/	va-ndeépeéla	'they will harvest for'
(b)	/a+[o/	/vá-[óloota/	vóóloota	'when pointing'
(c)	/a+[i/	/vandá-[itík-a/	va-ndeétiíka	'they will respond'
(d)	/a+[u/	/vandá-[ukúl-a/	va-ndoókuúla	'they will dig'

That is, as the result of coalescence VHH is not always surface true. The context for harmony (i.e., a Root-initial mid vowel) occurs in the output of (3c,d), yet VHH underapplies in these examples. The input contrast between Root-initial mid vowels and Root-initial high vowels which is lost in the output due to coalescence is indirectly maintained through the underapplication of VHH.

The lengthening accompanying coalescence has also been argued to illustrate opacity, due to what Goldrick (2000) terms overapplication of mora projection. If you assume that vowels are not associated with moras in the input, the long vowel that results from coalescence in (3) is opaque. As only one vocalic root node survives coalescence, there is no motivation in the output for projecting two moras, one for each input vowel:

(4) Opaque compensatory lengthening (Goldrick 2000, fig. (2))

	μ μ	μ μ	μ μ	
	| |	ǂ |	⋱ |	
/ V_1 V_2 C/ →	V_1 V_2 C →	V_1 V_2 C →	V_1 V_2 C →	[V_2: C]
Input	by Project-μ	Hiatus Resolution	Re-association	Output

As Liphola (2001) shows, another source of apparent opacity in Shimakonde VHH comes from a process of unstressed vowel reduction. As we see in (5), the penult vowel is systematically stressed (realized as lengthening). Beside the canonical, unreduced VHH pattern, we find a reduced pattern. Pretonic mid vowels optionally neutralize to the low vowel [a]; pretonic peripheral vowels 'i, u, a' do not alternate.

(5) Shimakonde (Liphola 2001: section 5.2)

			Unreduced	Reduced
(a)	-píínd-a	'bend'	-pind-íik-a	*-pand-íik-a

(b)	-púút-a	'wash'	-put-ííl-a		*-pat-ííl-a
(c)	**-tóót-a**	**'sew'**	**-tot-éék-a**	**OR**	**-tat-éék-a**
(d)	**-péét-a**	**'sift'**	**-pet-éél-a**	**OR**	**-pat-éél-a**
(e)	-páát-a	'get'	-pat-ííl-a		

Stressed mid vowels retain the input harmonic quality contributed by the Root-initial syllable, not the [+high] value expected following peripheral [a]. That is, as a result of vowel reduction the context of VHH is not always surface apparent. VHH overapplies in the reduced forms in (5c, d): the context for VHH is not met, yet the process applies to the stressed vowels. (They harmonize with the input mid quality of the Root initial syllable which is lost in the output due to vowel reduction.)

Liphola (2001: 191-194) argues that vowel reduction must be considered a neutralizing process in Shimakonde. As shown in the table in (6), which summarizes a phonetic analysis that Liphola carried out based on his own speech and that of two other native speakers, reduced [a] is not distinct from underlying [a] in either realization or perception. There is no significant difference in the realization of the long [aa]'s, whatever their source and whether they are stressed or unstressed. There is also no significant difference among the short [a]'s, whatever their source:

(6)

Type of vowel	Avg F1	Avg F2
Stressed [aa]	**931**	**1530**
[aa] < aa	926	1528
[aa] < ee	916	1538
[aa] < oo	911	1552
[a] < a	741	1558
[a] < e	761	1542
[a] < o	814	1575

There is, however, a significant difference in the F1 of the long vs. short [a]'s which Liphola (2001) attributes to duration-induced *undershoot*. Perceptual evidence also shows that reduced [a] neutralizes with phonemic [a]. In a perception experiment of neutralized minimal pairs, Liphola (2001) found that native listeners could only correctly distinguish phonemic and reduced [a]'s 31.86% of the time, below chance (= 33.3%). Following Liphola (2001) and

Crosswhite (2001, 2003), we can conclude, then, that reduction in Makonde involves phonemic neutralization of all non-low vowels to [a] accompanied by duration-induced phonetic undershoot of short [a]. The phonological analysis in section 4, below, accounts only for the phonemic neutralization.

Reduction also interacts with coalescence. Since the output of coalescence is a mid vowel, we would expect all coalesced vowels to undergo pre-tonic mid vowel reduction like other unstressed mid vowels do. However, Liphola (2001) shows that only some coalesced mid vowels (7a,b) reduce to [a], while others do not (7c, d):

(7) Shimakonde vowel coalescence, harmony and reduction (Liphola 2001: 101, 178)

		Input	Coalesced	Reduced	Gloss
(a)	/a+[e/	/vandá-[ep-íl-a/	va-ndeépeéla~	va-ndaápeéla	‘they will harvest for’
(b)	/a+[o/	/vá-[óloota/	vóóloota ~	vááloota	‘when pointing’
(c)	/a+[i/	/vandá-[itík-a/	va-ndeétiíka ~	*va-ndaátiíka	‘they will respond’
(d)	/a+[u/	/vandá-[ukúl-a/	va-ndoókuúla ~	*va-ndaákuúla	‘they will dig’

Coalescence makes vowel reduction derivationally opaque. The context for reduction (i.e., a pre-tonic mid vowel) is met in the output of (7c,d), yet reduction underapplies in examples like these. The output contrast between pre-tonic mid vowels that do not reduce to [a] (7c,d) and pre-tonic mid vowels that do reduce (7a,b) makes reduction not surface true.[2]

In section 4, I show that if we follow some other previous studies of Bantu VHH, like Goldsmith (1985), Harris (1987), and Steriade (1995), and adopt the element theory of vowel representations, all of these sources of opacity vanish. In the next section, I first motivate element theory by showing how it accounts for the basic VHH patterns illustrated in (2).

[2] As Barnes (2006) and Nevins (2007) note, the fact that some non-stressed long mid vowels resulting from coalescence can undergo vowel reduction confirms that reduction is not a purely phonetic process of duration-induced undershoot. Only long stressed vowels resist reduction, showing the process has been phonologized.

3 Element theory analysis of VHH in Shimakonde

In this section, I develop an OT analysis of the central properties of Shimakonde VHH: that it is Root-controlled, involves only mid vowels, and the low vowel is inert and opaque.[3] As we shall see, these properties fall out from the element theory of representation adopted, and from the related hypothesis that harmony involves positional licensing of markedness asymmetries. To highlight the advantages of the approach adopted here, I begin by presenting an alternative, Beckman's (1997) standard OT analysis of VHH in Shona.

3.1 Positional faithfulness account of Bantu VHH: Beckman (1997)

In Beckman's (1997) analysis of VHH in Shona – which basically translates into OT terms underspecified, autosegmental approaches like those reviewed in van der Hulst & van de Weijer (1995) – the motivation for harmony is gestural uniformity. VHH minimizes the number of different vowel height specifications in the Stem, as some are shared. The direction of harmony falls out from the Root>>Affix asymmetry defined by the ranked POSITIONAL FAITHFULNESS constraints in (8a, b). Only mid vowels are involved in VHH because of the HEIGHT MARKEDNESS HIERARCHY in (8c):

(8) *Faithfulness constraints defining Root-controlled harmony* (Beckman 1997: 14)
(a) IDENT-S1(hi)
A segment in the root-initial syllable in the output and its correspondent in the input must have identical values for the feature [high].
(b) IDENT(hi)
Correspondent segments in the output and input must have identical values for the feature [high].
(c) HEIGHT MARKEDNESS : DOMINANCE HIERARCHY: *MID >> *HIGH, *LOW

This hierarchy defines Mid vowels as the most marked, and penalizes having more than one height specification within the stem. Every height specification incurs a violation. The tableaux in (9) and (10) show how these constraints

[3] The analysis of VHH developed in this section will not address the front-back asymmetry illustrated in (2f) vs. (2g), as my account is not essentially different from that of Beckman (1997). Indeed, as Kaun (1995, 2004) shows, this asymmetry is common in [round] vowel harmony, cross-linguistically.

account for harmony in *–petela* 'sift for' vs. *–pindika* 'be bendable' (see (2a) and (2d), above):[4]

(9) (Beckman 1997: fig (38), using Shimakonde stem)

/-pet-il-a/	IDENT-S1(hi)	*MID	*HIGH	IDENT(hi)
a.-Ce C iC - | | Mid Hi		*	* !	
☞b. CeC e C - \ / Mid		*		*
c Ci C i C - | / High	*!		*	*
d Ce C e C - | | Mid Mid		**!		*

In tableau (9), candidate (b) is optimal, as harmonizing to the Root-initial vowel through autosegmental feature sharing minimizes the number of distinct vowel height specifications in the harmonizing domain while maintaining the input height of the Root-initial vowel.

(10) ROB tableau (Beckman 1997: fig (30), using Shimakonde stem)

/-pind-ek-a/	IDENT-S1(hi)	*MID	*HIGH	IDENT(hi)
a.-C i C i C - | | Hi Hi			**!	*
☞b. Ci C i C - \ / Hi			*	*
c.C eC e C - | / Mid	*!	*		*
d.-Ci C e C - | | Hi Mid		*!	*	

4 In the tableaux, Mid is an abbreviation for the Aperture features [-high], [+low]. Beckman (1997) assumes standard SPE-style vowel height specifications, grouped into an Aperture node.

The tableau in (10) illustrates that Root-initial high vowels optimally trigger the same kind of autosegmental spreading harmony, even if we assume, following Richness of the Base (ROB) that the affixal vowel is Mid in the input. Candidate (b) is again optimal, as the Root-initial vowel height feature is optimally shared by all vowels in the harmonizing domain.

To account for the inertness and opacity of [a], Beckman (1997) stipulates that the Faithfulness constraint, IDENT(low), is high ranked:

(11) IDENT(low):
The input and output of a vowel must have the same [low] specification.

This constraint must be considered a stipulation, as otherwise [high] is the active feature in VHH. As shown in (12) and (13), ranking this constraint above the height markedness constraints blocks [a] from being an optimal trigger or target for harmony:

(12) Beckman (1997: fig (34), with Shimakonde stem in (2e))

/-pat-el-a/	IDENT-S1(hi)	IDENT(low)	*MID	*HIGH	IDENT(hi)

☞a.-Ca C i C -
      | |
     Lo Hi
				*	*

b. CaC a C
      \ /
     Lo
		*!			*

c.-Ca C e C -
     | |
    Lo Mid
			*!		*

Candidate (12a) is optimal, as it does not violate any of the highest ranked constraints. The role of IDENT(low) in optimizing the inertness of [a] is highlighted by candidate (b), which shows that harmonizing to agree with a low Root-initial vowel violates this constraint. Candidate (c) is non optimal as the vowels in the harmonizing domain are not identical in height.

Similarly, in tableau (13), IDENT(low) optimizes candidate (a), in which [a] does not undergo harmony to the Root-initial vowel and also blocks harmony to subsequent vowels in the harmonizing domain.

(13) Beckman (1997: fig (35), with Shimakonde stem)

/-lekan-il-a/	IDENT-S1(hi)	IDENT (low)	*MID	*HIGH	IDENT(hi)
☞a..-Ce C a C i C - | | | Mid Lo Hi			*	*	
b. Ce Ce C e C- \ \ / Mid		*!	*		*
c.-C e C a C e C - | | | Mid Lo Mid			**!		*

As we can see from these tableaux, if we took IDENT(low) out of the ranking, total assimilation of the suffix vowel(s) would be optimal, as in (12b), (13b).

To sum up Beckman's (1997) analysis, a high-ranked Positional Faithfulness constraint (IDENTS1) accounts for the Root-controlled property of VHH. The motivation for harmony is to minimize the number of different vowel height specifications in the Stem (this minimizes violations of the HEIGHT MARKEDNESS : DOMINANCE HIERARCHY (8c)). Only Mid vowels are involved in VHH because these vowels are most marked. Low vowels are inert and opaque due to a special high-ranked Faithfulness constraint, IDENT(low).

While the analysis works very well for these basic patterns, there are conceptual reasons to be dissatisfied with it. First, it provides no explanation for why Mid vowels are marked: this is just a stipulation of the HEIGHT MARKEDNESS : DOMINANCE HIERARCHY (8c). Neither the inertness nor the opacity of [a] receive an explanation. This is just stipulated by the ranking of IDENT(low).[5] Crosswhite (2003) shows that, even though Shimakonde has the same basic VHH pattern as Shona, Beckman's (1997) analysis cannot be extended to account for the way that vowel reduction interacts with VHH in Shimakonde. Notice in (5) that the vowel reduction data obviously contradict Beckman's proposal that Root-driven harmony is motivated by Positional Faithfulness. The Root-initial syllable undergoes neutralization – in violation of IDENTS1, yet the penult suffixal vowel retains the Root-controlled harmonic quality. Further, no factorial typology based on Beckman's analysis of the unreduced VHH pattern can account for the choice of reduced vowel. For [a] to be the optimal reduced vowel, some constraint defining [a] as unmarked must

[5] These problems for Beckman's (1997) analysis are partly inherent to the SPE features that she adopts. See Goldsmith 1985, Harris 1994, 1997, Harris & Lindsey 1995, 2000, Hyman 1999 and Steriade 1995 for detailed discussion.

outrank IDENT(low) to account for why non-alternating Root [a] remains inert and opaque. As Crosswhite (2003) points out, simply adding the high-ranked licensing constraint in (14) to the analysis does not automatically optimize [a] as the reduced vowel because [a] is not the unmarked vowel in Beckman's analysis; [+high] vowels are.

(14) LIC/STRESS: A mid vowel is licensed if it is associated with a stressed syllable.

Other rankings cannot be changed. Recall that the relative ranking of *MID >> *HIGH is fixed. The ranking of IDENT(low) is also fixed, in order to account for the consistent inertness and opacity of [a]. This inflexibility of most of the constraint rankings is what makes it impossible to devise a factorial typology based on Beckman's analysis that can account for the reduced vowel harmony pattern. These points are made explicit in the tableau in (15), where the bomb indicates the candidate which is incorrectly chosen as optimal in the analysis:

(15)

/-pet-il-a/	LIC/ STRESS	IDENT-S1(hi)	IDENT (LOW)	*MID	*HIGH	IDENT(hi)
a.-Ce C iC - | | Mid Hi	*!			*	*	
b. CeC e C - \ / Mid	*!			*		*
💣c Ci C i C - | / High		*			*	*
d Ci C e C - | | High Mid		*		*!		*
👎e. CaC e C - | | Low Mid		*	*!	*		**

As we can see, simply adding high-ranked LIC/STRESS (14) to the analysis wrongly predicts reduction in both the Root-initial vowel and suffix vowel. Further, the optimal reduction target is a [+high] vowel due to the fixed ranking of the HEIGHT MARKEDNESS : DOMINANCE HIERARCHY (8c).

Crosswhite's (2001, 2003, 2004) theory of vowel reduction typologies also cannot explain why reduction results in [a] and not some other vowel. In Crosswhite's typology, there are two types of vowel reduction. In prominence reducing languages, reduction results in either [+high] vowels or schwa, the shortest vowels. This type of reduction is typically found in languages where unstressed vowels are 'extremely short.' In contrast-enhancing languages, reduction results in the peripheral vowels – [i, u, a]. Shimakonde reduction fits the contrast-enhancing pattern. The problem is that since all peripheral vowels are potential targets, there is no explanation for why [a] – the longest and most sonorous peripheral vowel – is the optimal target for reduction. As we shall see, one of the advantages of the analysis developed in the next sections is that it provides a ready explanation for why only [a] is the target for reduction in Shimakonde. Another is that it defines a ranking typology that accounts for the interaction of vowel reduction with VHH.

3.2 Element theory and positional licensing of markedness drive VHH

Both the reduced and unreduced Shimakonde vowel harmony patterns can be straightforwardly related in a non-derivational OT analysis, if we follow a different way of thinking about Bantu VHH argued for in work like: Goldsmith (1985), Harris (1987, 1994, 1997, 2005), Harris and Lindsey (1995), Hyman (1998, 1999), and Steriade (1995). What these researchers observe is that Root-initial syllables contain the full set of vowel contrasts. Harmony represents a neutralization of contrasts, as it makes the quality of suffix vowels predictable from the Root. Further, non-peripheral vowels occur in suffixes only if a non-peripheral vowel also occurs in Root-initial position.

The element theory of vowel representation (see, e.g., Anderson & Ewen 1987; Goldsmith 1985; Harris 1990, 1994, 1996, 1997; Harris & Lindsey 1995, 2000; van de Weijer 1994) and theories of complexity licensing which they implement (see, too, van der Hulst & Ritter 1999a, b) makes the markedness asymmetry between peripheral and non-peripheral vowels which drives Bantu vowel harmony explicit. The essential proposals of these works which are crucial to this analysis are summarized in (16):

(16) Element theory and licensing

Elements

(a) Peripheral vowels are simplex elements: A, I, U

(b) Mid vowels are complex, consisting of a head (underlined) and a dependent:
e = [A, I]
o = [A, U]

Licensing

(c) Simplex segments with elements on only one tier, Head or Dependent, have no special constraints on their distribution, but complex segments, with elements on both the Head and the Dependent tier, often require special licensing:

(ci) Either they are directly licensed, by occurring in a strong position, like Root-initial position or a stressed syllable;

(cii) Or they are indirectly licensed, for example, by being linked to a complex vowel in a strong position.

That is, in this theory, peripheral vowels are less marked because they are simplex, while the Mid vowels are complex. Mid vowels are the only ones involved in harmony, because the dependent element of complex vowels requires licensing, and can be licensed by spread. [A] is inert for the same reason that [I, U] are inert: it is simplex and only complex vowels need to be licensed. [A] is opaque because it is simplex: it cannot license adjacent complex vowels.

These generalizations about VHH are formalized by the constraints below:

Faithfulness Constraint:

(17) FAITH S1-[A] (adapted, Beckman 1997): Dependent input and output vocalic elements of the Root-initial syllable must be identical.

Licensing Constraint motivating VHH:[6]

(18) LIC/SPREAD (Harris 1994, 1997; Walker 2005): A dependent vocalic element is licensed if it is associated (by multiple-linking) to every vowel in the relevant prosodic domain (e.g., harmonizing domain).

6 Precedents for this constraint include Hyman's (1998) PLATEAU constraint and Goldsmith's (1985) suggestion that complex vowels are unstable, and can be "propped up" by being linked to other complex vowels. And work like Steriade (1994, 1995), Kaun (1995, 2004) and Walker (2005) argues there is a perceptual reason why marked vowels are optimally linked together: marked vowels are harder to perceive and having a longer duration – by occurring in a sequence of equally marked vowels – enhances their perceptibility.

The tableau in (19) exemplifies the analysis with a stem with a Root-initial Mid vowel; cf. (9), above:

(19)

/-pet-il-a/	FAITH S1-[A]	LIC/SPREAD
a. A | – Ce C i:C – | | I I		*!
☞b. A | ╲ – Ce C e:C – | | I I		
c. – C i C i: C – | / I	*!	

Candidate (b) is optimal, as it satisfies FAITHS1 and LIC/SPREAD, which require the dependent element of the Root-initial Mid vowel to be multiply linked. Candidate (a) is not optimal, as the dependent element is not licensed by multiple linking, while candidate (c) is not optimal, as the Root-initial vowel has not retained its input elements.

Tableau (20) illustrates that the analysis also correctly accounts for the lack of harmonic spreading found with [high] vowels:

(20) ROB tableau

/-pind-ek-a/	FAITH S1-[A]	LIC/SPREAD
a. A | – C i C e: C – | | $\underline{I}$ $\underline{I}$		*!
☞b. – Ci C i:C – | | $\underline{I}$ $\underline{I}$		
c. A / | – Ce C e:C – | | $\underline{I}$ $\underline{I}$	*!	

Candidate (b) is optimal, as simplex vowels do not need to be licensed. Candidates (a) and (c) each violate one of the two highest ranked constraints.

The analysis of the inertness and opacity of [a] requires two new constraints:

(21) OCP: $*[\underline{X}_i, X_i]$: Complex vowels with identical Head and Dependent element are marked. (adapted, Harris 1994, 1997)

(22) FAITH-HEAD: A vowel must be associated with the same Head element(s) in the input and output.

(23) Inertness of [a]

/-pat-il-a/	FAITH S1-[A]	LIC/SPREAD	FAITH-HEAD
☞a. – C a C i:C – | | A I			
b. – C a C a: C – \ / A			*!*
c. – C a C e: C – | / | A I			*!

Candidate (a), with no harmony or assimilation, is optimal, as it does not violate any constraints. Candidates (b) and (c), which both illustrate vowel harmony, are non-optimal as they violate FAITH-HEAD.

The next tableau shows how the analysis also accounts for the blocking effect of [a]:

(24) Blocking by [a]

/-lekan-il-a/	FAITH S1-[A]	OCP	LIC / SPREAD	FAITH-HEAD
☞a.- A | C e C a C i:C – | | | I A I			**	
b. A / \\ CeC aCe:C – | | | I A I		*!		
c. A | C e C a C e:C – | | / | I A I			**	*!

Candidate (a) with no harmony is optimal, as the only complex vowel in this output is licensed by association to Root-initial position. Candidate (b) crucially shows the role of the OCP constraint (21) in optimizing opaque [a] in VHH, while candidate (c) fatally violates FAITH-HEAD.

To sum up this section, we have seen that both Beckman's (1997) analysis and the element analysis of Bantu VHH developed here are equally successful in accounting for the basic harmony patterns. In the next section, I show that the element analysis can easily be extended to account for the interaction of vowel reduction and coalescence with VHH. Recall that Crosswhite (2003) shows this is impossible in Beckman's (1997) approach. As we will see, the interactions of both vowel reduction and vowel coalescence with VHH are not opaque in this analysis, a further advantage of the licensed elements approach.

4 Extending the analysis to clarify reduction, coalescence, and their interaction with VHH

4.1 Reduction

As shown in (25), in the Element theory of vowel representation, what unifies the unreduced and reduced harmony patterns in Shimakonde is that a dependent [A] must be associated with every vowel in the harmonic domain (in square brackets) in both patterns. Only Mid vowels are involved in both harmony and reduction, as they have a dependent [A] which requires licensing. What distinguishes the patterns is that the stressed syllable – the lengthened penult – is the primary licensor of complex vowels in the reduced pattern, not the Root-initial syllable:[7]

```
(25)          I̲     I̲                            I̲
(a)      [– p e t – eel ]– a        (b)   [– p a t – eel ] – a
              |/                               |/
              A                                A
```

In (25) we see reduction in a disyllabic stem, where unreduced (25a) and reduced (25b) are the only two variants. However, as Liphola (2001: 170) shows, more variation is possible in longer stems:

[7] As an aside, we might speculate that the reduction pattern developed when speakers at some point decided that the stressed penult was not just an 'accidental' end point for the domain of VHH, morphologically-defined, but the licensor of a harmonizing quality that is contributed, in the input, by a vowel at the opposite edge of the domain.

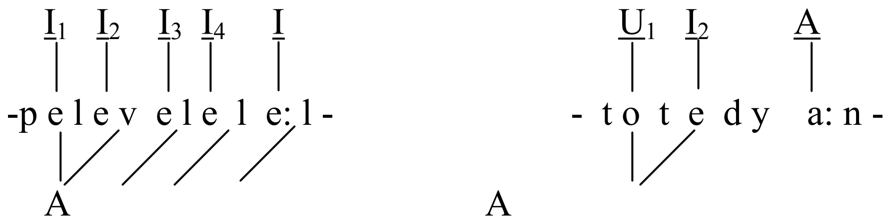
(26)

(a) *kú-pélévéléléela* ‘to not reach a full size for’
~ *kú-pálévéléléela*
~ *kú-pálávéléléela*
~ *kú-páláváléléela*
~ *kú-pálávúláléela*
~ **kú-pélévéláléela* (etc.).

(b) *kú-tót-édy-aana* ‘to cause each other to sew’
~ *kú-tátédyaana*
~ *kú-tátádyaana*
~ * *kú-tótádyaana*

Liphola (2001) observes that the generalization describing possible reduction patterns is that reduction applies contiguously from the Root-initial vowel towards the stressed vowel. In the analysis developed here, one can recast this generalization in terms of licensing of complex vowels. The element representations of the unreduced forms in (26), is given below:

(27) (a) (b)



Comparing these structures with the reduced variants, one observes a progressive strengthening of the capacity to license Heads and complex vowels as one moves towards the stressed vowel from the Root-initial vowel.

Similar patterns of stressed vowels attracting or licensing a harmonizing feature have been described for Spanish and Italian metaphony dialects (Hualde 1989; Walker 2005; Zubizaretta 1979). I adopt here the left-branching binary metrical representation of stress licensing motivated in Zubizaretta’s (1979) analysis of Andalusian Spanish vowel harmony. As we can see in (28), below, this theory of metrical representations allows the metrical motivation for contiguity of reduction to be very clearly formalized. The Root-initial syllable is metrically weakest, and metrical strength increases contiguously from the Root-initial through the stressed vowel. The progressive increase in metrical strength mirrors the contiguous potential realization of reduced vowels:

(28)

```
(a)  - [p e [l   e [v   e [l  e  l    e: ]1 ]2 ]3 ]4 -
                             [w       s]1
                        [w            s1]2
                 [w                   s2]3
         [w                           s3]4

(b)  - [ t  o  [t e  d y  a: ]1 ] 2 -
                [w        s]1
           [w             s1]2
```

The constraints in (29) formalize the proposal that the relative ability of a position to licence a complex (Mid) vowel or Head element depends on its relative metrical strength (as represented in (28)). STRESS/LICENSING A (29a) optimizes reduction by penalizing an output in which the stressed vowel is not at least as complex as the unstressed vowels in the domain. One could reduce a complex vowel to a simplex vowel by deleting either the Head element or the Dependent element. As we see in (25), above, vowel reduction in Shimakonde involves deleting a Head element, maintaining the multiply-linked Dependent element of complex Mid vowels. STRESS/LICENSING B (29b) accounts for the contiguity of reduction: it is violated if a vowel with a Head element is followed by a (metrically stronger) vowel lacking a Head element:

(29) *Licensing Constraints motivating reduction*

(a) STRESS/LICENSING A (adapted, Harris 1990, 1996, 1997, 2005):
Within the licensing domain, a vowel in the metrically strongest position must not be elementally less complex than vowels in metrically weaker positions.

(b) STRESS/LICENSING B:
Within the licensing domain, a Head element for a vowel in a metrically weak position is licensed by a Head element in the adjacent metrically stronger position.

To complete the analysis, we must account for the variability we find in reduction: any contiguous string of vowels from the Root to the stressed vowel can be reduced. The constraint conjunction in (30) optimizes this variation.[8]

---

8 See work like Downing (1998, 2000), Itô & Mester (2003b), Odden (2006), along with references cited in these works, for detailed motivation and exemplification of constraint conjunction in OT.

(30) FAITHHEAD (22) ∪ STRESS/LICENSING A
This constraint conjunction is satisfied if FAITHHEAD is satisfied (and then violations of STRESS/LICENSING A do not count) OR if STRESS/LICENSING A (and then violations of FAITHHEAD do not count). The conjunction is violated if the constraint in each half incurs violations.

FAITHHEAD (22) is involved in the conjunction, because reduction necessarily violates this constraint. The conjunction in (30) must be ranked below LIC/SPREAD (18), as this constraint is the one which optimizes retaining the dependent vowel element under reduction. (The dependent element is the one licensed by multiple linking.) The conjunction makes it optimal for variants to violate FAITHHEAD, if that leads to more optimal satisfaction of STRESS/LICENSING A. That is, reduction is optimal if the output contains less complex vowels in the metrically weak positions compared to the strongest position. Impossible variants are 'the worst of the worst': candidates which are non-optimal both for FAITHHEAD *AND* STRESS/LICENSING A. These would be candidate outputs where metrically strong positions have undergone reduction while metrically weak positions have not. STRESS/LICENSING B is ranked above the conjunction, as it is not violated in optimal candidates. That is, acceptable reduction variants must satisfy STRESS/LICENSING B by having Head elements contiguously cluster at the stressed right edges of the harmonic domain. (And, of course, they must also optimally satisfy higher ranked constraints.) These constraint rankings are summarized below:

(31) Constraint ranking accounting for vowel reduction
FAITH S1-[A], LIC/SPREAD >> STRESS/LICENSING B >> FAITHHEAD ∪ STRESS/LICENSING A

The tableau in (32) exemplifies the analysis. Candidates (b) - (d) are optimal variants because they satisfy STRESS/LICENSING B as well as the more highly ranked constraints. Candidate (b) satisfies FAITHHEAD, as no reduction occurs in it. Candidates (c) and (d) show reduction, so necessarily violate FAITHHEAD. However, these reduction patterns – with the reduced vowels contiguously clustered at the opposite edge of the domain from the stressed vowels – optimally satisfy STRESS/LICENSING A, B. Candidate (e), with reduction in the vowel next to the stressed vowel only, is a non-optimal variant, as it incurs violations of STRESS/LICENSING B: vowels in relatively weak positions are complex (and have Heads) while a relatively strong position has a reduced vowel. Candidates (a) and (f) are non-optimal because failing to license the dependent element through multiple linking violates the high-ranked constraints optimizing VHH.

(32)

/-pelIvIl-Il-a/	FAITH S1-[A]	LIC/ SPREAD	STRESS/ LIC B	**FAITH -HEAD**	∪	**STRESS/ LIC A**
a. A | - C e C i C i C i: C- | | | | I I I I		*!**				*
☞b. A | \ \ \ - C e C e C e C e: C- | | | | I I I I						
☞c. A | \ \ \ - C a C e C e C e: C- | | | I I I				*		
☞d. A | \ \ \ - C a C a C e C e: C- | | I I				**		
e. A | \ \ \ - C e C e C a C e: C- | | | I I I			*!	*		
f. - C i C i C i C i: C- | | | | I I I I	*!					

In sum, tableau (32) shows that the interaction of reduction with VHH need not be opaque. The analysis correctly chooses [a] as the optimal target for reduction and accounts for the attested range of variability in reduction in a non-derivational way. Recall that all of these were problems for previous analyses.

The tableau in (33) shows that the same analysis straightforwardly extends to account for the reduction pattern in (26b), data that Liphola (2001) and Ettlinger (2008) argue are problematic for an OT analysis. (Highest ranked FAITHS1-[A] is omitted from this tableau to improve readability, as it is never violated by the optimal output variants.) Candidates (b) - (d) are optimal variants because they satisfy STRESS/LICENSING B and also best satisfy the higher ranked constraints. Candidate (b) satisfies FAITHHEAD, as no reduction occurs in it. Candidates (c) and (d) show reduction, so necessarily violate FAITHHEAD. However, these reduction patterns – with Head elements contiguously clustered at the same edge of the domain as the stressed vowel – optimally satisfy STRESS/LICENSING B. Candidate (e), with reduction in the vowel next to the stressed vowel only, is a non-optimal variant, as it violates STRESS/LICENSING B: the metrically weak Root-initial position has a complex vowel, while the relatively stronger position of the following vowel has undergone reduction. Candidate (a) is non-optimal, as failing to license the dependent element through multiple linking violates the high-ranked constraints optimizing VHH. (Recall that the OCP constraint accounts for the blocking effect of [a], optimizing failure to spread Dependent [A] to the final syllable in the harmony domain.)[9]

9 As Liphola (p.c.) points out, this analysis does not capture that VHH is obligatory while the reduced variants are optional. How best to formalize this kind of asymmetry between variants is a topic for future research.

(33) Opaque'a' in harmony, reduction

/-totIdy-an-a/	OCP	LIC/ SPREAD	STRESS/ LIC B	**FAITH -HEAD**	∪	**STRESS/ LIC A**
a. A | - C o C i C a:C- | | | U I A		**!				*
☞b. A |\\ - C o C e C a: C- | | | U I A		*				*
☞c. A |\\ - C a C a C a:C - | A		*		*		
☞d. A |\\ - C a C e C a:C- | | I A		*		*	*	*
e. A |\\ - C o C a C a:C- | | U A		*	*!	*	*	*

4.2 Coalescence and its interaction with VHH and reduction

As shown in section 2, the interaction between vowel reduction and VHH is not the only apparent source of derivational opacity in Shimakonde. Coalescence and its interaction with VHH and reduction, illustrated in figures (3), (4), and (7), above, also illustrate derivational opacity. Coalesced outputs contain opaque long vowels, and some also contain Mid vowels which do not participate in VHH or vowel reduction. In this section I show that both sources of opacity vanish, given an alternative set of representational assumptions.

4.2.1 Clarifying coalesced vowel length

Goldrick's (2000) claim that the lengthening which accompanies vowel coalescence is opaque crucially rests on the assumption – borrowed from McCarthy & Prince (1986) via Rosenthall (1994: 26) – that moras are redundantly underspecified for short vowels in the input. All short vowels are optimally associated with a single mora in the output to satisfy PROJECT-μ (or VOWEL-μ):

(34) VOWEL-μ (Rosenthall 1994: 26)
For every vocalic root node rt_i, there is a mora μ_i.

This analysis was illustrated in (4), above, repeated below for convenience:

(35) Opaque compensatory lengthening analysis (Goldrick 2000, fig. (2))

	μ μ | |	μ μ ǂ |	μ μ ⟍ |	
/ V_1 V_2 C/ →	V_1 V_2 C →	V_1 V_2 C →	V_1 V_2 C →	[V_2: C]
Input	by Project-μ	Hiatus Resolution	Re-association	Output

As Goldrick (2000) notes, however, *Richness of the Base* (Prince & Smolensky 2004) requires us to consider two possible inputs for non-contrastive output properties like short vowel-mora association: not only the representation in (35), repeated below for ease of comparison, but also the representation in (36), with short vowels fully specified for moras (assumed by, e.g., Broselow et al. 1997):

(36) Transparent compensatory lengthening analysis

μ μ | |	μ μ ǂ |	μ μ ⟍ |	
V_1 V_2 C →	V_1 V_2 C →	V_1 V_2 C →	[V_2: C]
input	Hiatus Resolution	Re-association	Output

As we can see in (36), fully specifying moras in the input makes lengthening transparent in an OT account, just as it is in earlier accounts, like Clements

(1986). Input 'timing' (CV slots or moras) of the V_1 V_2 sequence is preserved in the output (bimoraic) long V, even though the input quality of both output vowels is not preserved. The fully-specified input in (36) not only renders lengthening transparent, it also better satisfies the principle of *Lexicon Optimization* (Prince & Smolensky 2004, Kager 1999: 33) which proposes that inputs should always match outputs, barring evidence to the contrary. This principle defines redundant underspecification of inputs as generally non-optimal, since inserting the underspecified features in the output necessarily leads to FAITHFULNESS violations which are not incurred if the input and output are identical.

To sum up this section, lengthening under coalescence is fully transparent if the optimal, fully specified input representation of vowel–mora association is adopted. Opacity is a characteristic of the underspecification analysis, not inherent to the process of compensatory lengthening.

4.2.2 Coalescence and the representations of Mid vowels

The data in (3) and (6), above, illustrate that coalescence can make both VHH and vowel reduction derivationally opaque. The output of coalescence is a Mid vowel, yet some coalesced Mid vowels do not participate in VHH and reduction while others do. I propose that this difference falls out in the approach adopted here, as Mid vowels which have coalescence of [a+i/u] as their source have a different representation from other Mid vowels.

In Element theory, coalescence can be analyzed as the fusion of two heads (see, e.g., Anderson & Ewen 1987), to satisfy the high-ranked constraint in (37):

(37) COALESCE: *Onsetless syllables and *Diphthongs.

The output of coalescence maintains the input headedness of the vowels, as shown in (38):

(38) Representation of coalesced mid vowels (compare with (16b))
e = [$\underline{A}$, $\underline{I}$]
o = [$\underline{A}$, $\underline{U}$]

This output violates the markedness principle in (39a) on element combinations, requiring them to have one and only one head. To make coalescence, rather than, say, vowel deletion the optimal repair for violations of the constraints abbreviated in (37), the FAITH constraint in (39b) must outrank HEAD:

(39)
(a) HEAD: Phonological entities have one and only one head.
outranked by:
(b) MAX-VOC: All input vocalic elements must have an input correspondent.

As harmony crucially involves the licensing of complex Mid vowels – ones with Head and Dependent elements – the inertness of Mid vowels resulting from the coalescence of [a+i] or [a+u] falls out transparently from the representations in (38). These Mid vowels have no dependents requiring licensing. Mid vowels resulting from coalescence of [a+e] or [a+o], however, have dependent elements – their representation is identical to (16b) – so they still trigger VHH and undergo reduction.

The constraint rankings for coalescence are summarized in (40), and the tableaux in (41) and (42) exemplify the analysis. (HEAD (39b) is omitted from the tableaux, as it is too low-ranked to play a role in choosing optimal candidates):

(40) Coalescence, plus canonical vowel harmony:
COALESCE>> >> FAITH [A]-S1, LIC/SPREAD >> STRESS/LICENSING B >> FAITHHEAD ∪ STRESS/LICENSING A >> MAX-VOC >> HEAD

(41) a+i coalescence

/-va-nda-[im-il-a/	COA-LESCE	FAITH [A]-S1	LIC/ SPREAD	STRESS/ LIC B	**FAITH HEAD**	∪	**STRESS/ LIC A**	MAX-VOC
a. - nd a [i m - i:l - | | | A I I	*!							
☞b. - nd e[: m i: l - |╲ | A I I								
c. - C [i : m - i:l - | | I I					*			*!

Recall that vowels are defined as complex if they have an element on both the Head and the Dependent tier. For this reason, candidate (b) in tableau (41) is optimal. Note that it does not violate any constraints. Candidate (a), which matches the input, is non-optimal as it violates high-ranked COALESCE (37). Candidate (c), which resolves hiatus by deleting one of the input vowel elements, is non-optimal, as it violates MAX-VOC (39b). Note in (41) that the Mid vowel which is the output of coalescence is not associated with a Dependent vowel, so VHH is not expected.

In contrast, VHH is optimal if the Mid vowel which is the output of coalescence contains an input Mid vowel with a Dependent element to be licensed. This is shown in (42):

(42) a+e

/-va-nda-[ep-il-a/	COA-LESCE	FAITH S1-[A]	LIC/ SPREAD	STRESS/ LIC B	**FAITH HEAD**	∪	**STRESS/ LIC A**	MAX-VOC
a. A | – nd a [e p - i:l – | | | <u>A</u> <u>I</u> <u>I</u>	*!		*				*	
☞b. A \\ | \\ – nd e[: p e: l – | | <u>I</u> <u>I</u>					*			*
c. – nd e [: p e: l– | \\ / | <u>I</u> <u>A</u> <u>I</u>		*!			*			*

Candidate (b) is optimal, as it violates none of the high-ranked constraints. Note that VHH is optimal in this candidate, as the coalesced Mid vowel has a dependent element which requires licensing. Candidate (a), which is identical to the input, is not optimal because it violates high-ranked COALESCE. Candidate (c), where the dependent element of the Root-initial syllable has been deleted, is not optimal because it violates FAITHS1-[A] (17).

The tableau in (43) shows that the analysis also transparently accounts for the interaction of [a+i] coalescence with reduction harmony:

(43) a+i plus reduction

/-va-nda- [im-il-a/	COA- LESCE	FAITH S1- [A]	LIC/ SPREAD	STRESS/ LIC B	**FAITH HEAD**	∪	**STRESS/ LIC A**	MAX- VOC
a. - nd a [i m - i:l - | | | A I I	*!							
☞b. – nd e[: m i: l – |╲ | A I I					*			
c. - nd a: [m - i:l - | | A I					*			*!

Candidate (b) is optimal, as it violates none of the constraints. Candidate (a), which is identical to the input, is not optimal because it violates high-ranked COALESCE. Candidate (c), with reduction to [a], is non-optimal because this violates MAX-VOC (39b).

To sum up, we have seen that adopting the element theory of vowel representations, with its crucial distinction between Head and Dependent elements, makes the relation between reduced and unreduced vowel harmony patterns transparent. In all surface variants, a dependent [A] element is licensed by spread to all vowels in the domain (subject to the OCP). Heads must also be licensed: they must either be in the metrically strongest position or adjacent to a Head in a metrically stronger position. This explains the Contiguity of reduced vowels. Coalesced Mid vowels only participate in VHH and reduction when they have a Dependent element to be licensed by harmony. None of the processes is derivationally opaque.

5 Conclusion

I have shown that the choice of features and the relations among segments that define optimal combinations of features are crucial to an analysis of Shimakonde vowel harmony. Element theory explains why only non-peripheral vowels participate in harmony. It also explains why [a] is the target of reduction. And it allows non-high vowels, crucially, to have two different output representations – this is why the interactions of coalescence and reduction with VHH are not opaque. Low vowels can be represented either as Head [A] or, more marked, as Dependent [A] (in reduction). Mid vowels can be represented either as Head/Dependent complexes or, more marked, as the fusion of two Heads (in coalescence). Further, we have a choice of representations of unbounded stress. A binary branching model, we have seen, clearly formalizes the progressive weakening in the licensing strength of vowels which are farthest from the stressed vowel. It is this choice of metrical representation which allows us to formalize contiguity of reduction in a way that is not derivationally opaque.

In short, the analysis shows that choice of input and output representations is as important in OT as the choice of constraints evaluating the representations. Opacity can vanish with the proper choice of representations.

Acknowledgements

Earlier versions of this paper were presented at the Sound Circle, Amsterdam, the 14th Manchester Phonology Meeting, OCP 3 in Budapest, a Potsdam Phonology Colloquium and the Workshop on the Phonological Bases of Phonological Features, Tromsø. I would like to thank the audiences at these talks – in particular, Jill Beckman, Ricardo Bermúdez-Otero, Paul Boersma, Ben Hermans, Larry Hyman, Bruce Morén, David Odden, Marc van Oostendorp, Donca Steriade, Jochen Trommer and Jeroen van de Weijer – as well as Marc Ettlinger, Marcelino Liphola, Sabine Zerbian, Eva Zimmermann and the editors of this volume – for helpful comments and questions. Any errors of fact or interpretation are, of course, my responsibility.

References

Anderson, John & Colin Ewen (1987). *Principles of dependency phonology*. Cambridge: Cambridge University Press.

Baković, Eric (2007a). A revised typology of opaque generalisations. *Phonology* 24, 217-259.

Baković, Eric (2007b). Local assimilation and constraint interaction. In: Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 335-352.

Barnes, Jonathan (2006). *Strength and Weakness at the Interface: Positional Neutralization in Phonetics and Phonology*. Berlin: Mouton de Gruyter.

Beckman, Jill N. (1997). Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* 14, 1-46.

Bermudez-Otero, Ricardo (In preparation). *Stratal Optimality Theory*. (Oxford Studies in Theoretical Linguistics). Oxford: Oxford University Press.

Broselow, Ellen, Su-I Chen & Marie Huffman (1997). Syllable weight: convergence of phonology and phonetics. *Phonology* 14, 47-82.

Casali, Roderic F. (1998). *Resolving Hiatus*. New York: Garland.

Clements, G.N. (1986). Compensatory lengthening and consonant gemination in LuGanda. In: Leo Wetzels & Engin Sezer, eds. *Studies in Compensatory Lengthening*. Dordrecht: Foris, 37-77.

Crosswhite, Katherine (2001). *Vowel reduction in Optimality Theory*. London: Routledge.

Crosswhite, Katherine (2003). Vowel reduction in Shimakonde. CASTL Kick-off Conference, University of Tromsø, 2-4 October 2003.

Crosswhite, Katherine (2004). Vowel reduction. In: Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically Based Phonology*, Cambridge: Cambridge University Press, 191-231.

Downing, Laura J. (1995). *The Tonal Phonology of Jita*. Munich: Lincom Europa.

Downing, Laura J. (1998). On the prosodic misalignment of onsetless syllables. *NLLT* 16, 1-52.

Downing, Laura J. (1999). Prosodic Stem ≠ Prosodic Word in Bantu. In: T. Alan Hall & Ursula Kleinhenz (eds.), *Studies on the Phonological Word*. Amsterdam: John Benjamins, 73-98.

Downing, Laura J. (2000). Morphological and prosodic constraints on Kinande verbal reduplication. *Phonology* 17, 1-38.

Downing, Laura J. (2005). On the ambiguous segmental status of nasals in homorganic NC sequences. In: Marc van Oostendorp & Jeroen van de Weijer (eds.) *The Internal Organization of Phonological Segments*. Berlin: Mouton de Gruyter, 183-216.

Downing, Laura J. (2006). *Canonical Forms in Prosodic Morphology*. Oxford: Oxford University Press.

Ettlinger, Marc (2008). *Input Drive Opacity*. Ph.D. dissertation, University of California-Berkeley.

Goldsmith, John (1985). Vowel harmony in Khalkha Mongolian, Yaka, Finnish and Hungarian. *Phonology Yearbook* 2, 253-275.

Goldrick, Matthew (2000). Turbid Output Representations and the Unity of Opacity. In: *Proceedings of NELS* 30, 231-245.

Green, Antony D. (2007). *Phonology Limited*. Potsdam: Universitätsverlag Potsdam.

Hargus, Sharon & Virginia Beavert (2006). Word-initial clusters and minimality in Yakima Sahaptin. *Phonology* 23, 21-58.

Harris, John (1987). Non-structure-preserving rules in Lexical Phonology: Southeastern Bantu Harmony. *Lingua* 73, 255-292.

Harris, John (1990). Segmental complexity and phonological government. *Phonology* 7, 255-300.

Harris, John (1994). Monovalency and opacity: Chichewa height harmony. *UCL Working Papers in Linguistics*, 509-547.

Harris, John (1996). Phonology output is redundancy-free and fully interpretable. In: Jacques Durand & Bernard Laks (eds.), *Current Trends in Phonology: Models and Methods*. Salford: ESRI, 305-332.

Harris, John (1997). Licensing inheritance: an integrated theory of neutralization. *Phonology* 14, 315-270.

Harris, John (2005). Vowel reduction as information loss. In: Philip Carr, Jacques Durand & Colin J. Ewen (eds.), *Headhood, Elements, Specification and Contrastivity*. Amsterdam: John Benjamins, 119-132.

Harris, John & Geoff Lindsey (1995). The elements of phonological representation. In: Jacques Durand & Francis Katamba (eds.), *Frontiers of Phonology*. London: Longman, 34-79.

Harris, John & Geoff Lindsey (2000). Vowel patterns in mind and sound. In: Noel Burton-Roberts, Philip Carr & Gerard Docherty (eds.), *Phonological Knowledge: Conceptual and Empirical Issues*, Oxford: Oxford University Press, 185-205.

Hualde, Jose Ignacio (1989). Autosegmental and metrical spreading in the vowel-harmony systems of northwestern Spain. *Linguistics* 27, 773-805.

van der Hulst, Harry & Jeroen van de Weijer (1995). Vowel harmony. In: John A. Goldsmith (ed.), *Handbook of Phonological Theory*. Cambridge, Mass.: Blackwell, 495-534.

van der Hulst, Harry & Nancy A. Ritter (1999a). Head-driven phonology. In: Harry van der Hulst & Nancy A. Ritter (eds.), *The Syllable: Views and Facts*. Berlin: Mouton de Gruyter, 113-167.

van der Hulst, Harry & Nancy A. Ritter (1999b). No sympathy for opacity. *CLS* 35:1, 153-173.

Hyman, Larry M. (1998). Positional prominence and the 'prosodic trough' in Yaka. *Phonology* 15, 41-75.

Hyman, Larry M. (1999). The historical interpretation of vowel harmony in Bantu. In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu Historical Linguistics: Theoretical and Empirical Perspectives*. Stanford: CSLI, 235-295.

Itô, Junko & Armin Mester (1997). Sympathy theory and German truncations. In: Viola Miglio & Bruce Morén (eds.), *University of Maryland Working Papers in Linguistics 5*, 117-139.

Itô, Junko & Armin Mester (2003a). On the sources of opacity in OT: coda processes in German. In: Caroline Féry & Ruben van de Vijver (eds.), *The Syllable in Optimality Theory.* Cambridge: Cambridge University Press, 271-303.

Itô, Junko & Armin Mester (2003b). *Japanese Morphophonemics: Markedness and Word Structure*. Cambridge, Mass.: The MIT Press.

Kager, René (1999). *Optimality Theory*. Cambridge: Cambridge University Press.

Kaun, Abigail (1995). *The Typology of Rounding Harmony: An Optimality Theoretic Approach*. Ph.D. dissertation, UCLA.

Kaun, Abigail (2004). The typology of rounding harmony. In: Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically Based Phonology*. Cambridge: Cambridge University Press, 87-116.

Kiparsky, Paul (2000). Opacity and cyclicity. *The Linguistic Review* 17, 351-365.

Kraal, Peter (2005). *A Grammar of Makonde*. Ph.D. thesis, University of Leiden.

Krämer, Martin. 2008. English schwa insertion before liquids and phonological opacity. *Proceedings of CLS* 41.1, 267-282.

Liphola, Marcelino (2001). *Aspects of Phonology and Morphology of Shimakonde*. Ph.D. dissertation, OSU.

McCarthy, John J. (1999). Sympathy and phonological opacity. *Phonology* 16, 331-399.

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

McCarthy, John J. (2003). Comparative markedness. *Theoretical Linguistics* 29, 1-51.

Mielke, Jeff, Mike Armstrong & Elizabeth Hume (2003). Looking through opacity. *Theoretical Linguistics* 29, 123-139.

Morén, Bruce (2006). Consonant-vowel interactions in Serbian: Features, representations and constraint interactions. *Lingua* 116, 1198-1244.

Mutaka, Ngessimo (1994). *The Lexical Tonology of Kinande*. Munich: LINCOM EUROPA.

Nevins, Andrew (2007). Review of Barnes, Jonathan, *Strength and Weakness at the Interface: Positional Neutralization in Phonetics and Phonology*. *Phonology* 24, 461-469.

Odden, David (2006). Minimality and onsetless syllables in Zinza. *Phonology* 23, 431-441.

Orgun, Cemil Orhan (2001). English r-insertion in Optimality Theory. *NLLT* 19, 737-749.

Prince, Alan & Paul Smolensky (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell.

Rose, Yvan (2000). *Headedness and Prosodic Licensing in the L1 Acquisition of Phonology*. PhD dissertation, McGill University.

Rosenthall, Sam (1994). *Vowel/Glide Alternation in a Theory of Constraint Interaction*. PhD dissertation, University of Massachusetts-Amherst.

Rubach, Jerzy (2003). Polish palatalization in Derivational Optimality Theory. *Lingua* 113, 197-237.

Steriade, Donca (1994). Positional neutralization and the expression of contrast. Ms. UCLA.

Steriade, Donca (1995). Underspecification and markedness. In: John A. Goldsmith (ed.), *Handbook of Phonological Theory*. Cambridge, Mass.: Blackwell, 114-174.

Uffmann, Christian (2005). Optimal geometries. In: Marc van Oostendorp & Jeroen van de Weijer (eds.), *The Internal Organization of Phonological Segments*. Berlin: Mouton de Gruyter, 27-62.

Uffmann, Christian (2007). Restricting GEN. In: Sylvia Blaho, Patrik Bye & Martin Krämer, (eds.), *Freedom of Analysis?* Berlin: Mouton de Gruyter, 281-312.

Vaux, Bert & Andrew Nevins, eds. (2008). *Rules, Constraints, and Phonological Phenomena.* Oxford: Oxford University Press.

Walker, Rachel (2005). Weak triggers in vowel harmony. *NLLT* 23, 917-989.

Van de Weijer, Jeroen (1996). *Segmental Structure and Complex Segments*. Ph.D. dissertation, HIL.

Zubizarreta, Maria Luisa (1979). Vowel harmony in Andalusian Spanish. *MITWPL* 1,

ZAS Papers in Linguistics were originally published by the Forschungsschwerpunkt Allgemeine Sprachwissenschaft, Typologie und Universalienforschung (FAS, Research Center for General Linguistics, Typology and Universals). The Center is now known as *Zentrum für Allgemeine Sprachwissenschaft* (ZAS) under the auspices of the BMBF - the *Bundesministerium für Bildung und Forschung* (Federal Ministry of Education and Research) and the State of Berlin. The Center currently has research projects in syntax, semantics, morphology, phonology, phonetics as well as language contact and language acquisition. ZAS provides a forum for the exchange of ideas in the academic community of the Berlin area through lectures, seminars, workshops and conferences. The Center cooperates with other universities in Germany, and sponsors visits by scholars from Europe and America.

Director: Manfred Krifka

For further information about ZAS, please consult our website:
http://www.zas.gwz-berlin.de

or write to:
Manfred Krifka, Director
Zentrum für Allgemeine Sprachwissenschaft
Schützenstr. 18
D-10117 Berlin
Germany
E-mail: krifka@zas.gwz-berlin.de

ZAS Papers in Linguistics reflect the ongoing work at ZAS. They comprise contributions of ZAS researchers as well as visiting scholars. Issues are available on an exchange basis or on request. For further information, please write to:
Sekretariat
Zentrum für Allgemeine Sprachwissenschaft
Schützenstr. 18
D-10117 Berlin
Germany
E-mail: sprach@zas.gwz-berlin.de
Phone: +49 30 20 19 24 01
Fax: +49 30 20 19 24 02

Later issues can also in part be downloaded at the ZAS website:
http://www.zas.gwz-berlin.de/zaspil.html

Cover design: Mathias Krüger, Mechthild Bernhard and the CMS, HU Berlin.

ZAS Papers in Linguistics previous issues (please consult the ZAS website for full tables of content, and for availability):

ZASPiL 1 Artemis Alexiadou, Nanna Fuhrop, Paul Law and Sylvia Löhken (eds.): ***Papers on syntax and semantics.*** Contributions by Ewald Lang, Anna Cardinaletti & Michal Starke, Jaklin Kornfilt, Ewald Lang, Renate Steinitz and Chris Wilder.

ZASPiL 2 Artemis Alexiadou, Nanna Fuhrop, Paul Law and Sylvia Löhken (eds.): ***Papers on syntax and morphology.*** Contributions by Peter Ackema & Ad Neeleman, Gaberell Drachman, Ursula Kleinhenz, Sylvia Löhken, André Meinunger, Renate Raffelsiefen, Iggy Roca, M. M. Verhijde and Wolfgang Ullrich Wurzel.

ZASPiL 3 Artemis Alexiadou, Nanna Fuhrop, Paul Law and Sylvia Löhken (eds.): ***Papers on syntax and phonology.*** Contributions by Ulrike Demske, Damaris Nübling, Wolgang Sternefeld and Susan Olsen.

ZASPiL 4 Artemis Alexiadou, Nanna Fuhrop, Paul Law and Sylvia Löhken (eds.): ***Papers on syntax and learning***. Contributions by Artemis Alexiadou & Elena Anagnostopoulou, Hans-Martin Gärtner, Jaklin Kornfilt, Paul Law, André Meinunger, Ralf Vogel & Markus Steinbach and Chris Wilder.

ZASPiL 5 Artemis Alexiadou, Nanna Fuhrop, Paul Law and Sylvia Löhken (eds.): ***Papers on syntax***. Contributions by Artemis Alexiadou & Spyridoula Varlokosta, Elena Herburger, Paul Law, Alan Munn, Cristina Schmitt, Juan Uriagereka, Chris Wilder and Petra de Wit & Maaike Schoorlemmer.

ZASPiL 6 Artemis Alexiadou, Nanna Fuhrop, Paul Law and Sylvia Löhken (eds.): ***Papers on clitics***. Contributions by Artemis Alexiadou & Elena Anagnostopoulou, Piotr Banski, Monika Baumann, Loren A. Billings, Damir Cavar, Uwe Junghanns, Ursula Kleinhenz, Jaklin Kornfilt, Christine Maaßen, Cristina Schmitt, Petra de Wit & Maaike Schoorlemmer, Maaike Schoorlemmer, Chris Wilder and Ilse Zimmerman.

ZASPiL 7 Artemis Alexiadou, Nanna Fuhrop, Paul Law and Ursula Kleinhenz (eds.): ***Papers on phonetics and phonology***. Contributions by Loren Billings, Christina Kramer & Catherine Rudin, Janet Grijzenhout, T. A. Hall, Haike Jacobs, Peter M. Janker, Manuela Noske, Bernd Pompino-Marschall, Peter M. Janker and Christine Mooshammer.

ZASPiL 8 Artemis Alexiadou, Nanna Fuhrop, Paul Law and Ursula Kleinhenz (eds.): ***Papers on syntax, semantics, phonology and acquisition***. Contributions by Artemis Alexiadou & Elena Anagnostopolou, Artemis Alexiadou & Melita Stavrou, Dagmar Bittner, Hans-Olav Enger, Manuela Friedrich, Wladimir D. Klimonow and Heike Wiese.

ZASPiL 9 Artemis Alexiadou, Nanna Fuhrop, Paul Law and Ursula Kleinhenz (eds.): ***Papers on focus and ellipsis***. Contributions by Loren A. Billings, Horst-Dieter Gasde, Uwe Junghanns, André Meinunger, Kerstin Schwabe and Ning Zhang.

ZASPiL 10 Artemis Alexiadou, Nanna Fuhrop, Paul Law and Ursula Kleinhenz (eds.): ***Papers on syntax of clefts, pseudo-clefts, relative clauses, and the semantics of present perfect*** Contributions by Artemis Alexiadou & Anastasia Giannakidou, Marcel den Dikken, André Meinunger and Chris Wilder, Caroline Heycock & Anthony Kroch, Jason Merchant, Renate Musan, Wolfgang Sternefeld, Peter Svenonius and Chris Wilder.

ZASPiL 11 Artemis Alexiadou, Nanna Fuhrop, Ursula Kleinhenz and Paul Law (eds.): ***Papers on morphology and phonetics.*** Contributions by H.G. Tillmann, K.J. Kohler, P.A. Keating, F. Schiel & A. Kipp, Ch. Draxler, A. Mengel, R. Benzmüller & M. Grice, A. P. Simpson, L. Ellis & W. J. Hardcastle, K. Russell, E. Farnetani, M. Jessen, B. Kröger, L. Faust and B. Pompino-Marschall & P. M. Janker.

ZASPiL 12 Artemis Alexiadou, Nanna Fuhrop, Ursula Kleinhenz and Paul Law (eds.): ***Papers on morphology and phonology.*** Contribution by Ursula Kleinhenz.

ZASPiL 13 Artemis Alexiadou, Nanna Fuhrop, Ursula Kleinhenz and Paul Law (eds.): ***Papers on morphology.*** Contributions by Werner Abraham, Nanna Fuhrhop, Livio Gaeta, Rüdiger Harnisch, Heinrich Hettrich, Bernhard Hurch, Wladimir D. Klimonow, Ekkehard König & Peter Siemund, Elisabeth Leiss, Elke Ronneberger-Sibold, Peter Schrijver, Richard Schrodt, Anja Voeste and Wolfgang Ullrich Wurzel.

ZASPiL 14 Ewald Lang and Ljudmila Geist (eds.): ***Papers on semantics of the copula.*** Contributions by Ewald Lang, Ljudmila Geist, Claudia Maienborn, Gerhard Jäger, Johannes Dölling, Ilse Zimmermann, Ning Zhang, Renate Musan, Renate Steinitz and Cristina Schmitt.

ZASPiL 15 Artemis Alexiadou, Nanna Fuhrop, Ursula Kleinhenz and Paul Law (eds.):
Papers on language change and language acquisition.Contributions by Werner Abraham, Nanna Fuhrhop, Gregory K. Iverson & Joseph C. Salmons, Wladimir Klimonow, Michail Kotin, Peter Suchsland, Letizia Vezzosi, Dagmar Bittner, Manuela Friedrich, Natalia Gagarina, Insa Gülzow and Theodore Marinis.

ZASPiL 16 Ewald Lang (ed.*)*:
Papers on copular- and AUX-constructions. Contributions by Ewald Lang, Gerhard Jäger, Michail Kotin, Cristina Schmitt, Nanna Fuhrhop, Ljudmila Geist and Joanna Blaszczak

ZASPiL 17 Cathrine Fabricius-Hansen, Ewald Lang and Claudia Maienborn (eds.):
Approaching the grammar of adjuncts. Proceedings of the Oslo conference. Contributions by Assinja Demjjanow & Anatoli Strigin, Johannes Dölling, David Dowty, Thomas Ernst, Marina V. Filipenko, Werner Frey, Graham Katz, Claudia Maienborn, Barbara Partee & Vladimir Borschev, Karin Pittner, Inger Rosengren, Susan Rothstein, Benjamin Shaer, Arnim von Stechow and Ilse Zimmermann.

ZASPiL 18 Dagmar Bittner, Wolfgang U. Dressler and Marianne Kilani-Schoch (eds.):
First verbs: On the way to mini-paradigms. Contributioins by Dagmar Bittner, Wolfgang U. Dressler & Marianne Kilani-Schock, Sabine Klampfer, Insa Gülzow, Klaus Laalo, Barbara Pfeiler, Marianne Kilani-Schoch, Carmen Aquirre, Antigone Katicic, Pawel Wójcik and Natalia Gagarina.

ZASPiL 19 T. A. Hall and Marzena Rochon (eds.):
Investigations in prosodic phonology. Contributions by Bozena Cetnarowska, Laura J. Downing, T. A. Hall, David J. Holsinger, Arsalan Kahnemuyipour, Renate Raffelsiefen, Marzena Rochon and Caroline R. Wiltshire.

ZASPiL 20 Kerstin Schwabe, André Meinunger and Horst-Dieter Gasde (eds.):
Issues on topics. Contributions by André Meinunger, Yen-Hui Audrey Li, Liejiong Xu, Danqing Liu, Marie-Claude Paris, Kleanthes K. Grohmann, Artemis Alexiadou, Werner Frey and Michael Grabski.

ZASPiL 21 Oliver Teuber and Nanna Fuhrhop (eds.):
Papers for Ewald Lang. Contributions by Dagmar Bittner and Klaus-Michael Köpcke, Werner Frey, Nanna Fuhrhop, Michael Grabski, Kleanthes Grohmann, Tracy Alan Hall, Wladimir D. Klimonov, Paul Law, Kerstin Schwabe, Patrick O. Steinkrüger, Oliver Teuber and Wolfgang Ullrich Wurzel.

ZASPiL 22 Gerhard Jäger, Anatoli Strigin, Chris Wilder and Ning Zhang (eds.):
Papers on Predicative Constructions. Contributions by John F. Bailyn, Misha Becker Patrick Brandt, Assinja Demjjanow & Anatoli Strigin, Roland Hinterhölzl, Orin Percus, Susan Rothstein, Sze-Wing Tang, Wei-Tien Dylan Tsai and Ning Zhang.

ZASPiL 23 Klaus von Heusinger and Kerstin Schwabe (eds.):
Information Structure and the Referential Status of Linguistic Expressions. Contributions by Franz-Josef d'Avis, Carsten Breul, Dina Brun, Daniel Büring, Donka F. Farkas, Hans-Martin Gärtner, Michael Hegarty, Jeanette K. Gundel & Kaja Borthen, Jügen Lenerz, Horst Lohnstein, Norberto Moreno & Isabel Pérez, Paul Portner, Ingo Reich, Elisabeth Stark, Anita Steube and Carla Umbach.

ZASPiL 24 Klaus von Heusinger and Kerstin Schwabe (eds.):
Sentence Type and Specificity. Contributions by Raffaella Zanuttini & Paul Portner, Horst-Dieter Gasde, Kleanthes K. Grohmann, Remus Gergel, Kerstin Schwabe, Klaus von Heusinger, Bart Geurts, Nicholas Asher and Werner Frey.

ZASPiL 25 Anatoli Strigin and Assinja Demjjanow (eds.):
Secondary Predication in Russian. Contributions by Anatoli Strigin and Assinja Demjjanow.

ZASPiL 26 Ning Zhang (ed.):
The Syntax of Predication. Contributions by David Adger & Gillian Ramchand, Tor A. Åfarli & Kristin M. Eide, Ana Ardid-Gumiel, Kleanthes K. Grohmann, Youngjun Jang & Siyoun Kim, Jaume Mateu, Joan Rafel, Kylie Richardson, Peter Svenonius and Ning Zhang.

ZASPiL 27 Ewald Lang und Ilse Zimmermann (eds.):
Nominalizations. Contributions by Fritz Hamm & Michiel von Lambalgen, Veronika Ehrich, Veronika Ehrich & Irene Rapp, Ulrike Demske, Artemis Alexiadou, Klaus von Heusinger and Ilse Zimmermann.

ZASPiL 28 T. A. Hall, Bernd Pompino-Marschall and Marzena Rochon (eds.):
Papers on Phonetics and Phonology: The Articulation, Acoustics and Perception of Consonants. Contributions by Hansook Choi, Silke Hamann, Kenneth de Jong, Kyoko Nagao & Byung-jin Lim, Lisa M. Lavoie, Jeff Mielke, Marianne Pouplier & Louis Goldstein, Daniel Recasens, Rachid Ridouane, Zoë Toft, Nathalie Vallée, Louis-Jean Boë, Jean-Luc Schwartz and Pierre Badin & Christian Abry.

ZASPiL 29 Dagmar Bittner and Natalia Gagarina (eds.):
The Acquisition of Aspect. Contributions by Dagmar Bittner, Annerieke Boland Dina Brun & Babyonyshev, Sophia Delidaki & Spyridoula Varlokosta, Alison Gabriele, Gita Martohardjona & William McClure, Miren Hodgson, Linae Jeschull, Claire Martinot, Maja Andel & Sunil Kumar, Ayumi Matsuo, Barbara Schmiedtová, Yasuhiro Shirai and Ursula Stephany & Maria Voeikove.

ZASPiL 30 Regine Eckardt (ed.):
Questions and Focus. Contributions by Florian Schwarz and Markus Fischer.

ZASPiL 31 Dagmar Bittner (ed.):
Von starken Feminina und schwachen Maskulina. Contribution by Dagmar Bittner.

ZASPiL 32 T. A. Hall and Silke Hamann (eds.):
Papers in Phonology and Phonetics. Contributions by Karen Baertsch, Stuart Davis, Jana Brunner, Susanne Fuchs, Pascal Perrier, Hyeon-Zoo Kim, Antony Dubach Green, T. A. Hall, Silke Hamann, Jaye Padgett and Marzena Zygis.

ZASPiL 33 Natalia Gagarina and Dagmar Bittner (eds.):
Studies on the Development of Grammar in German, Russian and Bulgarian. Contributions by Dagmar Bittner, Natalia Gagarina, Milena Kühnast, Velka Popova, Dimitar Popov and Franziska Bewer.

ZASPiL 34 Paul Law (ed.):
Proceedings of AFLA 11, ZAS, Berlin 2004. Contributions by Edith Aldridge, Loren Billings & Daniel Kaufman, Chun-Mei Chen, Wen-yu Chiang & Fang-mei Chiang, Wen-yu Chiang & I Chang-Liao, Mark Donohue, Nelleke Goudswaard, Nikolaus Himmelmann, Arthur Holmer, Arsalan Kahnemuyipour & Diane Massam, Daniel Kaufman, Tomoko Kawamura, Edward Keenan & Cecile Manorohanta, Yuko Otsuka, Ileana Paul, Matt Pearson, Eric Potsdam, Craig Thiersch.

ZASPiL 35 Ben Shaer, Werner Frey and Claudia Maienborn (eds.):
Proceedings of the Dislocated Elements Workshop, ZAS Berlin, November 2003. Contributions by Maria Alm, Olga Arnaudova, Betty Birner, Ariel Cohen, Cécile de Cat, Judit Gervain, Beáta Gyuris, Liliane Haegeman, Konstantina Haidou, Anke Holler, Ruth Kempson & Ronnie Cann & Jieun Kiaer, Anikó Lipták, Eric Mathieu, Sam Mchombo & Yukiko Morimoto, Nicola Munaro & Cecilia Poletto, Frederick J. Newmeyer, Andreas Nolda, Javier Pérez-Guerra & David Tizón-Couto, Benjamin Shaer & Werner Frey, Nicholas Sobin, Augustin Speyer, Malte Zimmermann.

ZASPiL 36 Anatoli Strigin:
Blocking Resultative Secondary Predication in Russian.

ZASPiL 37 Susanne Fuchs and Silke Hamann (eds.):
Papers in Phonetics and Phonology. Contributions by Laura J. Downing, Christian Geng, Antony D. Green, T. A. Hall, Silke Hamann, Al Mtenje, Bernd Pompino-Marschall, Christine Mooshammer, Sabine Zerbian, and Marzena Zygis.

ZASPiL 38 Jason Mattausch:
On the Optimization and Grammaticalization of Anaphora

ZASPiL 39 Jana Brunner:
Supralaryngeal mechanisms of the voicing contrast in velars

ZASPiL 40 Susanne Fuchs, Pascal Perrier and Bernd Pompino-Marschall (eds.):
Speech Production and Perception: Experimental analyses and models. Contributions by Susanne Albert, Jérôme Aubin, Pierre Badin, Sophie Dupont, Sascha Fagel, Roland Frey, Alban Gebler, Cédric Gendrot, Julia Gotto, Abraham Hirschberg, Ian S. Howard, Mark A. Huckvale, Bernd J. Kröger, Ines Lopez, Shinji Maeda, Lucie Ménard, Christiane Neuschaefer-Rube, Xavier Perlorson, Pascal Perrier, Hartmut R. Pfitzinger, Bernd Pompino-Marschall, Nicolas Ruty, Walter Sendlmeier, Willy Serniclaes, Antoine Serrurier, Annemie Van Hirtum and Ralf Winkler.

ZASPiL 41 Susanne Fuchs:
Articulatory correlates of the voicing contrast in alveolar obstruent production in German.

ZASPiL 42 Christian Geng, Jana Brunner and Daniel Pape (eds.):
Papers in Phonetics and Phonology. Contributions by Jana Brunner, Katrin Dohlus, Susanne Fuchs, Christian Geng, Silke Hamann, Mariam Hartinger, Phil Hoole, Sabine Koppetsch, Katalin Mády, Victoria Medina, Christine Mooshammer, Pascal Perrier, Uwe D. Reichel, Anke Sennema, Willy Serniclaes, Krisztián Z. Tronka, Hristo Velkov and Marzena Zygis.

ZASPiL 43 Laura J. Downing, Lutz Marten, Sabine Zerbian (eds.):
Papers in Bantu Grammar and Description. Contributions by Leston Buell, Lisa Cheng, Laura J. Downing, Ahmadi Kipacha, Nancy C. Kula, Lutz Marten, Anna McCormack, Sam Mchombo, Yukiko Morimoto, Derek Nurse, Nhlanhla Thwala, Jenneke van der Wal and Sabine Zerbian.

ZASPiL 44 Christian Ebert and Cornelia Endriss (eds.):
Proceedings of the Sinn und Bedeutung 10. Contributions by Stavros Assimakopoulos, Maria Averintseva-Klisch, Kata Balogh, Sigrid Beck & Arnim von Stechow, Adrian Brasoveanu, Ariel Cohen, Paul Dekker, Ljudmila Geist, Wilhelm Geuder, Wilhelm Geuder & Matthias Weisgerber, Elsi Kaiser, Elsi Kaiser & Jeffrey T. Runner & Rachel S. Sussman & Michael K. Tanenhaus, Dalina Kallulli, Mana Kobuchi-Philip, Sveta Krasikova & Ventsislav Zhechev, Eric McCready, Telmo Móia, Karina Veronica Molsing, Fabrice Nauze, Francesca Panzeri, Doris Penka, Daniel Rothschild, Florian Schwarz, Torgrim Solstad, Stephanie D. Solt, Tamina Stephenson, Rachel Szekely, Lucia M. Tovena, Anna Verbuk, Matthias Weisgerber, Hedde Zeijlstra, Malte Zimmermann, Eytan Zweig.

ZASPiL 45 Sabine Zerbian:
Expression of Information Structure in the Bantu Language Northern Sotho

ZASPiL 46 Ines Fiedler & Anne Schwarz (eds.):
Papers on Information Structure in African Languages. Contributions by Klaus Abels & Peter Muriungi, Enoch O. Aboh, Robert Carlson, Bernard Caron, Klaudia Dombrowsky-Hahn, Wilfrid H. Haacke, Angelika Jakobi, Susie Jones, Gregory Kobele & Harold Torrence, H. Ekkehard Wolff & Doris Löhr.

ZASPiL 47 Barbara Stiebels (ed.):
Studies in Complement Control

ZASPiL 48 Dagmar Bittner & Natalia Gagarina (eds.):
Intersentential Pronominal Reference in Child and Adult Language. Proceedings of the Conference on Intersentential Pronominal Reference in Child and Adult Language. Contributions by Jeanette K. Gundel, Dimitris Ntelitheos & Melinda Kowalsky, H. Wind Cowles, Peter Bosch & Carla Umbach, Gerlof Bouma & Holger Hopp, Petra Hendriks, Irene Siekman, Erik-Jan Smits & Jennifer Spenader, Dagmar Bittner, Natalia Gagarina, Milena Kühnast, Insa Gülzow & Natalia Gagarina.

ZASPiL 49 Marzena Zygis & Susanne Fuchs (eds.):
Papers in Phonetics and Phonology. Contributions by Claire Brutel-Vuilmet & Susanne Fuchs, Marzena Zygis, Laura Downing, Elke Kasimir, Daniel Recasens, Silke Hamann & Susanne Fuchs, Anna Bloch-Rozmej, Grzegorz Nawrocki, Cédric Patin.

ZASPiL 50 Hristo Velkov:
Akustische Analysen zur koartikulatorischen Beeinflussung des frikativischen Teils stimmloser Plosive im Deutschen und im Bulgarischen

ZASPiL 51 Anton Benz & Reinhard Blutner (eds.):
Papers on Pragmasemantics. Contributions by Anton Benz, Reinhard Blutner, Michael Franke, Elena Karagjosova, Tom Lenz, and Henk Zeevat.