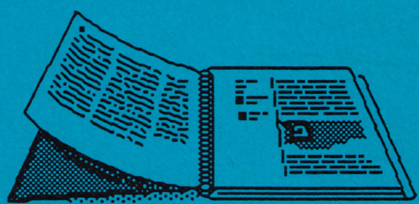
Zentrum für Allgemeine Sprachwissenschaft,
Sprachtypologie und Universalienforschung

ZAS Papers in Linguistics

Volume 7
December 1996



Edited by

Artemis Alexiadou
Nanna Fuhrhop
Paul Law
Sylvia Löhken

ISSN 0947-7055

The FAS Papers in Linguistics was 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, Typologie und Universalienforschung (ZAS)* under the auspices of the Deutsche Forschungsgemeinschaft (The German Research Foundation) and the Land of Berlin. The Center currently has research projects in syntax, semantics, morphology, phonology, phonetics as well as language change 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: Ewald Lang

For further information on the Center, please write to:

Ewald Lang, Director
ZAS
Jägerstr. 10/11
D-10117 Berlin
Germany

Telephone: +49 30 20 19 24 01
Fax: +49 30 20 19 24 02

E-mail: sprach@fas.ag-berlin.mpg.de

ZAS Papers in Linguistics is intended to reflect the on-going work at the Center. It publishes papers by its staff members and visiting scholars. The publication is available on an exchange basis. For further information, please write to:

The editors
ZAS Papers in Linguistics
Jägerstr. 10/11
D-10117 Berlin
Germany

Telephone: +49 30 20 19 24 04
Fax: +49 30 20 19 24 02

E-mail: faspil@fas.ag-berlin.mpg.de

*Accordingly, the title of the series changes to *ZAS Papers in Linguistics*.

Table of Contents

Syntax and prosody: *li* questions in Macedonian and Bulgarian

Loren Billings, Christina Kramer, and Catherine Rudin

Abstract. The distribution of the yes/no-interrogative clitic *li* in Macedonian and Bulgarian reveals a complex interaction of syntax with prosodic factors. The underlying syntactic uniformity of *li* questions in the two languages is obscured by a series of prosodic phenomena affecting one language or the other. In Macedonian two prosodic factors affect the placement of *li*: optional stressing of auxiliaries and optionally allowing certain sequences of words to have only one stress. In Bulgarian two different prosodic phenomena are relevant: stressing of clitics after the negative element *ne* and inversion of initial clitics with the following verb. When these four factors are controlled for, the syntax of *li* questions in the two languages is startlingly homogeneous. If no element is focused (i.e., moved to SpecCP), then, in both languages, the tensed verb head-incorporates into *li* in C. Usage differences complicate the picture somewhat as well.

1. Introduction

Much recent work in theoretical linguistics relies upon cross-linguistic comparison to elucidate the limits and causes of variation in Universal Grammar; that is, to establish just what structures and processes are possible in human languages. It has often proven especially useful to compare related languages, thus teasing out differences between languages which share much of their grammatical structure. The two languages considered in this article, Macedonian and Bulgarian, are very closely related genetically, both being South Slavic. In addition, they are areally related: both participate in the Balkan *Sprachbund*, sharing many of the contact-induced grammatical features common to the Balkan area. It is thus of particular interest to note cases in which the grammars of the two languages diverge. We expect that differences between such closely related languages will involve rather superficial factors, changes which could be introduced into a grammar in a relatively short time, even if their surface effects are considerable.

We examine here one construction, the yes/no question formed with the interrogative clitic *li*, which displays interesting similarities and differences in the two languages. Superficially, Macedonian and Bulgarian appear to have distinct conditions on the placement of *li*, particularly with respect to the other clitics. We argue, however, that the syntactic behavior of *li* is fundamentally identical in the two languages, as is their clitics' placement. Several prosodically realized differences between the two languages interact with their common syntax to

* Thanks to the following colleagues for their assistance: P. Ambrosiani, W. Browne, D. Ćavar, L. Hammer, R. Izvorski, T. King, B. Nilsson, J. Toman, O. Tomić, M. Votruba, D. Zec and audiences at annual AATSEEL meetings (Billings 1994, Rudin 1993b, Rudin and Kramer 1994), Cornell University (Billings 1996a), FASL-5 (Rudin 1996), FLSM-4 (Rudin 1993a), and—in part—the University of Massachusetts (Rudin, King and Izvorski 1995). Any inconsistencies, mistakes, or misinterpretations of their worthwhile contributions, of course, are only our fault. Billings's contribution was supported primarily by a visiting fellowship at ZAS, Berlin, during the summer of 1996; thanks especially to Tracy Hall for making this visit possible. Additional support to Billings came from a Foreign Language Area Studies fellowship awarded by Cornell University.

produce contrasting surface orders. Differences in the usage of *li*, as well as other ways of forming yes/no questions in Macedonian, further differentiate the two languages. Thus, apparent syntactic differences turn out to be due to non-syntactic factors, both phonological (prosodic) and pragmatic.

This paper has the following organization: Following a description of some of the more striking surface differences between Macedonian and Bulgarian (in §2), we present an analysis (in §3) unifying the underlying syntactic distribution of *li* in the two languages. Then (in §4) we discuss four prosodic peculiarities in the two languages that obscure this syntactic uniformity. Following a brief background on the prosodic systems of the two languages (§4.1), we discuss four phenomena which obscure the placement of *li*: In Macedonian two lexical words, under certain circumstances, share a single word stress (§4.2). In the same language clitic forms of 'be' are sometimes stressed (§4.3). In Bulgarian the negative element *ne* causes the following element, even an otherwise-unstressable clitic, to be stressed (§4.4). Additionally, only in Bulgarian, pronominal and auxiliary clitics are prohibited from being clause-initial (§4.5). Finally (in §5), we show how pragmatic factors limit the acceptability of *li* in Macedonian.

2. The Problem

In both Macedonian and Bulgarian, as well as several other Slavic languages, yes/no questions may be formed by adding *li* to a declarative sentence. But at first glance, the position of *li* in some types of Macedonian questions appears to be radically different from its position in Bulgarian. Examples (1) and (2) display precisely opposite grammaticality judgments for positive yes/no questions with complementary word orders in Macedonian and Bulgarian, while (3) and (4) show the same disparity in negative questions, and (5) and (6) in conditional questions. (Throughout the paper we have boldfaced *li* in examples.)

(1) a. Go vide **li**? (Macedonian)
him$_{ACC}$ saw$_{3.SG}$ Q

b. *Vide **li** go?

'Did (s)he see him?'

(2) a. *Go vidja **li**? (Bulgarian)

b. Vidja **li** go?
saw$_{3.SG}$ Q him$_{ACC}$

'Did (s)he see him?'

(3) a. Ne go vide **li**? (Macedonian)
NEG him_{ACC} $\text{saw}_{3.SG}$ Q

b. *Ne go **li** vide?

c. Ne **li** go vide?[1]

'Didn't (s)he see him?'

(4) a. *Ne go vidja **li**? (Bulgarian)

b. Ne go **li** vidja?
NEG him_{ACC} Q $\text{saw}_{3.SG}$

c. *Ne **li** go vidja?

'Didn't (s)he see him?'

(5) a. Bi mi dal **li** pari? (Macedonian)
would me_{DAT} $\text{given}_{M.SG}$ Q money

b. *Bi **li** mi dal pari?

'Would he give me money?'

(6) a. *Bi mi dal **li** pari? (Bulgarian)

b. Bi **li** mi dal pari?
$\text{would}_{3.SG}$ Q me_{DAT} $\text{given}_{M.SG}$ money

'Would he give me money?'

Examples (1) through (6) all involve pronominal clitics; *go* 'him_{ACC}' or *mi* 'me_{DAT}'.[2] In *li* questions without other clitics, the strikingly divergent pattern of grammatical word orders disappears; in fact, normal word order is often identical in the two languages; compare the (a) and (b) versions of (7) through (9):

(7) a. Zboruvate **li** angliski? (Macedonian)
$\text{speak}_{2.PL}$ Q English

b. Govorite **li** angliski? (Bulgarian)
$\text{speak}_{2.PL}$ Q English

'Do you speak English?'

(8) a. Kniga **li** procita Anna? (Macedonian)
book Q $\text{read}_{3.SG}$

b. Kniga **li** procete Anna? (Bulgarian)
book Q $\text{read}_{3.SG}$

'Did Anna read a book?'

[1] Not all Macedonian speakers accept (3c). Olga Tomić has suggested to us that those who do may be confusing it with *neli* 'isn't it?'. See also our discussion of Enlarged Stress Domain in §4.2 below. Regarding the gloss of (3c), cf. §4.2 below, especially ex. (28b).

[2] Cf. Avgustinova 1994 and Hauge 1976 for a full discussion of the clitic system of Bulgarian. Cf. the end of §4.4 below regarding the differing accentuation of *bi* in (5)-(6).

(9) a. Od dve poluistini stanuva **li** celinja? (Macedonian)
from two half-truths becomes$_{3.SG}$ Q whole

b. Ot dve poluistini stava **li** edna cjala? (Bulgarian)
from two half-truths becomes$_{3.SG}$ Q one whole

'Do two half truths make a whole one?'

At this point one might posit special rules in each language for the placement of *li* relative to other clitics, or different positions for clitics in Macedonian and Bulgarian. This is unnecessary, however. We argue below that the syntactic position of *li* is the same in both languages, not only in cases like (7) through (9) where it appears identical, but also in (1) through (6). The conditions on word order with clitics do differ between the two languages, but the relevant conditions are prosodic, not syntactic. We adopt the analysis proposed by Rudin, King, and Izvorski 1995 and Izvorski, King, and Rudin 1996 for Bulgarian *li* questions, and show that it accounts for the corresponding Macedonian construction as well. In fact, the previously proposed structure is arguably clearer in Macedonian, where it is not obscured by certain prosodically controlled word-order changes. Such an analysis of *li* questions in Bulgarian is strengthened by this cross-linguistic comparison.

Questions with *li* in Bulgarian have received quite a bit of attention in the recent theoretical literature (see especially Rivero 1993, King 1993: 146–55/1994: 113–18/1995: 156–63, Izvorski 1994, Penčev 1993 and Rudin 1992; 1993a). But to the best of our knowledge, Macedonian *li* questions have not been analyzed in any detail,[3] although reference grammars contain brief descriptions, Friedman's (1993: 286–87) being the most complete.[4]

3. Syntactic Analysis

In both Macedonian and Bulgarian, *li* is a yes/no-interrogative particle which can also check a focus feature. As an enclitic, it is suffixed to a stressed element. We assume, with many recent analyses, that *li* is in C. (For arguments, see Rudin 1993a and Rivero 1993.) When C is [+focus], it checks a [+focus] feature on a fronted focus phrase in SpecCP (so-called specifier-head agreement), as in the tree in (10). Otherwise, when no focus phrase precedes *li*, the verb must raise and be

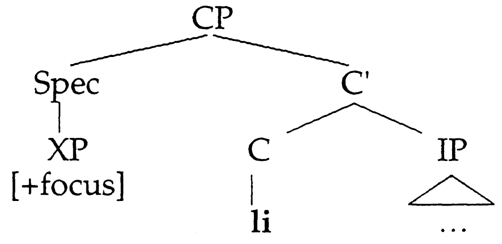
---

[3] Englund 1977 includes a relatively large corpus study of yes/no questions in both Bulgarian and Macedonian. While empirically very complete, Englund's study stops short of any extensive analysis. Tomić (1996a: 511) discusses Macedonian *li* briefly. (We were not able to locate Englund 1979 or Tomić 1996b in time for this paper. From their titles, these works promise to provide such analysis.)

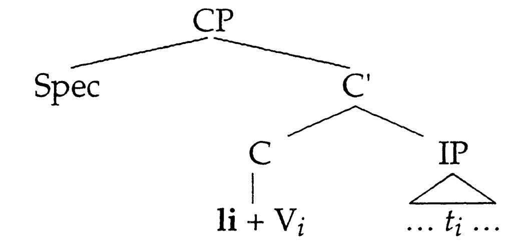
[4] Much of the this analysis holds for Russian as well (King 1993: 134–44/1994: 92–110/1995: 137–53; Rudin, King and Izvorski 1995), which has both the structures in (10) and (11). (Cf. Billings 1996b, however, for a phrasal-affix analysis of Russian *li*.) In Serbo-Croatian (Wayles Browne p.c., Mihaljević 1996, Rivero 1993, Wilder and Ćavar 1994) and Czech (Toman 1996: 508–09), *li* is limited to the structure in (11); (10) is not attested in those languages. Unlike modern (Serbo-) Croatian, Mihaljević 1996 reports, Croatian Church Slavonic did allow non-verbal elements to precede *li*. The other remaining South Slavic language, Slovenian, apparently no longer uses *li* as an independent morpheme (cf., however, *SSKJ* 1975: 600, which reports archaic examples, as well as one apparently clause-initial example of *li*!).

head-incorporated into C (V-to-C movement) to check this [+focus] feature, as in the tree in (11).[5]

(10)



(11)



Structure (10) is exemplified by the sentences in (8), where *kniga* is the phrase under focus. Structure (11) is seen in both (7) and (9); the portion of the tree labeled "*li* + V" is realized in these sentences as verb **followed** by *li* (*zboruvate li*, for instance). For reasons which will become clear below, we analyze this as right-adjunction of V to *li*, followed by prosodic inversion (following Halpern 1992/1995). In (9), in addition to V-raising, a non-focused XP (a topic phrase, *od dve poluistina*) is fronted to a position higher than SpecCP (probably adjoined to CP). It is not in a specifier-head relation with *li*, and therefore has no features checked by *li*. V-to-C raising is necessary to support the interrogative feature of *li* here, showing that *li* is not simply a second-position clitic (or clausal affix), but rather an element in the phrase structure.

The problematic cases with clitics, seen in (1) through (6), also have the structure in (11). The structure here, however, is less obvious; *li* does not always appear suffixed to the verb. We believe, however, that the facts in all of these cases are accounted for automatically, assuming the structure in (11) and given the existence in one or the other language of four prosodic distinctions.

4. Prosodic Distinctions

Macedonian and Bulgarian each manifest certain prosodic phenomena which affect the position of *li*. Before discussing these four phenomena, however, however, we provide some background on the prosodic system of the two languages' clitics. (Numbered examples throughout this section show clitics *italicized*, lexical words <u>underlined</u>, stressed syllables in ALL-CAPS, and syllable breaks with dots.)

[5] Slightly different proposals, such as those of Izvorski 1993, in which *li* is the head of FocusP rather than CP, would also be compatible with our analysis of Macedonian. We will not defend the details of the structures sketched in (10) and (11) here; the point is simply that in Macedonian, as in Bulgarian, *li* either incorporates a verb or checks a focus feature on an immediately preceding focused phrase in *li*'s specifier (SpecCP).

4.1 Primer on the Languages' Clitic-Prosody Systems

The two languages' prosodic systems, especially the stress system within the word, differ significantly, requiring separate descriptions. We begin with Macedonian, showing how the antepenultimate-stress system is especially useful in elucidating the uniqueness of *li* as a clitic. We then follow with a point-by-point comparison with Bulgarian, showing the specific ways in which the two differ.[6]

4.1.1 The Macedonian Clitic-Prosody System

Our description here of the Macedonian system is drawn from several works, many of which deal with the theoretical problem of accounting for antepenultimate stress—a somewhat exotic pattern cross-linguistically.[7] We refrain from entering into the somewhat lively debate about just how antepenultimate stress is to be formalized, assuming simply that such a mechanism exists. We concentrate instead on how to fit *li* into such a system, something that has been ignored to date in the literature, to our knowledge.

Macedonian's antepenultimate stress[8] allows us to assess whether a particular clitic is part of the same prosodic word (PrWd) with the lexical word

---

[6] Comparative works of the Macedonian and Bulgarian nominal (type-II and-III) clitics include Tomić 1996a and Elson 1976. See these especially for discussion of their distribution. For example, Macedonian shows a tripartite proximal-neutral-distal distinction, while some Bulgarian noun classes show NOM/object distinctions.

[7] These works on the theoretical issues surrounding Macedonian antepenultimate stress (some of which discuss the related issue of penultimate stress in Polish) include the following: Baerman (1996), Comrie (1976), Deevy (1995), Franks (1987; 1989; 1991), Garde (1968), Halle and Kenstowicz (1991), Hammond (1989), Kager (1993) and Kenstowicz (1991).

[8] See works listed in the preceding footnote for discussion of words with exceptional stress. These have either penultimate or final stress. As the following two representatives of each type of exception show (respectively), however, the addition of syllables (such as the plural marker or an article) to the end of the word regularize them to antepenultimate stress:

	INDEF.SG	DEF.SG	INDEF.PL	DEF.PL	Gloss
(i)	te.le.VI.zor	te.le.VI.zo.r*ot*	te.le.VI.zo.ri	te.le.vi.ZO.ri.*te*	'television'
(ii)	de.le.GAT	de.le.GA.t*ot*	de.le.GA.ti	de.le.GA.ti.*te*	'delegate'

That is, as shown by (i), whereas the INDEF.SG has exceptional penult stress, the addition of one syllable (in either the DEF.SG or INDEF.PL columns) keeps the stress peak on the same vowel, thus giving these forms non-exceptional antepenultimate stress. Moreover, the addition of a second syllable (shown in the DEF.PL column) actually shifts the stress peak to the next vowel, keeping this form antepenultimate in stress. Example (ii) shows this more gradually: The INDEF.SG form is two syllables out of place, while the DEF.SG and INDEF.PL columns are one syllable out of place. Finally, the DEF.PL form shows regular stress, on the antepenult. Neither of these exception types, therefore, involves stress *earlier* than the antepenult. There is one other type of exception of a different type, exhibited by the verbal-adverb (= gerund) form only, as shown in (iii) and (iv):

(iii)	no.SEJ.ḱi	no.SEJ.ḱi *go*	no.SEJ.ḱi *mu ja*		
	'bringing'	'bringing him'	'bringing him it(fem)'		[≈ exx. 56a-c in Franks (1987: 128)]
(iv)	VI.vaj.ḱi	VI.vaj.ḱi *go*	so.op.ŠTU.vaj.ḱi	so.op.ŠTU.vaj.ḱi *mu*	
	'calling'	'calling him'	'announcing'	'notifying him'	[≈ exx. 57a-d in *Ibid.*]

As we show below, the addition of enclitics to non-finite verbal forms shifts stress rightward. Not so with this part of speech, which traditionally has fixed stress on the first syllable of the *-ejḱi/-ajḱi* suffix, as shown in (iii). Any added clitics fail to attract stress, even to the point of having pre-antepenultimate stress. In another style the verbal adverb has acquired regular antepenultimate preceding it (= prosodically hosted by the preceding word) or not. That is, if the clitic is one of only two syllables following the stress, then it is within the PrWd, as shown in (12) through (15):

(12) The yes/no-interrogative particle *li*: Not part of the PrWd — Type I
 a. SVE.kr.va.***li*** 'Is it (a) mother-in-law?' b. *sve.KR.va.***li***

(13) The definite article *-ta* (and other allomorphs): Part of PrWd — TypeII
 a. sve.KR.va.*ta* 'the mother-in-law' b. *SVE.kr.va.*ta*

(14) Possessive clitics (homophonous with DAT clitics): Part of PrWd — TypeIII
 a. sve.KR.va.*mi* 'my mother-in-law' b. *SVE.kr.va.*mi*

(15) Other clitics, incl. pronominals: Part of PrWd — TypeIV
 a. do.NE.si.*go* 'Bring it_{ACC}!' b. do.ne.SI.*mi.go* 'Bring me_{DAT} it_{ACC}!'

Thus, (12) differs from (13) through (15) in that *li* is not part of the so-called trisyllabic stress window. That is, in each of (13) through (15) the addition of a monosyllabic clitic shifts the stress by one syllable rightward, to the antepenult of the lexical word plus the enclitics.

This distinction has been missed in several descriptions of Macedonian enclitic prosody in the literature. For example, Elson (1976: 276) claims, "For Macedonian, there is nothing of significance to be gained from this [comparing the accentual properties of type-II and other, undisputed enclitics/BK&R] because non-enclitics, enclitics, and the forms of the article all behave the same way with regard to the assignment of antepenult stress ..."; Elson (1976) unfortunately fails to consider *li*, just clitics of types II through IV.[9]

(16)

		Verbal Clitics	Nominal Clitics
a.	No clitic:	DO.ne.si KNI.ga 'Bring (a) book!'	SVE.kr.va 'mother-in-law'
b.	Type-I clitic:	do.NE.se.te.***li*** ... 'Do you bring ...?'	SVE.kr.va.***li*** 'Is it a mother-in-law?'
c.	Type-II clitic:	N/A	sve.KR.va.*ta* 'the mother-in-law'
d.	Type-III clitic:	N/A	sve.KR.va.*mi* 'my mother-in-law'
e.	Type-IV clitic:	do.NE.si.*go* 'Bring it_{ACC}!'	N/A

[(16a, c, e) from Elson (1976); (16b, d) elicited/BK&R]

Clearly, (16b) shows that not all clitics affect stress.

---

speech, as in (iv). Note that in these examples additional syllables still do not shift stress rightward, again resulting in pre-antepenultimate stress, as shown in the rightmost example in (iv). The exceptions in (i) through (iv) are the only deviations from regular antepenultimate stress in Macedonian. We present only regular-antepenultimate-stress data below.

[9] Elson (1976) does, however, consider **Bulgarian** *li*, but ignores its Macedonian counterpart.

Franks (1989) likewise assesses the accentuation of Macedonian clitics but fails to consider *li* altogether. Moreover, other studies discuss the accentual properties of *li* in Macedonian without actually mentioning on which syllable the stress is pronounced (cf. Englund 1977: 119-120; Friedman 1993: 287); Tomić (1996a) initially discusses the similarity of type I with types II and III, but then discontinues consideration of type I (where, citing Halpern 1992 extensively, she discusses junctural phonology at the word-clitic boundary).

Another distinction among these four types is whether these clitics can be initial. By this criterion, types I through III are distinct from type IV. If the verb is finite, then type-IV clitics precede the verb and can then be clause-initial in Macedonian. Those in types I through III must always be non-initial in a particular syntactic domain.[10]

Type-II and -III clitics must be non-initial the **noun phrase**, as shown in examples (17) and (18). The use of type-III clitics in Macedonian is limited to kinship terms and lexified expressions, but their use is consistent.[11]

(17) u.BA.vi.*ot* MAŠ
[handsome+the]$_{M.SG}$ man
'the handsome man' (NB: -*ot* is an allomorph of -*ta* 'the')
[≈ Elson (1976: 279, n. 8), who uses phonetic transcription]

(18) MAJ.ka.*mu* STA.ra (**also:** sta.RA.*ta.mu* MAJ.ka)
mother+his old$_{F.SG}$ [old+the]$_{F.SG}$+his mother
'his old mother' (cf. *MAJ.ka STA.ra.*mu*, *STA.ra.*ta* MAJ.ka.*mu*)
[Friedman (1993: 286)]

Any other placement of these clitics—i.e., proclitic to any word or enclitic to any word but the first one—is ungrammatical.[12]

Unlike type-II and -III clitics, *li* (in type I) must follow a word of the **clause** (specifically, "clause" here is the complementizer phrase, not counting any adjuncts to it), as in (19):

(19) RAZ.bi.raš.***li*** TI ma.KE.don.ski
understand+Q you$_{NOM.SG}$ Macedonian
'Do you understand Macedonian?'
[Englund (1977: 93), stresses elicited/BK&R]

[10] Ćavar (1996a, 1996b) and Ćavar and Wilder (1994) distinguish between Wackernagel's (1892) Law and the so-called Tobler-Mussafia Effect. Wackernagel's Law requires particular constituents to be in second position, while the Tobler-Mussafia Effect requires merely that certain elements be non-initial, based on observations about Romance. Cf. Mussafia 1898; we've been unable to locate Adolf Tobler's ca. 1880 work. Ćavar and Wilder also discuss the notion of prosodic subcategorization in Zec and Inkelas (1990: 369): "[[ ]PrWd __]PrWd".

[11] We distinguish between type II and III for reasons that aren't crucial to the main text of this paper: First, the two are not in complementary distribution, as the alternative form in (18) shows. Next, these two types have differing segmental-junctural properties, discussed in Elson 1976, Sadock 1991 and Tomić 1996a. See also the next footnote.

[12] More specifically, type-II and -III clitics are encliticized to the first word in the nominal expression with nominal features (including adjectives and numerals). That is, an NP-initial adverb will be skipped over by the article and possessive clitics. Our description here is intended to be primarily descriptive.

(NB: in (19) *ti* is not a clitic, but rather the word-stressed, nominative-case pronoun.) We do not, however, assume that *li* is a clausal affix, merely prosodically adjoined to the clause's first word, as proposed for Russian *li* by Billings (1994, 1996b), for example. Instead, as shown in (10) and (11), *li* is in C; prosodic inversion takes place only if no element dominated by CP precedes *li.*

Type-IV clitics, unlike types I through III, are unique in being able to appear initially., as shown in (20a-c): [13]

(20) a. *mi* *go* <u>KA.ža.le</u> (three-syllable verb)
me$_{DAT}$ it$_{ACC}$ said$_{PL}$
'They said it to me.'

b. *mi* *go* <u>DA.le</u> (two-syllable verb)
me$_{DAT}$ it$_{ACC}$ gave$_{PL}$
'They gave it to me.'

c. *mi* *go* <u>DAL</u> (one-syllable verb)
me$_{DAT}$ it$_{ACC}$ gave$_{M.SG}$
'He gave it to me.'

In other words, there exists no non-initiality requirement on type-IV clitics in Macedonian.

Note also that whereas post-verbal type-IV clitics affect the place of stress on the verb, *mi* and *go* in (20) do not. This stress asymmetry has been observed in Romance languages.[14] Macedonian type-IV clitics, as in some of these Romance languages as well, precede only finite verbs.[15]

One (always) pre-verbal element, the negative particle *ne*, which in some cases looks like a clitic, is inherently accented in Macedonian (one of the main

---

[13] Examples (1a) and (5a) also show this point. See fn. 2 regarding the clitichood of *bi*.

[14] Cf. Peperkamp 1995 for a summary of the literature on the dual-position clitics in Romance. Specifically, Macedonian appears to be very similar to certain Lucanian dialects of Italian. As Peperkamp (1995: 122, citing Lüdtke 1979: 31 in part) reports, Lucanian words show stress on one of the last three syllables and the addition of enclitics and suffixes both regularize the stress of the suffixed and/or encliticized stem to penultimate stress. (Macedonian, as a preceding footnote mentions, has words with stresses in the same final-trisyllabic window, with the addition of suffixes and non-*li* enclitics regularizing the stress, in the case of Macedonian, to **ante**penultimate position.) Garde (1968: 32) and Kenstowicz (1991) also point out this Macedonian-Italian similarity.

[15] We have uncovered the following two exceptions, of the same kind, to this description:

(i) <u>ne</u> *ME* <u>da.vaj</u> <u>MAJ.ko</u> 'Don't give me (in marriage), mother!'
NEG me(acc) give(imperative) mother(vocative) [Lunt (1952: 22, fn. 1)]

(ii) Ne go gledaj! 'Don't look at him!'
NEG him(acc) look(imperative) [= ex. 28 in Alexander (1994: 10)]

Lunt implies that this order is marked (but shows ESD diacritics!). Alexander marks (ii) with a question mark and shows the preferred order, *Ne gledaj go!* [= her ex. 29], without marking stress in either example. These forms are significant from a prosodic standpoint, as Alexander points out, because those speakers who use(d) them appear to merely have a non-initiality requirement on clitics of non-finite verb forms—Wackernagel's (1892) Law in Alexander's terminology—because non-finite verbal forms are frequently the first lexically accented word of the clause . Those who don't use forms like (i) and (ii) appear to order their verbal clitics as in Romance, discussed in the preceding fn.

assertions of Garde 1968).[16] As (21) shows, whereas the type-IV clitics *mu* and *go* are inherently unaccented, *ne* is accented.

(21) NE *mi* *go* DA.le
NEG me$_{DAT}$ it$_{ACC}$ gave$_{PL}$
'they didn't give it to me'

Example (21) does not show clearly which lexical (underlined) word hosts the clitics. Two plausible prosodic organizations are shown in (22a-b):

(22) a. [NE] [*mi* *go* DA.le]
b. [NE *mi* *go*] [DA.le]

There are a few reasons to favor the bracketing in (22a): First of all, these clitics are syntactically ordered preceding the finite verb; the null hypothesis is that the prosody matches the syntax. Another reason to favor (22a) is the behavior of such clitics when the verb is clause-initial, as in (20a-c). If these clitics are hosted by any word prosodically, it must be by the verb, perhaps by so-called stray adjunction.

Another argument in favor of (22b) is shown in (23), in which there is no overt verb: [17]

(23) ka.mo *MI* *ti* *go*
where-to me$_{DAT}$ you$_{DAT.SG}$/your$_{SG}$ it$_{ACC}$
'where (should) I (put) it for you?' [Elson (1993: 157)]

In the case of (23), the only lexical word is the *wh*-interrogative *kamo*. Once there are three syllables worth of clitics, then the stress shifts off of the lexical word and on to the third (monosyllabic) clitic from the end. This suggests that the clitics in (21) are prosodically hosted by the **preceding** lexical word *ne*.

While the data in (20) and (23) suggest that clitics more naturally adjoin to a preceding prosodic host, these data are inconclusive. Example (23) also looks like stray adjunction, as in (20a-c). That is, when clitics have a lexical-word neighbor on only one side, then they will be hosted prosodically by that word. What we really need is an example like (21), but with at least **three** syllables worth of verbal clitics between the two lexical words, like (24): [18]

---

[16] Anderson (1996: 188–89) offers a different account for the unique properties of *ne* in Macedonian, considering it a clitic which is positioned differently from the other clitics.

[17] Ex. 19d in Franks (1989: 561) glosses this same example as 'Where did that thing of yours get to on me?' (Elson 1993, incidentally, was apparently written unaware of Franks 1989.) Both Elson and Franks apparently got this example from Koneski [1957: 123/]1967: 163, which doesn't have glosses, because it's written in Macedonian. This example, furthermore, does not appear in Lunt or Garde (which would have glosses, at least in French in the latter case). We have added the word-for-word glosses and part-of-speech labels, which shows the type-IV and -III interpretations of *ti*, respectively.

[18] Elson (1993: 152–53) reports the following pair of examples; see §4.2 regarding (ii).

(i) ne *SME* *mu* *go* ZE.le (ii) ne *sme mu GO* ze.le
NEG are$_{1.PL}$ him$_{DAT}$ it$_{ACC}$ took$_{PL}$
'We didn't take it from him.'

(24) NE *sme* *mu* *go* DA.le
NEG are$_{1.PL}$ him$_{DAT}$ it$_{ACC}$ given$_{PL}$
'We didn't give it to him.'

This example shows that three intervening clitics fail to draw stress off of *ne*, which indicates that the bracketing in (22a) is correct, with the proviso that only prosodic enclitics affect stress.

To summarize this subsection, Macedonian has clitics with four different kinds of properties: *li*, in type I, is always prosodically enclitic and never initial, but, unlike all other clitics, fails to affect stress. Like *li*, the definite article and possessive clitics (types II and III) are non-initial (within a noun phrase) and prosodically enclitic, but unlike *li*, always affect stress. Finally, the type-IV clitics precede finite verbs and follow non-finite ones and are always hosted by the verb.

4.1.2 The Bulgarian Clitic-Prosody System

In this subsection we show the aspects of the Bulgarian clitic system that differ form the corresponding Macedonian ones described in the preceding subsection. In more than one aspect we simply note that the Bulgarian system is different, without giving as detailed a description as we did of the Macedonian system. We defer instead to the very detailed account in Hauge 1976.[19]

In Bulgarian, stress in largely unpredictable. That is, the location of stress (or its pattern of assignment) is lexically encoded. Stress may fall on any syllable of a word, and is usually unaffected by the addition of clitics. (A partial exception is the definite article, which in some stress paradigms takes stress on itself:

(Stress notation modified; word-for-word glosses added; sentential gloss unchanged.) Elson says clearly that the three clitics are hosted prosodically by *ne*. Our informants rejected example (i) , preferring instead the order in (50a). Elson (1993: 157, n. 1) glosses and attributes this and other data in the article with the following acknowledgment: "Items cited for illustrative purposes **or their models**, are from Lunt 1952: 21–25, Koneski 1967: 139–210, or Garde 1968." Indeed, example (ii)—the acceptability of which we don't dispute—appears in Garde 1968: 31 and Lunt 1952: 23. The source of Elson's assumption, that the clitics are hosted by *ne*, seems to be Garde (1968: 36):

(iii) — / *NE* — *sme mu go* / *ZELE* — → — / *NE sme mu gó zele* — [sic.]

Garde's abstract notation implies that whereas on the left side of the arrow [= some sort of underlying representation prior to application of ESD] the (all-caps) lexically accented stems *ne* and *zele*, with the clitics hosted by *ne*, the right side of the arrow [= the combined underlying and surface representations of the ESD form] has only *ne* with accent and everything hosted prosodically by it, with actual stress indicated by the acute accent on *gó*. Elson (1993: 152–53), while arguing against much of Garde's proposal, appears to espouse the left-hand side of (iii), assuming the right side of (iii) and the form in (i) to be attested. Our data on *li*-insertion below (in §4.2), specifically in (35a), show additional evidence that the clitics are hosted prosodically by the verb. Still, Macedonian—and Balkan in general—being a very diverse dialectal situation, we leave open the possibility that (iii) is attested for some speaker somewhere.

[19] Anderson (1996: 188) incorrectly characterizes the ordering of type-IV languages in Bulgarian (and Macedonian) "follow gerunds, infinitives and imperatives." (i) neither Bulgarian nor Macedonian has an infinitive; (ii) as both Hauge (1976: 5) and Alexander (1994: 9) point out, such clitics precede non-initial imperative verbs as well in Bulgarian.

(25) a. GRAD 'city' b. gra.d*ÂT* 'the city'

(26) a. PA.met 'memory' b. pa.met.*TA* 'the memory'

(27) a. SE.dem 'seven' b. se.dem.*TE* 'the seven'

[≈ exx. in Elson (1976: 276)]

The addition of other clitics does not shift stress.

Bulgarian has all four types described in the preceding subsection. Type I (*li*) is used far more frequently than in Macedonian, as we discuss (in §5) below. Type II has roughly the same distribution, with the stress peculiarity shown in (25) through (27), while type III is far more frequently used in Bulgarian than in Macedonian. Types I through III have the same prosodic properties in both languages—all required to be non-initial (*li*, in the non-adjoined-to CP; types II and III, in the NP).[20]

Two major distinctions of Bulgarian we discuss below in separate subsections: Type-IV clitics in Bulgarian are, like *li* is in both languages, required to be non-initial in the clause (§4.4) and the negative element *ne* idiosyncratically stresses the following constituent (§4.5), even a clitic. Both of these phenomena drastically obscure the syntax and relatively simple prosody of *li* in Bulgarian. Before Proceeding to those two phenomena, however, we return to two distinctions in Macedonian.

4.2 Enlarged Stress Domain in Macedonian

We show in above (in §4.1.1) that *ne* in Macedonian is inherently accented (unlike *ne* in Bulgarian; cf. §4.5 below). It is also possible, however, for *ne* and the following verb to share a single PrWd stress; this possibility of stressing both lexical words as one PrWd accounts for the acceptability of both (3a) and (3c) above, and results in contrasts like (28):

(28) a. NE *li* SA.kaš *da* O.diš? (Macedonian)
NEG Q want$_{2.SG}$ to go$_{2.SG}$
'Don't you really want to go?' (or 'Do you really not want ... ?')

b. NE sa.kaš *li da* O.diš? (Macedonian)
NEG want$_{2.SG}$ Q to go$_{2.SG}$
'Don't you want to go?'

In (28a) the speaker is sure of a confirmative answer; the *ne* is being questioned. In (28b), however, the speaker is unsure whether you want to go or not; *sakaš* is being questioned. In Bulgarian, only *Ne iskaš li da otideš?*—the order corresponding to (28b)—is possible.

[20] Cf. other works, however, which discuss interesting interactions of these non-initiality domains. Halpern (1992/1995: 227–31), devotes an appendix to overlapping domains, citing Ewen (1979) and Hauge (1976) with Bulgarian examples. Halpern discusses what happens when the leading edge of a clause coincides with an NP's leading edge. He claims, contrary to our proposals here, that type-II and -III clitics must be peninitial in that NP, while type-I clitics must be clause-peninitial. Halpern also discusses briefly the apparent ability of possessive (type-III) clitics to appear out of their NP, within the cluster of type-IV clitics. Note also that within the nominal expression the article precedes the possessive clitics, as the alternative form in (18) shows.

This possibility of stressing two lexically accented words as a single PrWd has been called "enlarged stress domain" (Franks 1987, also referred to as "collocational stress" in Elson 1993 and "accentual units" in Alexander 1994), as shown in (29) through (32). In the (a) examples each (underlined) lexical word receives the predictable stress (antepenult if at least trisyllabic; otherwise, initial). In the (b) examples, however, the entire two-word domain receives a single PrWd stress, on the antepenult. Aside from adjective + noun, shown in (29), the other word pairs reported in the literature are preposition + noun, *wh*-interrogative + verb, and negation + verb, shown in (30) through (32), respectively. These four environments are elaborated as well in Elson (1993) and Franks (1989).[21]

		Without ESD		With ESD
(29)	a.	LE.va.*ta* NO.ga left+the foot 'the left foot'	b.	le.va.*TA* no.ga left+the foot 'the left foot'
(30)	a.	O.ko.lu TR.lo near sheep-pen	b.	o.ko.LU tr.lo near sheep-pen
(31)	a.	KOJ RE.če who$_{NOM}$ said$_{3.SG}$	b.	KOJ re.če who$_{NOM}$ said$_{3.SG}$
(32)	a.	NE *mi* *go* DA.le NEG me$_{DAT}$ it$_{ACC}$ gave$_{PL}$ 'they didn't give me it'	b.	ne *mi* *GO* da.le NEG me$_{DAT}$ it$_{ACC}$ gave$_{PL}$ 'they didn't give me it'

Additionally, as (29) and (32) show, clitics can appear between the lexical words. In (29) the definite article *-ta* is enclitic to the first word of the noun phrase; in (32) the clitics *mi* and *go* are syntactically ordered before the finite verb. The environment in (31) also allows medial clitics.[22]

As we show above (at the end of §4.1.1), the verb is the **prosodic** host of the clitics *mu* and *go* in (32a). We repeat example (24) as (33a), adding its ESD

---

[21] The ESD forms in (9b) and (10b) are now quite marked in Contemporary Standard Macedonian, considered as either archaic or dialectal. We cite them just once, nonetheless to report the extent of ESD in the language. Alexander (1994: 11) also lists the following ESD example. Cf. Lunt (1952: 24–25) for further discussion.

(i) dobre *TE* najdov
well you(acc) found(1.sg)
'Welcome!' (literally: 'I found you well')

Alexander (and we) cannot explain this ESD form, a fixed expression, in syntactic terms.

[22] Two peculiarities of ESD occur when the latter lexical stem is monosyllabic: The first—which Franks (1989) calls the "monosyllabic-head effect"—prevents the stress from preceding the beginning of the second stem by more than one syllable, as shown in (i):

(i) ne *sum mu GO* zel — 'I didn't take it from him.' [Lunt (1952: 23)]
NEG am him it took — (cf. *ne *sum MU go* zel)

(ii) ne *BI* dal **or** NE *bi* dal — '(He) should not have given ....'
NEG should gave NEG should gave — [= ex. 18a in Franks (1989: 559)]

The second peculiarity is an exception to Franks's monosyllabic-head effect just in case the entire ESDomain consists of exactly three syllables, as shown in (ii). Examples (i) and (ii) also have one non-ESD variant each: NE *sum mu go* ZEL, NE *bi* DAL. Kepeski and Pogačnik (1968) provide an opportune pair, similar to (ii): doBAR den 'good day', DObra nok 'good night'.

counterpart in (33b). The bracketings in (34a-b) represent the prosodic organizations of (32a-b). We avoid the debate in the ESD literature about whether one of the two lexical words becomes a clitic.

Without ESD — With ESD

(33) a. NE *sme mu go* DA.le — b. ne *sme mu GO* da.le
NEG are$_{1.PL}$ him$_{DAT}$ it$_{ACC}$ gave$_{PL}$
'We didn't give it to him.'

(34) a. [NE] [*sme mu go* DA.le] — b. [ne *sme mu GO* da.le]

In addition to the three-syllables-worth-of-clitics test in (24), the bracketing in (34a) can be tested by turning the clause into a yes/no question using *li*, as shown in (35) and (36): [23]

(35) a. NE ***li*** *sme mu go* DA.le — b. *ne *sme mu GO* ***li*** da.le
NEG Q are$_{1.PL}$ him$_{DAT}$ it$_{ACC}$ gave$_{PL}$

(36) b. *NE *sme mu go* ***li*** DA.le — b. ne *sme mu GO* da.le ***li***
'Didn't we give it to him?'

The grammatical positions of *li*, depending on whether there is ESD, are in (35a) and (36b). There are three plausible hypotheses about the placement of *li* worth considering—assuming a prosodic-inversion account such as Halpern's (1992/1995)—shown in (37a-c):

(37) a. Hypothesis A: *li* follows first PrWd **stress** [Wackernagel 1892].
b. Hypothesis B: *li* follows first PrWd **domain**,
assuming the structure in (22b) above: [NE *mi go*] [DA.le].
c. Hypothesis C: *li* follows first PrWd **domain**,
assuming the structure in (22a) above: [NE] [*mi go* DA.le].

Hypothesis A predicts that *li* will appear at the first available syntactic boundary following the first word stress; it correctly predicts the form in (35a), but incorrectly predicts (35b). Hypothesis B correctly predicts the form in (36b), but incorrectly predicts (36a). Hypothesis C is the only one to correctly predict both of the attested forms: (35a) and (36b). The clitic *li*, therefore, corroborates the correctness of our analysis above in (22a), as shown first by the datum in (33a).

(38) a. ne *sum ti go* DA.la ***li***? (Macedonian, with ESD)
NEG am$_{1.SG}$ you$_{DAT.SG}$ it$_{ACC}$ given$_{F.SG}$ Q
b. ne ***li*** *sum ti go* DA.la? (Macedonian, without ESD)
c. ne *SÂM* ***li*** *ti go* DA.la? (Bulgarian)
'Haven't I given it to you?'

---

[23] We discuss *neli* vs *ne li* in fn. 1 above. Cf. also the following ungrammatical forms:
(i) NE *mi go* DA.le ***li*** (ii) *ne ***li*** *mi GO* dale
In (i) *li* follows two stresses, while in (ii) *li* follows no stresses—both illicit inversion.

The only order allowed in Bulgarian is shown in (38c).

A similar example, from a folk song, is (39), in which the speaker believes the addressee **should** be ashamed.

(39) NE *li ti* TE.be SRA.mo.ta? (Macedonian)
NEG Q you$_{DAT.SG}$ you$_{DAT.SG}$ ashamed
'Aren't you ashamed?' [from *Acana mlado nevesto*]

The corresponding Bulgarian question, *Ne te **li** e sram?*, must place *li* after *te*, not right after *ne* (for reasons we present in §4.5 below). This example does not have to do with ESD, since no such option exists here (i.e., no verb after *ne* + clitics).

In (31b) above we show that ESD can take place between an initial *wh*-interrogative stem and the verb. As (the title of) Rudin 1992 shows for Bulgarian, *li* can appear in *wh*-interrogative clauses as well, lending an emphatic 'on earth' or 'the hell' meaning to the question. This is true as well for Macedonian. It should thus be possible for *li* to appear in a clause like (31b). Indeed, we've found only the following example, in (40a):

(40) a. Koj li ḱe bide toj? —... (Macedonian)
who Q MOD is$_{3.SG}$ that$_{M.SG}$/he
b. KOJ *li ḱe* BI.de TOJ
'Who on earth would that/he be?'

In light of the widely accepted model that *wh*-interrogatives like *koj* 'who' are in SpecCP and *li* is in C, which we adopt here, then it should follow that there cannot be an ESD counterpart of this example (i.e., **Koj ḱe bide li toj?*), since *li* already follows some stressed constituent, *koj*, at S-structure. Whereas *ne* + verb constructions (with intervening clitics) have both ESD and non-ESD variants (with *li* ordered accordingly), *wh* + verb constructions are attested in only this example in Englund's *li* corpus of Mac, which our informants pronounced only **without** ESD.[24]

As Elson (1993: 158, n. 4) points out, while all other ESD pairs in (29) through (32) constitute a syntactic constituent of some sort, *wh* + verb seems not to be a constituent. That problem aside, there is a functional problem: If *wh* + verb ESD means **de**-emphasis of some sort on the *wh* element, and *wh* + *li* represents emphasis on the *wh* element, then imposing the non-ESD onto such a clause de-emphasizes the *wh* element, contrary to *li*'s purpose in a *wh* clause. The hypothetical ESD version should be bizarre because question words are always focused and the remainder of the *wh* question clause is presupposed (i.e., *What did you see?* presupposes that *you* saw something; *Who is he?* presupposes that *he* is someone). This is the focus-presupposition structure characteristic of the XP-*li* construction in (10); it would be pragmatically weird at best to have a *wh* question with V-*li* structure in (11)!

To summarize this subsection, we have shown that a stress domain in Macedonian is enlarged, allowing certain pairs of lexically accented words, along with any clitics between them, to be stressed as one PrWd. The addition of *li* to

[24] One informant rejected this example, but appears to reject all *wh* + *li* questions.

ESD and non-ESD environments disproves one aspect of the now famous Wackernagel's Law (Wackernagel 1892), which describes peninitial clisis as following the first stress **peak**. In fact, as these data show, such clitics follow the first stress **domain**, passing up the chance to follow immediately after the stress. We also discuss the interaction of *li* in *w h* questions with ESD. This unique phenomenon in Macedonian also explains one constituent order difference with Bulgarian.

4.3 Optional Auxiliary-Accent in Macedonian

Clitic forms of 'be' in Macedonian appear to be accented, at least in some contexts, while the corresponding forms in Bulgarian are obligatorily stressed after *ne* (cf. §4.5 below) and obligatorily unstressed elsewhere.

Englund (1977: 111) reports example (41a), quoting Živko Čingo's *Paskvelija*, without stress in the standard orthography, but with *ti* alone not in italics. We have merely transliterated the example:

(41)	a.	—	*Si*	*li*	ti	*član*	*na*	*mladinata?*	(Macedonian)
			are$_{2.SG}$	Q	you$_{NOM.SG}$	member	of	youth+the	
	b.		*SI*	***li***	TI	ČLEN	*na*	mla.DI.na.*ta*	(Macedonian)
			'Are you a member of the youth/young-people?'						

(Note again, as in (19) above, that *ti* here is the lexically accented NOM-case form, homonymous with the unaccented DAT clitic.) Responses varied when we elicited stress from informants: Those who did accept (41a) without comment invariably stressed it as shown in (41b). This is predicted if one assumes that *li* must follow the first stressed word in structures **without** a focused XP in SpecCP, as shown in the tree in (10).[25]

This example is especially interesting in light of the head-movement account we adopt (in §3) above. It shows what happens when the tensed verb is a clitic auxiliary, which moves to C (in order to check *li*'s focus feature) and there is no other verbal stem to host *li*. For those speakers who accept (41), the best means of keeping *li* from being initial seems to be a last-resort strategy of stressing the clitic auxiliary.[26]

Those who rejected this order suggested (42a-b) instead:

(42) a. DA.li *si* ČLEN *na* mla.DI.na.*ta* (Macedonian)

b. TI ***li*** *si* ČLEN *na* mla.DI.na.*ta* (Macedonian)

Using *li*'s stressed allomorph *dali*, in (41c), is discussed (in §5) below.

[25] The stress on *ti* in (41b) is not very pronounced phonetically. Some authors (cf. Hauge 1976) list NOM-case pronouns as separate PrWds but don't mark stress on them. Lunt (1952) doesn't however indicate word stress on such pronouns.

[26] Examples similar to the preceding Macedonian ones are also attested in Bulgarian. Hauge (1976: 2-3) lists examples in which clitic forms of 'be' are stressed, accounting for most of these by either ellipsis or displacement of the would-be host of the clitic, leaving it stranded in clause-final position. Hauge has no explanation, however, for one example with prosodic and syntactic structure and word order seemingly identical to that of (41).

It is not clear, however, what the structure of (41d) is. Two possible S-structures—**prior** to any prosodic inversion—are shown in (43a-c):

(43) a. [$_{CP}$ [$_{SpecCP}$ ti] [$_{C}$ ***li***] [$_{IP}$ [$_{I}$ *si*] … [$_{PP}$ *na* mla.di.na.*ta*]$_{PP}$]$_{IP}$]$_{CP}$

b. [$_{CP}$ [$_{C}$ ***li***] [$_{IP}$ [$_{SpecIP}$ ti] [$_{I}$ *si*] … [$_{PP}$ *na* mla.di.na.*ta*]$_{PP}$]$_{IP}$]$_{CP}$

In (43a) *ti* is focused (i.e., moved to SpecCP), as diagrammed arboreally in (10) above. This structure does not require prosodic inversion and merely stresses the correct syllable in each PrWd, as shown in (42b). The structure in (43b) shows *ti* in SpecIP. This structure merely requires *li* to invert past the first stressed word, *ti,* in order to keep *li* prosodically enclitic to some word in this clause.[27] The choice between the structures in (43a-b) might be distinguished phonetically by differing accentuation on *ti.* We have not conducted such tests and merely present both.

Interestingly, Macedonian has lexically accented 'be' forms, as does Bulgarian. These prosodically heavy forms, formed from the stem /bid-/ (/bâd-/ in Bulgarian). These accented stems, as is apparent from Kramer 1993, are used only as auxiliaries, never as copulas.[28] Korubin (1974: 247–48) also points this distinction, supplying additional examples corresponding to (41) and (42).[29] Thus, these accented stems are not available as ways of making these clitics accented.

Before leaving the issue of copula-stressing, we have one example which appears to involve the interaction of copula-stressing and ESD (cf. in §4.2 above):

(44) a.	…,	ama	*ne*	*e*	*li*	toa	*otvoren* […] *?*
		but	NEG	is$_{3.SG}$	Q	it$_{F.SG}$/she$_{SG}$	closed
	'…, but isn't it/she closed […]?'						
b.		ama	NE	*e*	***li***	TO.a	*OT.vo.ren*

All our informants accepted (44a), a transliteration of Englund's example (1977: 115, quoting Taško Georgievski's *Zmiski vetar*), each of them supplying the stresses in (44b), which is, at first glance, problematic for the syntactic and prosodic-inversion accounts we adopt here. The problem is that if *ne* is inherently accented, then it should bear stress, with *e* encliticizing to it prosodically. Generally speaking, when *li* prosodically inverts, it appears between the first stressed word and any of its enclitics. This pattern would result in the order **ama ne* ***li*** *e …* We suggest instead that ESD, between *ne* and the now-accented *e,* may be involved. If so, then *li* inverts to the attested place, after *ne.*

[27] The following structure might also be the S-structure of (42b):

(i) [$_{CP}$ [$_{C}$ ***li***] [$_{IP}$ [$_{I}$ *si*] … ti … [$_{PP}$ *na* mla.di.na.*ta*]$_{PP}$]$_{IP}$]$_{CP}$

In such a structure *li* would have to invert prosodically past the first stressed word, *ti,* resulting in the form **si* TI *li* ČLEN *na* mla.DI.na.*ta;* this leaves *si* in clause initial position. In our discussion of Macedonian type-IV clitics above in §4.1.1—cf. especially exx. (20), (21), (23), and surrounding discussion and footnotes—we show that type-IV clitics can be initial only if proclitic to a (lexically accented) finite verb form. When no such finite-verb host appears in the clause, as in (23), then these clitics appear to be enclitic to the first PrWd. If this is the case, then the structure in (i) is still plausible; *si* inverts as well.

[28] Hauge (1976: 16, 36–44) makes explicit the auxiliary/copula distinction in Bulgarian.

[29] Korubin (1974: 246) also shows that auxiliary clitics lack third-person forms.

To summarize this subsection, we have shown that clitic copulas can be accented under conditions that are still not clear to us. We also show one potential interaction of this phenomenon with ESD. In future work we hope to pursue the conditions under which such forms are possible, as well as how (if at all) their Bulgarian counterparts differ.

4.4 Inversion of Verbal Clitics in Bulgarian

Macedonian type-IV clitics precede the verb only when it finite (and are both prosodically and syntactically proclitic to it). Bulgarian type-IV clitics are also essentially syntactically proclitic to the verb. Unlike those of Macedonian, however, clitic pronominals and auxiliaries (not all type-IV clitics!) are restricted from being clause initial. In addition to *li* in both languages, type-IV clitics in Bulgarian are prohibited from being initial.[30] If no accented word appears in front of these clitics, prosodic inversion takes place (following Halpern 1992/1995). The effect of this process is clearly seen in the difference between the non-*li* examples in (45) through (48), as well as in the *li* questions in (1) and (2) above.

(45) *me* BO.li U.vo.*to* (Macedonian)
me$_{ACC}$ hurts$_{3.SG}$ ear+the
'My ear hurts.'

(46) a. bo.LI *me* u.XO.*to* (Bulgarian)
hurts$_{3.SG}$ me$_{ACC}$ ear+the
'My ear hurts.'

b. u.XO.*to me* bo.LI (Bulgarian)
ear+the me$_{ACC}$ hurts$_{3.SG}$
'My **ear** hurts.'

c. *me bo.LI u.XO.*to* (Bulgarian)

(47) *ti* *ja* DA.dov (Macedonian)
you$_{DAT.SG}$ it$_{ACC}$ gave$_{1.SG}$
'I gave it to you.'

(48) a. DA.dox *ti* *ja* (Bulgarian)
gave$_{1.SG}$ you$_{DAT.SG}$ it$_{ACC}$
'I gave it to you.'

b. AZ *ti* *ja* DA.dox (Bulgarian)
I$_{NOM}$ you$_{DAT.SG}$ it$_{ACC}$ gave$_{1.SG}$
'I gave it to you.'

c. **ti ja* DA.dox (Bulgarian)

[30] Unlike *li* (in both languages), Bulgarian type-IV clitics must be non-initial in a slightly different domain. As we show in §4.1 above, *li* inverts if no other element dominated by CP appears in front of *li*. That is, adjuncts to CP do not count. Bulgarian type-IV clitics, however, can make use of adjuncts to CP, coordinating conjunctions and other material not dominated by the CP node. As Hauge (1976: 5) points out, however, type-IV clitics in Bulgarian are also prohibited from following a clause-internal pause. Thus, it seems that the crucial non-initiality domain for these is some sort of phonological phrase.

Clitic inversion is relevant to the syntax of *li* questions because such clitics raise to C along with the verb. We assume (following Rudin 1996) that clitics are functional heads which incorporate V. In Macedonian, *li* is straightforwardly suffixed to the verbal complex (i.e., the complex prosodic word consisting of the verb and its preceding type-IV clitics), as in (49); see also (1), (3) and (5) above.

(49) a. [*gi* NAJ.de] ***li*** PA.ri.*te* (Macedonian)
them$_{ACC}$ found$_{3.SG}$ Q money+the
'Did he find the money?'

b. [*ste* *go* GLE.da.le] ***li*** O.voj FILM (Macedonian)
are$_{2.PL}$ it$_{ACC}$ seen$_{PL}$ Q this$_{M.SG}$ film
'Have you seen this film?'

In (49b), for example, the verbal complex is *ste go gledale*. Right adjunction of the verb to *li* results in the S-structure string *li* + *ste go gledale* in C; prosodic inversion then produces the surface order.

In Bulgarian, exactly the same process occurs if a topic phrase (or any other material not belonging to a preceding clause) precedes C. In (50), the verb complex *ste go gledali* can appear pre-verbally

(50) VI.e [*ste* *go* GLE.da.li] TO.zi FILM (Bulgarian)
you$_{NOM.PL}$ are$_{2.PL}$ it$_{ACC}$ seen$_{PL}$ this$_{M.SG}$ film
'You seen this film.'

However, when no topic or other material precedes C, rendering the complex verb sentence-initial, then clitic inversion is required to provide a prosodic host for the clitics. In (51b)—the Bulgarian counterpart of (49a)—the verbal complex *ste go gledali* is adjoined to C, resulting in the string *li* + *ste go gledali.* Simple prosodic inversion of *li* here would produce a sentence with initial clitics, which is not possible in Bulgarian. Clitic inversion is required, resulting in the surface order *gledali ste go.* In both Bulgarian and Macedonian, when a complex V raises, *li* cliticizes to the first stressed element of the V; that is, prosodic inversion of *li* is to the end of the first stressed phonological word to its right. In the examples in (49) the only stressed element is the verb, so *li* follows it, in both languages; the result for Bulgarian is (51):

(51) a. na.ME.ri ***li*** *gi*? (Bulgarian)
found$_{3.SG}$ Q them$_{ACC}$
'Did he find them?'

b. GLE.da.li ***li*** *ste* *go*? (Bulgarian)
seen$_{PL}$ Q are$_{2.PL}$ it$_{ACC}$
'Have you seen it?'

In constructions which involve a stressed auxiliary instead of clitic forms of 'be', *li* follows the auxiliary. These auxiliaries include *ima-* 'have' in Macedonian, the past tense forms of the 'be' auxiliary in Bulgarian, and the negative-future particle *nema-/njama-* 'won't' :

(52) a. I.ma *li* DOJ.de.no *kaj* NAS? (Macedonian)
has$_{3.SG}$ Q come to us
'Has (s)he been to our place?'

b. BE.še *li* do.ŠLA? (Bulgarian)
was$_{3.SG}$ Q come$_{F.SG}$
'Had she arrived?'

(53) a. NE.ma *li da* VR.ne? (Macedonian)
won't$_{3.SG}$ Q to rain$_{3.SG}$

b. NJA.ma *li da* va.LI? (Bulgarian)
won't$_{3.SG}$ Q to rain$_{3.SG}$
'Isn't it going to rain?'

As we mention at the beginning of this subsection, not all type-IV clitics in Bulgarian are restricted from being clause-initial. In both languages the future particle *ḱe/šte* is an unstressable clitic, but in Bulgarian it differs from the other type-IV clitics in being able to begin a sentence; this then shields the other (non-*li*) clitics from being clause-initial:

(54) a. *ḱe go* ZA.vr.šat *li* (Macedonian)
MOD it$_{ACC}$ finish$_{3.PL}$ Q

b. *šte go* SVÂR.šat *li* (Bulgarian)
MOD it$_{ACC}$ finish$_{3.PL}$ Q
'Will they finish it?'

On the other hand, the conditional stem /bi-/ is lexically stressed in Bulgarian but not in Macedonian, leading to the contrast shown above in (5) and (6). Whereas *bi* appears clause-initially in both (5a) and (6b), this element is distinct in the two languages. The irrealis element in Bulgarian is conjugated, showing person- and number-agreement and each form is stressed; the form *bi* in (6b) is the 3.SG (homophonous with the 2.SG) form (cf. Hauge 1976: 36). In Macedonian *bi* is an invariant clitic, behaving like any other type-IV clitic in that language.

To summarize this subsection, we have shown that Bulgarian, unlike Macedonian, has a non-initiality requirement on its clitic pronominals and auxiliaries (a subgroup of type IV). The modal *ḱe* is not, however, subject to non-initiality, and the conditional stem /bi-/ is not a clitic at all. These differences cause many clitic sequences to differ between the languages.

4.5 Negative Stress Shift

The second prosodic peculiarity of Bulgarian we will discuss has to do with the unique properties of the negative particle *ne*. In Bulgarian *ne* carries an inherent stress, which is realized on the **following** syntactic constituent, even if that following constituent happens to be a normally-unstressable clitic (see Halpern 1992/1995, Scatton 1984). The relevance of this fact for *li* questions should be clear at this point, given the role of stress in determining the position of *li*. (Izvorski 1994; Izvorski, King, and Rudin 1996; King 1993/1994/1995; Rudin 1990–91, 1992, 1993a, 1993b, 1996; Rudin, King, and Izvorski 1995; and Rudin and Kramer 1994 have recognized this.) Since *li* cliticizes to the first stressed element of the verb word, *li* in Bulgarian is automatically placed after the constituent immediately following *ne*. In Macedonian, however, *ne* does not induce stress on the following word, and thus doesn't influence the placement of *li*. So, when *li* cliticizes to the first stressed element of the verb word (as in Bulgarian), the result is different: It ends up encliticized to the verb (in ESDs) or to *ne*, since the clitics are not stressed.

We have already seen examples of this in (3) and (4) above; another set is given in (55):

(55) a. <u>ne</u> *mu* *GO* <u>da.de</u> ***li*** (Macedonian, with ESD)
NEG him_{DAT} it_{ACC} $\text{gave}_{3.SG}$ Q

b. ne *MU* ***li*** *go* <u>da.DE</u> (Bulgarian)
NEG him_{DAT} Q it_{ACC} $\text{gave}_{3.SG}$
'Didn't (s)he give it to him?'

In (55b) *mu* is stressed in Bulgarian, as is *go* in (4b), because of the preceding *ne*, and *li* therefore must follow them. This does not happen in Macedonian.[31] If no type-IV clitic is present, then the Bulgarian order is *ne* + verb + *li*, which appears like Macedonian ESD, but is instead the result of *ne* not having stress and (vacuously) stressing the following accented verb stem.

Thus, constructions with *li* and *ne* are complicated by two independent phenomena, causing Bulgarian and Macedonian to diverge markedly. ESD causes Macedonian *li* to appear after the second inherently accented word, while negative stress shift causes Bulgarian *li* to follow up to one clitic after *ne*.

To summarize this section, then:

- The two languages differ significantly in their word prosody. Macedonian has antepenultimate word stress, while Bulgarian has lexically-encoded stress location (§4.1).
- Macedonian allows certain two-word combinations, causing *li* to follow what appears to be two words in that language (§4.2).
- Macedonian also apparently allows certain copulas to be stressed, likewise affecting the position of *li* (§4.3).
- Bulgarian requires clitic auxiliaries and pronominals to be non-initial in the clause. Additionally, the conditional head in Bulgarian, unlike in Macedonian, is not a clitic. This does not affect the placement of *li* as such, but greatly confuses comparisons with Macedonian (§4.4).
- Bulgarian *ne*, the Neg head, has a special property of stressing the following element. Thus, when *ne* follows *li* at the beginning of the S-structure order, *li* must invert prosodically past not just *ne* but the following, stressed element as well (§4.5).

These prosodic differences cause Macedonian and Bulgarian to appear very divergent in their placement of *li*. In fact, the syntactic structure of the two languages, as shown in the preceding section, is quite uniform.

[31] Englund (1977: 112), quoting Jovan Boškovski's *Izbor*, does list the following example: "*Ne ti li se čini* deka pticana kako da ni potskažuva za zaludnosta na ova naše metkanje po ulicive?" All of our informants rejected this order, putting *li* instead after *ne* (non-ESD) or after the verb (ESD).

5. Usage

Macedonian yes/no questions differ from those of Bulgarian in another way as well, this time not prosodic, but pragmatic.[32] Unlike in Bulgarian, Macedonian *li* is optional; it alternates apparently rather freely with Ø as well as with the non-clitic yes/no-interrogative complementizer *dali*. Englund (1977) reports that in her literary corpus 60 percent of all yes/no questions in Bulgarian contained *li*, the remainder mostly formed with other question words, such as *nali* 'isn't it'. In Macedonian, only 30 percent of yes/no questions had *li*, and 44 percent (almost half!) had no question word; cf. Friedman 1993: 286–287 and Kramer 1986: 130–50.

This difference in the usage of *li* is borne out by a survey of questions in Kramer 1985, a phrase book of Macedonian. Out of 101 questions that would take *li* in Bulgarian, 52 have *li*, 29 have *dali*, and 20 have no overt question word. The three types of questions are apparently synonymous; when visiting the auto mechanic, for example, the tourist is advised to ask the three questions in (56):

(56) a. Imate maslo za avtomobil? **Ø** (Macedonian)
have$_{2.PL}$ oil for automobile
'Do you have oil for cars?'

b. Imate **li** auspuh? ***li*** (Macedonian)
have$_{2.PL}$ Q muffler
'Do you have a muffler?'

c. Dali imate svećkicki? ***dali*** (Macedonian)
Q have spark-plugs
'Do you have spark plugs?'

Only the second of these corresponds to a normal question in Bulgarian; a question formed with intonation alone, like (56a) is marginal if possible at all in Bulgarian, while *dali* in a main clause signals a rhetorical question in Bulgarian: 'I wonder if you might have spark plugs.'

The optionality of *li* does not bear directly on the syntax of *li* questions, but does affect judgments. Speakers may reject an example with *li* not because the syntax is wrong, but because they prefer *dali* or Ø. The reasons for this may be pragmatic or simply personal preference. One speaker we consulted, an 18-year-old woman, showed an especially strong preference for Ø questions, accepting questions with *li* only reluctantly. However, when forced to use *li* she had very clear intuitions about where in the sentence it could and could not go. Another speaker frequently commented that an intonation question would be more usual in everyday speech than one with *li*. This preference certainly leads to a difference in usage and the frequency of *li* questions in the Bulgarian and Macedonian, but it does not seem to be connected to any difference in the syntax of *li* itself.

6. Conclusion

Our analysis captures both the essential similarity between the grammars of two closely related languages, and the striking differences between them. As one

[32] Restan 1972 also discusses differences in *li* usage between the two languages.

might expect, the "deeper" syntax of the two languages is identical; the differences are due to the interaction of the syntax of *li* questions with a series of relatively "superficial" factors: differing prosodic constraints on clitics, idiosyncratic differences in the stress properties of particular lexical items, and differences in usage. This satisfying result underlines once again the utility of a parametric approach to the grammars of related languages.

References

Alexander, Ronelle. (1994) "The Balkanization of Wackernagel's law". *Indiana Slavic studies* 7: 1–8. (G. Fowler, H. R. Cooper, Jr., and J. Z. Ludwig, eds. *Proceedings of the 9th Biennial Conference on Balkan and South Slavic Linguistics, Literature, and Folklore*. Bloomington, IN: Indiana University, Dept. of Slavic Languages.)

Anderson, Stephen R. (1996) "How to put your clitics in their place, or why the best account of second-position phenomena may be something like the optimal one." *The linguistic review* 13(3–4): 165–91. (Special issue: M. Nespor, ed. *Phonology between words and phrases*.)

Avgustinova, Tania. (1994) "On Bulgarian verb clitics." *Journal of Slavic linguistics* 2(1): 29–47.

Avgustinova, Tania and Karel Oliva. (1991) *The structure of Bulgarian verb complex*. Saarbrücken: Universität des Saarlandes, Computerlinguistik. (*CLAUS* [Computerlinguistik an der Universität des Saarlandes], 4.)

Baerman, Matthew. (1996) "The emergence of prosodic constraints in Macedonian." Presented at Workshop on conflicting constraints, Groningen,.

Billings, Loren A. (1994) "An Optimality approach to the syntax and prosody of discourse particles." Presented at American Association of Teachers of Slavic and East European Languages conference, San Diego.

Billings, Loren A. (1996a) "Macedonian enlarged stress domain and peninitial clisis". Presented at Cornell University, September.

Billings, Loren A. (1996b) "Conflicting prosodic and syntactic constraints on special clitics". *ZAS papers in linguistics* 6: 43–50. (J. Kornfilt, ed. *Proceedings of the Workshop on the phonology, morphology and syntax of clitics*.)

Billings, Loren, Christina Kramer, and Catherine Rudin (1996) "The unique prosodic properties of *li* in Macedonian." To be presented at American Association of Teachers of Slavic and East European Languages conference, Washington, December.

Ćavar, Damir. (1996a) "On cliticization in Croatian: Syntax or prosody?" *ZAS papers in linguistics* 6: 51–65. (J. Kornfilt, ed. *Proceedings of the Workshop on the phonology, morphology and syntax of clitics*.)

Ćavar, Damir. (1996b) "Interface phenomena in Croatian and Polish" Presented at Universität Leipzig, November.

Ćavar, Damir and Chris Wilder. (1994) " 'Clitic third' in Croatian". *Linguistics in Potsdam* 1: 25–63.

Comrie, Bernard. (1976) "Irregular stress in Polish and Macedonian". *International review of Slavic linguistics* 1(2–3): 227–40.

Deevy, Patricia L. (1995) "An Optimality theoretic analysis of stress in Macedonian." J. N. Beckman et al, eds. *Papers in Optimality Theory*. Amherst: Graduate Linguistic Student Association, 138–66. (*University of Massachusetts occasional papers*, 18.)

Elson, Mark J. (1976) "The indefinite article in Bulgarian and Macedonian." *Slavic and East European journal* 20(3): 273–79.

Elson, Mark J. (1993) "Collocational stress in Contemporary Standard Macedonian." *Slavic and East European journal* 37(2): 149–61.

Englund, Birgitta. (1977) *Yes/no-questions in Bulgarian and Macedonian: Form*. Stockholm: Almqvist & Wiksell International. (Acta Universitatis stockholmensis, Stockholm Slavic studies, 12.)

Englund, Birgitta. (1979) "The function of yes/no-questions (on basis of Bulgarian and Macedonian material)." In *Papers in Slavonic linguistics*. Stockholm, 16–26. (Meddelanden från institutionen för slaviska och baltiska språk", 18.)

Ewen, Robert Charles. (1979) *A grammar of Bulgarian clitics*. Ph.D. dissertation, U. of Washington.

Franks, Steven L. (1987) "Regular and irregular stress in Macedonian." *International journal of Slavic linguistics and poetics* 35–36: 93–142.

Franks, Steven (1989) "The monosyllabic head effect." *Natural language and linguistic theory* 7(4): 551–63.

Franks, Steven (1991) "Diacritic extrametricality *vs.* diacritic accent: A reply to Hammond." *Phonology* 8(1): 145–61.

Friedman, Victor. (1993) "Macedonian". B. Comrie and G. C. Corbett, eds. *The Slavonic languages* (chap. 6). London: Routledge, 249–305.

Garde, Paul. (1968) "Note sur l'enclise et la proclise en macédonien." *Zbornik za filologiju i lingvistiku* [Novi Sad: Matica srpska] 11: 31–37.

Halle, Morris and Michael Kenstowicz. (1991) "The Free Element Condition and cyclic vs noncyclic stress." *Linguistic inquiry* 22(3): 457–501.

Halpern, Aaron Lars (1992/1995) *On the placement and morphology of clitics*. Stanford, CA: CSLI Publications. (Published version of his *Topics in the placement and morphology of clitics*. Ph.D. dissertation, Stanford U.)

Hammond, Michael (1989) "Lexical stress in Macedonian and Polish." *Phonology* 6(1): 19–38.

Hauge, Kjetil Rå. (1976) *The word order of predicate clitics in Bulgarian*. Oslo: Universitetet i Oslo, Slavisk-Baltisk Institutt. (*Meddelelser*, 10.)

Ilievski, P. Xr. and K. Peev (1983) "Neobična inovacija vo dolnovardarskite govori (Prokliza pri imperativ)." *Makedonski jazik* 34: 105–13.

Izvorski, Roumjana. (1993) "On WH-movement and focus movement in Bulgarian". Presented at CONSOLE 2, Universität Tübingen.

Izvorski, Roumyana. (1994) "Yes-no questions in Bulgarian: Implications for phrase structure". Presented at 9th Biennial Conference on Balkan and South Slavic Linguistics, Literature, and Folklore, Indiana University, Bloomington.

Izvorski, Roumyana, Tracy Holloway King, and Catherine Rudin. (1996) "Against *li* lowering in Bulgarian". To appear in *Lingua*.

Kager, René. (1993) "Stress in windows". ms., Universiteit Utrecht.

Kenstowicz, Michael. (1991) "Enclitic accent: Macedonian, Italian, Polish." P. M. Bertinetto et al, eds. *Certamen phonologicum II: Papers from the 1990 Cortona Phonology Meeting*. Torino: Rosenberg & Sellier, 173–85.

Kepeski, Petar and Jože Pogačnik. (1968) *Makedonščina*. Ljubljana: Državna založba Slovenije.

King, Tracy Holloway (1993/1994/1995) "Focus in yes-no questions." *Configuring topic and focus in Russian*. Stanford: CSLI Publications, 137-163 (chap. 6). (Published, somewhat revised version of (chap. 5 [sic.]) of her (1993) Ph.D.. dissertation, Stanford U. This chapter also appeared as her (1994) "Focus in Russian yes-no questions". *Journal of Slavic linguistics* 2(1): 92–120.)

Koneski, Blaže. (1957/1967) "Za akcentot". *Gramatika na makedonskiot literaturen jazik*. Skopje: Kultura. (Blaže Koneski: Izbrani dela vo sedum knigi, 6.)

Korubin, B. (1974) "Mesto glagola *sum* kak vspomogatel´nogo i kak svjazki v makedonskom literaturnom jazyke." N. Ju. Švedova et al, eds. *Grammatičeskoe opisanie slavjanskix jazykov: Koncepcii i metody*. Moskva: Nauka, 244–50.

Kramer, Christina. (1985) *Anglisko-makedonski razgovornik*. Skopje: Univerzitet "Kiril i Metodij".

Kramer, Christina E. (1986) *Analytic modality in Macedonian*. München: Verlag Otto Sagner. (Slavistische Beiträge 198.)

Kramer, Christina. (1993) "The future in Macedonian". Presented at American Association of Teachers of Slavic and East European Languages conference, Toronto.

Lüdtke, H. (1979) *Lucania*. Pisa: Pacini. (Profilo dei dialetti italiani 17.)

Lunt, Horace G. (1979) *A grammar of the Macedonian literary language*. Skopje: Državno Knigoizdatelstvo na NR Makedonija.

Mihaljević, Milan. (1996) "The interaction of li and negation in Croatian Church Slavonic". To appear in G. Zybatow and U. Junghanns, ed. *Formale Slavistik*. (Presented at the First European Conference on "Formal Description of Slavic Languages", 1995.)

Mussafia, A. (1898) "Enclisi o proclisi del pronome personale atono qual oggetto". *Romania* 27: 145–46.

Penčev, Jordan. (1993). *Bâlgarski sintaksis: Upravlenie i svârzvane*. Plovdiv: Plovdivsko universitetsko izdatelstvo.

Peperkamp, Sharon. (1996) "On the prosodic representation of clitics". U. Kleinhenz, ed. *Interfaces in phonology*. Berlin: Academie Verlag, 102–27. (Studia grammatica, 41.)

Restan, Per. (1972) *Sintaksis voprositel´nogo predloženija; obščij vopros (Glavnym obrazom na materiale sovremennogo russkogo jazyka [Syntax of the interrogative sentence; yes-no questions (Mainly on the basis of Russian material)*. Oslo: Universitetsforlaget.

Rivero, María-Luisa. (1993) "Bulgarian and Serbo-Croatian yes-no questions: V^0 raising to *-li* vs. *-li* hopping". *Linguistic inquiry* 24(3): 564–75.

Rudin, Catherine. (1990–91) "Topic and focus in Bulgarian". *Acta linguistica hungarica* 40(3–4): 429–47.

Rudin, Catherine. (1992) "Kakvo li e *li*: Interrogation and focusing in Bulgarian". To appear in *Balkanistika* 10. (Special issue: V. Friedman, ed., Gedenkschrift for Zbigniew Gołąb.)

Rudin, Catherine. (1993a) "On focus position and focus marking in Bulgarian questions". A. Davison et al, eds. *Papers from the 4th annual meeting of the Formal Linguistics Society of Midamerica*. Iowa City: Univ. of Iowa, 252–65.

Rudin, Catherine. (1993b) "On the syntax of *li* questions in Bulgarian". Presented at American Association of Teachers of Slavic and East European Languages conference, Toronto.

Rudin, Catherine. (1996) "Agr-O and Bulgarian pronominal clitics". To appear in S. L. Franks, ed. *Formal approaches to Slavic linguistics*, 5. Ann Arbor: Michigan Slavic Publications.

Rudin, Catherine, Tracy Holloway King, and Roumyana Izvorski. (1995) "Focus in Bulgarian and Russian yes-no questions". Presented at the Workshop on Focus, Univ. of Massachusetts, Amherst.

Rudin, Catherine and Christina Kramer. (1994) "*Li* questions in Macedonian". Presented at American Association of Teachers of Slavic and East European Languages conference, San Diego.

Sadock, Jerrold M. (1991) *Autolexical syntax: A theory of parallel grammatical representations*. Chicago: University of Chicago Press.

Scatton, Ernest. (1984) *A Reference Grammar of Modern Bulgarian*. Columbus: Slavica.

[*SSKG*] (1975) *Slovar slovenskega knjižnega jezika*. Ljubljana: Slovenska akademija znanosti in umetnosti, Institut za slovenski jezik, v. 2 (*I–Na*).

Toman, Jindřich. (1996) "A note on clitics and prosody". A. L. Halpern and A. M. Zwicky, eds. *Approaching second: Second position clitics and related phenomena.* Stanford, CA: CSLI Publications, 505–10 (chap. 15). (= CSLI lecture notes 61.)

Tomić, Olga M. (1996a) "The Balkan Slavic nominal clitics." A. L. Halpern and A. M. Zwicky, eds. *Approaching second: Second position clitics and related phenomena*. Stanford, CA: CSLI Publications, 511–35 (chap. 16). (= CSLI lecture notes 61.)

Tomić, Olga M. (1996b) "The Balkan Slavic clausal clitics." To appear in *Natural language and linguistic theory*. (Announced in *NLLT* 14(3).)

Usikova, R. P. (1964) "Ob udarenii v sovremennom makedonskom jazyke". V. D. Koroljuk et al, eds. *Slavjanskaja i baltijskaja akcentologija*. Moskva: Nauka, 27–39. (Kratkie soobščenija Instituta slavjanovedenija, 41.)

Usikova, R. P. (1985) *Makedonskij jazyk: Grammatičeskij očerk, teksty dlja čtenija c kommentarijami i slovarem*. Skopje: Makedonska kniga.

Wackernagel, Jacob (1892) "Über ein Gesetz der indogermanischen Wortstellung." *Indogermanische Forschungen* 1: 333-436.

Wilder, Chris and Damir Ćavar. (1994) "Long head movement? Verb movement in cliticization in Croatian". *Lingua* 93: 1–58.

Zec, Draga and Sharon Inkelas. (1990) "Prosodically constrained syntax." S. Inkelas and D. Zec, eds. *The phonology-syntax interface*. Chicago: University of Chicago Press, 365–78 (chap. 18).

[Billings]
Institut für Slavistik
Universität Leipzig
Augustusplatz 9
D–04109 Leipzig
billings@rz.uni-leipzig.de

[Kramer]
Dept. of Slavic Lang. and Lit.
University of Toronto
21 Sussex Ave.
Toronto, Ont. M5S 1A1 Canada
ce.kramer@utoronto.ca

[Rudin]
Linguistics
Humanities Division
Wayne State College
Wayne, NE 68787 USA
crudin@wscgate.wsc.edu

Weakening, Strengthening, and Nasalization;
the things that happen to consonants in morphological environments[i].

Janet Grijzenhout, Utrecht University

1 Introduction

In this paper I will present a phonological account of consonant alternations that are triggered by morphosyntactic or lexical environments. These alternations involve gradient changes in the degree of oral aperture or in the degree of sonority and cannot be explained by referring to the phonological environment. In some cases, these processes not only affect the degree of oral aperture or the degree of sonority, but also the place of articulation. The aim of this paper is twofold. First, I will attempt to provide a unified autosegmental account of morphosyntactically conditioned gradient changes in the manner of articulation and in sonority. The second aim is to provide a phonological account for the interaction between manner of articulation and place of articulation.

Consonant alternations that affect the manner of articulation involve changes along the scale in (1) below:

(1) Changes in the manner of articulation:
long plosive <---> short plosive <---> fricative/approximant <---> zero

Rightward changes are commonly referred to as 'weakening' or 'lenition' and occur, for instance, in Finnish, Irish, and Welsh. Changes in the reverse direction are known as 'strengthening' or 'fortition' and are found in West-Atlantic Fula, and Amerindian Southern Paiute.

Consonant alternations that affect the degree of sonority involve changes along the scales in (2):

(2) Changes in the degree of sonority:
a. voiceless obstruent ---> voiced obstruent ---> nasal sonorant
b. voiceless obstruent ---> prenasal obstruent
c. voiceless obstruent ---> nasal sonorant

In Irish, alternations between voiceless and voiced obstruents occur in the same morphological environments in which voiced stops alternate with nasal stops. In Southern Paiute, we find that voiceless stops alternate with prenasal stops in certain grammatical environments and in Welsh, voiceless aspirated stops alternate with voiceless nasals, while unaspirated stops alternate with voiced nasals.

In sections 1.1 to 1.3 I will present some examples of consonant weakenings, consonant strengthenings, and consonant nasalizations, respectively. The theoretical framework that I will adopt is introduced in section 2. Section 2.1 provides an analysis of consonant weakening processes in Finnish and Irish in this framework and section 2.1.1 is devoted to two types of consonant weakening processes in Welsh. Sections 2.2, and 2.3 provide an analysis of consonant strengthenings and consonant nasalizations, respectively. In section 3 I discuss some anomalies concerning the place of articulation and I will argue in favour of Rice's (1996) model of underspecification of place features to explain these anomalies. Section 4 concludes.

1.1 Consonant weakening

In Finnish, a phonetically long oral stop is realised as a phonetically short oral stop in the same morphological environment in which an underlying short stop is realised as an approximant. This process is commonly referred to as 'consonant gradation'. The examples below are from Kiparsky (1992),Skousen (1972), and Vainikka (1988). According to Vainikka (1988), [D] represents a coronal sonorant continuant:

(3) a.	piippu	pp=[p:]	'pipe' (nom)
b.	piipun	p =[p]	'pipe' (gen)
(4) a.	tapa	p =[p]	'custom' (nom)
b.	tavan	v =[w]	'custom' (gen)
(5) a.	hattu	tt =[t:]	'hat' (nom)
b.	hatun	t =[t]	'hat' (gen)
(6) a.	katu	t =[t]	'street' (nom)
b.	kadun	d =[D]	'street' (gen)

The chain-like changes from a phonetically long stop to a short stop and from a short stop to a continuant usually occur after a vowel or sonorant consonant in the onset of a closed syllable. However, there are some suffixes that close a syllable and do not trigger a change in the onset stop:

(7) a.	käte	t =[t]	'hand'
b.	kätensä	t =[t]	'his hand'

From this we may infer that consonant gradation is not conditioned by the phonological environment, but by the morphological environment. The phonological context does not play a role in the gradation process and need not be specified in our account of these alternations. The changes depicted in (3b, 4b, 5b, 6b) occur in exactly the same morphological context and constitute one phonological process. The change from a phonetically long stop to a short stop and from a short stop to an approximant in the same environment is difficult to express in a feature geometric model, because there is no unified set of features which accounts for it. In section 2.1 I will show how a theory which employs so-called 'aperture positions' (which are defined in terms of degrees of oral aperture) enables us to express Finnish consonant gradation as one process.

Another example of a consonant weakening process can be found in Modern Irish. In this language, a word which is pronounced with a short stop in isolation is realised with a fricative after certain function words or to mark a syntactic property (e.g. to indicate the past tense of a verb). In the same environments, a word with an initial labial fricative is realised without an initial consonant. The examples below are from Ní Chiosáin (1991):

(8)	a.	páipéar	p =[p]	'a paper'
	b.	mo pháipéar	ph=[f]	'my paper'
(9)	a.	fata	f =[f]	'a potato'
	b.	mo fhata	fh=Ø	'my potato'

The change from an underlying stop to a fricative in an environment that cannot be described in phonological terms is also attested in native American Southern Paiute. In this language, certain lexical items trigger a mutation in a consonant-initial suffix. In the examples below from Sapir (1930) the suffix *-pi* is realised with an initial fricative after the lexical items *sappi* 'belly' and *avi* 'shade' (other lexical items may induce other changes, see sections 1.2 and 1.3 below). Whether the fricative that results from this spirantization process is voiceless or voiced depends on the preceding vowel. Fricatives are voiceless after voiceless vowels (indicated below by small caps) and voiced after voiced vowels:

(10)	a.	sappɪ + pi---> sappɪɸi	'belly' (absolutive)
	b.	avi + pi ---> avivi	'shade' (absolutive)

The fact that spirantization of stops is not triggered by a preceding vowel in Southern Paiute, but by other non-phonological conditions can be illustrated by the following quotation. According to Press (1979:24): 'a few postpositions in Southern Paiute vary their initial consonant according to whether the noun to which they are attached is animate or not, regardless of the "phonological shape" of the stem'.

Stop-fricative alternations are frequently found in intervocalic positions. In Spanish, for instance, the voiced stops /b, d, g/ are realised as the corresponding voiced fricatives [β, ð, ɣ] in between two sonorants, and in certain dialects of American English the voiced stop /d/ is realised as the flap [ɾ] in intervocalic position. In Finnish, Irish, and Southern Paiute, however, the respective consonant changes cannot be attributed to surrounding sounds (although this was the case historically in Irish). I will account for these morphosyntactically or lexically conditioned weakening phenomena in a uniform way in section 2.1.

1.2 Consonant strengthening

The reverse of consonant weakening, i.e. consonant strengthening, is also attested, although gradient changes to sounds which are one degree more constricted are hard to find. There no changes from zero to continuant which are not triggered by the phonological environment. This is probably due to the following facts. First, onsetless syllables are rare cross-linguistically as well as within one language. Inserting a continuant at the left edge of the word where, in most instances, there is an underlying onset consonant already would consequently not be a very effective way of marking a morphological process. Second, to insert a continuant at the right edge of a word goes against the general tendency in languages to end a syllable by either a vowel or a non-continuant (i.e. an oral or nasal stop). Considering these facts, it is not surprising that we do not find the option of using an inserted continuant as a morphological or lexical marker in languages.

In contrast, the change from no underlying consonant to the presence of a continuant is frequently found in phonological contexts. In Dutch, for example, an approximant is inserted between two vowels (of which the first one determines the place of articulation of the approximant):

(11) a.	piano	pi[j]ano	c.	Inuit	Inu[w]it
b.	theater	the[j]ater	d.	boa	bo[w]a

The change from a continuant to a short plosive is found in morphological environments in West-Atlantic Fula (Arnott 1970, Paradis 1992).The stem *baat* 'needle' does not change its initial consonant in noun class 3, while *weer* 'host' appears with initial *b* in that class and *fow* 'hyena' is realised with initial *p*:

(12) a.	baat + el	--->	baatel	'needle' (diminutive sg.)
b.	weer+ el	--->	beerel	'host' (diminutive sg.)
c.	fow + el	--->	powel	'hyena' (diminutive sg.)

In Southern Paiute, some lexical items cause a suffix-initial short stop to lengthen:

(13)	a + pi	--->	ap:i	'horn' (absolutive)

The examples given in this section illustrate non-gradient leftward changes along the scale in (1) above. I will now turn to examples of changes along the scales in (2a-c).

1.3 Consonant nasalization

Irish has a morphological process know as 'initial nasalization' by which voiceless obstruents become voiced and voiced stops become nasal stops (there are no underlying voiced fricatives in Irish). The data below are from Ní Chiosáin (1991):

(14) a.	peann	p =[p']	'pen'
b.	a bpeann	bp=[b']	'their pen'
(15) a.	fear	f =[f']	'man'
b.	leis an bhfear	bhf=[v']	'with the man'
(16) a.	bó	b =[b]	'cow'
b.	ocht mbó	mb=[m]	'eight cows'

Another nasalization process is found in Amerindian Southern Paiute. In this language, a lexical item may trigger prenasalization in a stop-initial suffix:

(17)	ago + pi	--->	ago^{m}pi	'tongue' (absolutive)

There are no underlying oral voiced stops or prenasal stops in Southern Paiute. I know of no language in which voiceless stops alternate with prenasal stops in the same morphological context where prenasal stops alternate with nasals. The analysis that I will present in section 2.3 below predicts a system in which such chain-like changes occur. The reason why a language to illustrate this system is difficult to find may be due to two facts. First, languages with underlying prenasal stops are rare and, second, languages which utilise consonant nasalization as a morphosyntactic process are rare. Together, the two facts explain why we do not find languages in which voiceless stops turn into prenasal stops in the same context in which underlying prenasal stops turn into nasals.

In Welsh, voiceless aspirated stops are realised as voiceless nasals in the same environment in which voiceless unaspirated stops are realised as plain nasal stops (Ball and Müller 1992):

(18)	a.	cath	c =[kʰ]	'cat'
	b.	fy nghath	ngh=[ŋ̊]	'my cat'
(19)	a.	Fe'm gwelodd Wyn	g =[k]	'Wyn saw me'
	b.	Roedd Wyn yn fy ngweld	ng =[ŋ]	'Wyn was seeing me'

The examples presented in this section differ from the phonological phenomenon known as 'nasal harmony' in that there is no overt phonological trigger like a nasal consonant or nasal vowel. In most instances, the degree of sonority is increased by lowering of the velum for part of the segment's duration (Southern Paiute) or for the entire duration of the segment (Welsh). Irish displays a chain-wise increase of sonority by (i) voicing voiceless segments and (ii) nasalizing segments which are voiced underlyingly. In section 2.3 I will propose that the processes illustrated in this section always involve the addition of the feature 'Sonorant Voicing' (see Rice 1993).

2 Alternations in manner of articulation and autosegmental phonology

The consonant alternations presented in sections 1.1 and 1.2 involve changes in the manner of articulation. In autosegmental phonology, the so-called 'manner features' include [continuant] and [approximant]. These features pose a number of problems for phonological theory.

First, McCarthy (1988), among others, observes that the feature [approximant] does not participate in autosegmental processes like spreading or delinking. Incorporating the feature [approximant] into the root node explains why this feature cannot spread or delink on its own. Wetzels (1991) observes that we seldom find that the feature [+continuant] spreads from a fricative to a following stop as in the hypothetical case *aspa* --> *asfa*, or is given up when adjacent to another continuant, as in *afsa* --> *apsa*. Given the theory of autosegmental phonology in which spreading and delinking of features are the basic operations on segmental representations, the behaviour of the feature [continuant] is unexpected. A possible solution to this problem may be to incorporate this feature in the root node together with the feature [approximant], but this poses other problems, e.g. for the representation of affricates.

Second, Steriade (1993) observes that in a theoretical model which assumes the features [continuant] and [approximant] there is no explanation for the fact that complex segments like pre- and postnasals involve plosives (stops and affricates), rather than continuants (fricatives and approximants).

Third, consonant weakening processes may involve phonetically long plosives which are shortened in the same context in which short plosives are spirantised (Finnish), or short plosives which are spirantised in the same environment in which fricatives are deleted (Irish) and such chain-like changes are difficult to express in a theory which assumes the features [approximant] and [continuant]. Linguists have referred to these changes as 'bizarre, irregular, and quirky phenomena' (Lieber 1987). Many attempts have been made to express such weakening processes as a single generalisation (e.g. Ní Chiosáin 1991, Vainikka 1988), but the result has been analyses in which it has been proposed, for example, that [+continuant] is inserted in the case of a segment that is underlyingly underspecified for that feature, while additional rules ensure that an underlying [+continuant] specification results in deletion of all underlying features. Such rules are indeed highly 'irregular'. As an example of such an approach, I will discuss the autosegmental account of Modern Irish initial lenition proposed by Ní Chiosáin (1991).

Ní Chiosáin (1991) assumes the following underlying representations for Irish consonantal phonemes (palatalization, which is distinctive in Irish, is ignored):

(20) Underlying representations of Irish consonants:

a.	/p,t,k,b,d,g/	b.	/f/	c.	/h/
	[-son,+cons]		[-son,+cons]		[-son]
			|		
			[+cont]		

In environments where lenition does not apply, a redundancy rule inserts the value [-continuant] for stops and [+continuant] for /h/. In environments where lenition applies the value [+continuant] is inserted by the spirantization rule below:

(21) Spirantization

$[_{Wd}\ \alpha$

 |

[ ] ---> [+continuant]

The sound /f/ undergoes the spirantization rule vacuously, because it is underlyingly specified as [+continuant], (see 20b). For this case, Ní Chiosáin (1991:51) proposes a default rule 'Total Deletion', which only applies when the spirantization rule applies vacuously. This is theoretically 'suspect'. I will now discuss a theoretical framework which overcomes the problems which a theory with manner features has.

Steriade (1993) proposes to define the nodes to which features may attach in terms of degree of oral aperture. Instead of the feature-geometric notion of root node which may comprise the features [consonantal], [sonorant], and (perhaps) [approximant], Steriade assumes the following aperture positions:

(22) a.	A0:	complete obstruction in the oral cavity
b.	Afric:	degree of oral release sufficient to create a turbulent airstream
c.	Amax:	maximal degree of oral release for consonants

Some consequences of this proposal are, first of all, that the features [approximant] and [continuant] are no longer necessary in phonological representations and that the exceptional status of these features with respect to spreading or delinking processes does not need a special explanation. Second, since non-continuants (stops and affricates) involve two aperture nodes (one for the closure phase and one for the release phase), whereas continuants involve one node (for the release phase), it follows that features like [nasal], [spread glottis], etc. have two positions to associate to in the case of non-continuants and one in the case of continuants. This proposal predicts a four-way contrast among non-continuants and a two-way contrast among continuants:

(23) a. stops:	aspirated	preaspirated	postaspirated	unaspirated
	A0 Amax	A0 Amax	A0 Amax	A0 Amax
	[+spr. gl]	[+spr. gl]	[+spr.gl]	
b. fricatives:	aspirated	unaspirated		
	Afric	Afric		
	[+spr. gl]			

Third, Steriade's proposal to represent the manner of articulation by means of aperture positions has interesting consequences for consonant weakening, consonant strengthening, and consonant nasalization processes, as I will now go on to show in the following sections.

2.1 Consonant weakening processes and aperture theory

Languages that distinguish released stops and affricates (e.g. German) and languages that distinguish voiced fricatives and voiced approximants (e.g. Dutch) have a phonological distinction between two degrees of oral release. Steriade (1993) suggests the following segmental representations involving aperture positions for fricated and approximant release:

(24) a.	approximant:	Amax	c.	fricative:	Afric
b.	released stop:	A0 Amax	d.	affricate:	A0 Afric

In what follows I will not be concerned with different degrees of release. The idea that I will borrow from Steriade (1993) is that released non-continuants involve two positions in phonological representations, whereas continuants are represented by one position. Stops in utterance-medial positions are generally unreleased and are assumed to carry one position for closure in phonological representations (see 25a). For languages that do not have affricates, nor a phonological distinction between voiced fricatives and voiced approximants, the difference between voiceless fricatives and voiced approximants may be indicated by a feature for voicing under the aperture position for release. The aperture position that characterises fricatives as well as approximants in such languages is given in (25b) and the representation of released stops is given in (25c):

(25) a.	unreleased stop (in coda position):	A0
b.	continuant:	Arel
c.	released stop (in onset position):	A0 Arel

One of the consequences of the proposal to represent closure and release in the oral cavity by distinct aperture positions is that it enables us to express consonant weakening as one process, viz. the deletion of an aperture position, as I will illustrate for Finnish first.

Finnish has the following consonant inventory:

(26) *Finnish consonantal phonemes*

	Labial	Dental	Palatovelar	Laryngeal
stops	p	t	k	
fricatives		s		h
nasals	m	n	ŋ	
approximants	w	l/r	j	

In the case of released stops, the aperture position for closure is lost under gradation and what remains is the aperture position for release which characterises an approximant:

(27) t a p a+n ---> t a [w] a n ('custom')

$$\begin{array}{ccccccc} t & a & p & a+n & \longrightarrow & t\ a\ [w]\ a\ n \\ & & / \backslash & & & | \\ & & A_0\ A_{max} & & & A_{max} \end{array}$$

The same process affects phonetically long stops. These segments are usually referred to as 'geminates' and can be analysed as a sequence of an unreleased stop in coda position followed by a homorganic released stop in onset position. The result of gradation, i.e. deletion of A0, is a A0 Amax sequence, which characterises a phonetically short stop:

(28) l a p p u+n ---> l a [p] u n ('paper')

$$\begin{array}{ccccccc} l & a & p & p & u+n & \longrightarrow & l\ a\ [p]\ u\ n \\ & & | & / \backslash & & & / \backslash \\ & & A_0 & A_0\ A_{max} & & & A_0\ A_{max} \end{array}$$

To sum up, (29) schematically represents the analysis of Finnish gradation that I propose in this paper:

(29) Finnish Consonant Gradation as deletion of A_0:

long stop ---> short stop;		short stop ---> approximant
$A_0\ A_0\ A_{max}$ ---> $A_0\ A_{max}$		$A_0\ A_{max}$ ---> A_{max}

I will now argue that the same analysis accounts for Irish initial lenition. In (30) I present the relevant part of the phonological representation forIrish obstruents. Palatalization is ignored and I will not be concerned with the strident fricative (for a phonological account of the behaviour of /s/ in Irish, see Grijzenhout 1995). There are no underlying affricates, voiced fricatives, or glides in Irish and for this reason I will use Arel rather than Afric and Amax in the representations below:

(30) a. /p,t,k,b,d,g/ $A_0\ A_{rel}$ b. /f/ A_{rel}

The stem-initial lenition process illustrated in examples (8b) and (9b) can be analysed as deletion of a stem-initial aperture position as illustrated below:

(31) Modern Irish Initial Lenition as deletion of the initial Aperture Position:

initial stop	---> fricative;	initial fricative	---> zero
# $A_0\ A_{rel}$	---> A_{rel}	# A_{rel}	---> Ø

Welsh has a similar process of stem-initial consonant weakening and I will argue next that these also involve deletion of one aperture position. Welsh has a contrast between aspirated and unaspirated plosives and for this reason I will discuss the phonological representation of aspiration first.

2.1.1 Aspiration and aperture theory

In this section I will illustrate Steriade's (1994) account of word-initial aspiration in Huautla Mazateco, an Otomanguean language of Oaxaca, Mexico, and discuss the consequences for an account of word-initial aspiration and stem-initial lenition in Modern Welsh.

Huautla Mazateco has a complex pattern of onset consonants including pre- and postglottalised plosives, pre- and postaspirated plosives, prenasalised plosives and prenasalised aspirated plosives. In the present section, we concentrate on aspiration. Steriade (1994:219) presents the following consonant inventory:

(32) *Mazateco consonantal phonemes*

labial	alveolar	strident	postalveolar	retroflex	velar	laryngeal
b	t	ts	tʃ	tʂ̻	k	ʔ
		s	ʃ	ʂ̻		
m	n		ɲ			
	l		y			h

In onset position of Mazateco words, we find single consonants, but no consonant clusters. In word-initial positions, oral stops and affricates may be preaspirated, while nasal stops may be partly devoiced:

(33) a. hti ht =[ʰt] 'fish'

b. htse hts=[ʰts] 'a sore'

c. hno hn =[n̥n] 'corn'

The segment /h/ involves maximal aperture for consonants in the oral cavity, no place of articulation, and a spread glottis. Steriade (1994) therefore represents laryngeal /h/ as an A_{max} position with the feature [spread] for aspiration:

(34) /h/: A_{max}
|
[spread]

This segment may appear on its own in Mazateco. For reasons beyond the scope of the present paper, Steriade (1994:233-234) argues that onsets with preaspiration are generated by a merger process of the aperture position for /h/ and the aperture position for closure of the onset stop. This is compatible with the fact that aspiration is phonetically realised as simultaneous with at least the first half of the stop closure:

(35) a. Merger deriving preaspirated oral stops in onsets:

[spread] [spread]
| |
A_{max} A_0 A_{max} ---> A_0 A_{max}

b. Merger deriving partially devoiced nasal stops in onsets:

```
[spread]                 [spread]
 |                        |
Amax A0 Amax     --->    A0 Amax
        \ /                \ /
      [nasal]            [nasal]
```

In Mazateco, consonants are available landing sites for the features of aspiration and glottalisation, but the restriction seems to be that the one feature excludes the other. Onset consonants may be preaspirated or postaspirated and preglottalised or postglottalised, but not, for instance, preaspirated-postglottalised. In Steriade's framework this follows from the fact that aspirated and glottalised onset consonants are monosegmental and only one laryngeal feature is allowed per segment (below [spr] and [cstr] stand for the features for spread and constricted glottis, respectively):

```
(36)a. preaspirated stop   b. postglottalised stop   c. impossible
       [spr]                      [cstr]                [spr][cstr]
       |                          |                       |    |
       A0 Amax                    A0 Amax                *A0 Amax
```

The analysis of Huautla Mazateco aspiration finds independent support from Welsh, as will be shown next.

According to Ball & Jones (1984), the voicing contrast on Welsh plosives is phonetically realised as the presence or absence of aspiration. The fricatives exhibit a contrast in length. All speakers use length of frication as a distinguishing feature between what is usually notated as /f,θ/ and /v,ð/, respectively, while some also utilise an amount of voicing for the fricatives /v/ and /ð/ although they are still not fully voiced. To better represent the distinction in length, I will use /f:/ and /θ:/ for the phonetically long fricatives and /f/ and /θ/ for their shorter counterparts. The following table presents an inventory of Welsh consonants that may occur in word-initial non-mutation environments:[ii]

(37) *Modern Welsh initial consonants*

	labial	dental	strident	velar
stops	p^h, p	t^h, t		k^h, k
fricatives	f:, f	θ:, θ	s	
nasals	m	n		
liquids		ɬ, l, r̥, r		

As in Irish, word-initial consonants in Welsh may be weakened in certain morphological or syntactic environments. Even though the attested consonant alternations are different from Modern Irish, I will maintain that the phonological process is the same, viz. lenition entails the deletion of an initial aperture position.

In Welsh lenition environments, aspirated plosives are realised as unaspirated plosives (38b), unaspirated plosives as fricatives (39b), fricatives are not affected, and devoiced sonorants are fully voiced (40b). The following examples are from Willis (1986):

(38)	a.	pen	p =[p^h]	'head'
	b.	ei ben	b =[p]	'his head'
(39)	a.	brawd	b =[p]	'brother'
	b.	ei frawd	f =[f]	'his brother'
(40)	a.	llong	ll=[ɬ]	'ship'
	b.	ei long	l =[l]	'his ship'

With Steriade (1994), I assume that aspiration for plosives and devoicing of sonorants is carried by a separate aperture position for laryngeal release. In non-lenition environments the aperture position for laryngeal release merges with the aperture position for closure of the onset stop or the aperture position for sonorants, as illustrated for Mazateco aspirated onset stops in (35a).

In accordance with the analysis of consonant weakening processes in Finnish and Irish, I will analyse Welsh stem-initial lenition as deletion of an aperture position. In lenition environments, the initial aperture position is deleted in a segment characterised by more than one aperture position and this may be the aperture position for laryngeal release. Contrary to Irish, fricatives are not affected in Welsh due to the fact that an onset consonant is obligatory in Welsh. The consequence is that an aperture position for release cannot be deleted and must remain to fill the onset position: [iii]

(41) Welsh initial lenition as deletion of an initial A-position:

a. aspirated plosive ---> unaspirated plosive

[spread]
|
A_{rel} A_0 A_{rel} ---> A_0 A_{rel}

b. unaspirated plosive ---> phonetically short fricative

A_0 A_{rel} ---> A_{rel}

c. fricative unaffected (to fill the onset-position)

A_{rel}

d. devoiced sonorant ---> fully voiced sonorant

[spread]
|
A_{rel} A_{rel} ---> A_{rel}
| |
[lateral]   [lateral]

What emerges from this discussion is that lenition in Welsh and Irish is basically the same process, but, since the underlying representations are different in both languages, the outcome of the lenition process has different phonetic results. [iv]

This concludes the discussion of the phonological analysis of consonant weakening processes. In all the cases that we have come across, consonant weakening involves deletion of an aperture node. As independent support for the theory, I discuss the analysis of consonant strengthening processes in the following section.

2.2 Consonant strengthening processes and aperture theory

In section 1.2 it was illustrated that Fula approximants turn into voiced stops in the same environment in which voiceless fricatives turn into voiceless stops. This process is difficult to account for in a autosegmental theory. We may, for instance, suggest that the process involves the introduction of the feature value [-continuant] or the deletion of the feature value [+continuant], but that would leave unexplained why the approximant loses its sonorant status. In the theory proposed here, voiceless fricatives are characterised by one position for voiceless release and approximants are represented as one position for voiced release. Following Rice (1993) and in anticipation of the discussion on sonority in section 2.3, I will use the feature SV (sonorant voicing) to indicate voicing of obstruents and sonorants. The change from stem-initial continuant to released stop in certain noun classes in Fula can be analysed as the insertion of an aperture position for closure:

(42) Fula Strengthening as insertion of initial A0:

a. voiceless continuant ---> voiceless stop (e.g. f---> p)

Arel ... A0 Arel

b. voiced continuant ---> voiced stop (e.g. w ---> b)

$$\begin{array}{ccc} \text{Arel} & & \text{A0 Arel} \\ | & & | \\ \text{SV} & & \text{SV} \end{array}$$

I propose the same analysis for Southern Paiute lengthening of short plosives (see example 13):

(43) Southern Paiute Gemination as insertion of initial A_0:

short stop ---> phonetically long stop (e.g. p ---> p:)

A_0 A_{rel} ---> A_0 A_0 A_{rel}

All consonant weakening processes manifest themselves as a decrease of consonantal stricture and consonant strengthening processes manifest themselves as an increase of consonantal stricture. The theory developed on independent grounds in Steriade (1993) seems particularly suited to capture this fact, i.e. a decrease of stricture is analysed as the loss of an aperture position and an increase of stricture is analysed as insertion of an aperture position. I will now consider an analysis of consonant nasalizations in this framework.

2.3 Consonant nasalization processes and aperture theory

In standard phonological representations, voicing of obstruents is expressed by the feature [voice] and the "spontaneous voicing" (Chomsky and Halle 1968) which characterises voicing of sonorants is expressed by the feature [sonorant]. Rice (1993) shows that in languages in which voicing of obstruents and voicing of sonorants is not distinctive, we capture more generalisations if the two features are replaced by one phonological feature 'Spontaneous Voicing' or 'Sonorant Voice' (abbreviated as SV). I what follows I will adopt her proposal for the languages under discussion here.

Consider first Southern Paiute Prenasalization which turns suffix-initial oral stops into prenasal stops after certain lexical items. I will analyse this process as one of SV-association to the stem-initial segment as below:

(44) Southern Paiute Nasalization:

```
voiceless stop ---> prenasal
A0 Arel             A0 Arel
                    |
                    SV
```

When there is a complete obstruction in the oral tract (represented by A0), the air which passes the vibrating vocal cords (represented by SV) cannot escape through the oral cavity and has to be released through the nasal cavity. For this reason, a representation of A0 which dominates SV always indicates nasality. A sequence of A0 Arel where SV is associated to A0, represents nasality followed by oral release, i.e. this is the representation of a prenasal.

The same process of SV-association applies to Irish stem-initial nasalization. In this language there is a preference for association to Arel rather than to A0 (in Grijzenhout 1995 I argue that this is a cross-linguistic tendency). When there is a choice, the prefix SV will associate to an Arel position in Irish rather than to an A0 position:

(45) Modern Irish Initial Nasalization; the case of voiceless obstruents:

```
voiceless stop ---> voiced stop;   voiceless fricative ---> voiced fricative
   A0 Arel          A0 Arel               Arel                  Arel
                       |                                         |
                       SV                                        SV
```

Nasal consonants are more sonorant than voiced obstruents and this may be expressed in phonological representations by multiple association of SV to two aperture positions. In morphological contexts in which the prefix SV is associated to the aperture position for release in the case of voiceless obstruents, it is associated to the other position in the case of voiced obstruents in Irish:

(46) Modern Irish Initial Nasalization; the case of voiced stops:

```
voiced stop   ---> nasal stop
   A0 Arel         A0 Arel
   |                \ /
   SV               SV
```

Welsh initial nasalization differs from the Irish case in that SV is associated to both positions:

(47) Welsh Initial Nasalization:

```
aspirated stop  ---> voiceless nasal;     unaspirated stop ---> voiced nasal
[spr]                [spr]
  |                    |
A0 Arel              A0 Arel                A0 Arel          A0 Arel
                      \ /                                     \ /
                      SV                                      SV
```

To conclude, (48) summarises the analysis of consonant weakenings, consonant strengthenings, and consonant nasalizations, respectively:

(48) Consonant Weakening: deletion of Aperture Position

Southern Paiute Spirantization:	delete A0
Modern Irish Lenition:	delete A0, else Arel
American English Flapping:	delete Arel

Consonant Strengthening: insertion of Aperture Position

Fula strengthening, Southern Paiute gemination:	insert A0
Intervocalic glide insertion:	insert Arel

Nasalization: Association of SV to Aperture Position:

Southern Paiute and Fula Nasalization:	associate SV to A0
Modern Irish Nasalization:	associate SV to Arel or A0
Modern Welsh Nasalization:	associate SV to A0 and Arel

There are some residual questions concerning the place of articulation. For instance, in Finnish and Welsh, unaspirated labial and coronal short stops alternate with labial and coronal continuants, respectively (see 3a-6a for Finnish and 38a-39b for Welsh), but velar stops exhibit different alternations:

(49) Finnish velar stop gradation:

a.	kurki	--->	kurjen	j =[y]	'stork'
b.	luku	--->	luvun	v =[w]	'chapter'
c.	suka	--->	suan		'brush'

(50) Welsh lenition of the velar stop:

a.	gardd	---> ardd	g= [k]	'garden'
b.	glas	---> las	g= [k]	'blue'

In Irish, labial and velar stops correspond to labial and velar fricatives in lenition environments, but the coronal voiceless stop alternates with the laryngeal fricative and the coronal voiced stop alternates with the dorsal voiced fricative in lenition environments.

(51) Irish lenition of the coronal stops:

a.	teach	t =[t']	'a house'
b.	mo theach	th=[h']	'my house'
c.	dúnaimid	d =[d]	'we close'
d.	dhúnamar	dh=[ɣ]	'we closed'

These are not the only examples of the deviant behaviour of velar stops and coronal stops in phonological processes. I will present a few additional examples in section 3 and present a plausible solution to the coronal and velar asymmetries in section 3.1.

3 Asymmetries and coronal consonants

The first often cited example illustrating the special status of coronal consonants concerns the fact that a segment may be coronal in neutral environments, but take the place of articulation of a following or a preceding consonant:

(52) English: only nasal alveolar stops in coda position undergo place assimilation

a. iN + adequate ---> inadequate
b. iN + possible ---> impossible
c. iN + complete ---> i[ŋ]complete

(53) Korean: oral and nasal alveolar stops in coda position undergo place assimilation

a.	ət-ko	[ək.ko]	'get and'
b.	kət^{h}-poli	[kəp.pori]	'kind of barley'
c.	sinpal	[simbal]	'shoes'
d.	nun-mul	[nummul]	'tear'
e.	son-kalak	[soŋgarak]	'finger'

(54) Dutch: oral coronal stops in onset position of the diminutive morpheme assimilates in place of articulation to a preceding nasal stop

a.	zoon	+ Tje	---> zoontje	'son'
b.	zoom	+ Tje	---> zoompje	'border'
c.	koni[ŋ]	+ Tje	---> koni[ŋ]kje	'king'

The second example of the special status of coronals is that in some (mostly West African) languages spreading of vocalic place features takes place across coronal consonants, but not across consonants with other places of articulation.

Third, in languages with an epenthetic stop, the inserted consonant is a coronal stop rather than a labial or velar stop. The following example is from Axininca Campa:

(55) /i-N-koma-ako-aa-i-ro/ ---> [iŋkomatakotaatiro] 'he will paddle for it again'

Fourth, there are distribution facts which show that coronals are 'special'. For instance:

(56)	a.	Koyukon and Finnish:	/m/ and /n/ occur, but *only* coronal /n/ is found in rhymes
	b.	Kissi:	/m/, /n/, and /ŋ/ occur, but coronal /n/ is *not* found in rhymes

However, similar distribution facts indicate that velars are 'special' too:

(57)	a.	Japanese:	/m/, /n/, and /ŋ/ occur, but *only* velar /ŋ/ is found in rhymes
	b.	English and German:	/m/, /n/, and /ŋ/ occur, but /ŋ/ is *not* found in onsets

In the following section I will discuss a model which accounts for these facts.

3.1 Coronals and velars as the result of a place node with no dependants

Paradis and Prunet (1991:3), among others, state that 'the special status of coronals lies in the fact that they lack specifications for place features in underlying representations'. In a recent paper, Rice (1996) elaborates on this proposal. She argues that coronals as well as velars may result from a C(onsonantal)-Place node with no dependent. Failure to fill in a feature will result in a representation which is phonetically interpreted as a velar consonant (58a). Default fill-in of the unmarked place feature Coronal creates a coronal consonant (58b) and assimilation to an adjacent node will create a labial, coronal, or velar consonant (58c):

```
(58) a.  underlying          no default   surface velar
         Root                              Root
          |                                 |
         C-Place                           C-Place

     b.  underlying          default      surface coronal
         Root                 --->         Root
          |                                 |
         C-Place                           C-Place
                                            |
                                           Coronal

     c.  underlying          assimilation  surface labial
         Root      Root       --->          Root     Root
          |         |                        |        |
         C-Place   C-Place                  C-Place C-Place
                    |                           \     |
                   Labial                        Labial
```

According to Rice (1996), this approach correctly predicts that assimilation may affect velar and coronal consonants. Transparency of coronal (*and* velar) consonants follows from the fact that these segments lack a place specification in languages where these segments are transparent to the spreading of vocalic place features as in the following hypothetical examples:

```
(59)   a     t      i          a     p      i
       V    Root    V          V    Root    V
       |     |      |          |     |      |
       Place Place  Place      Place Place  Place
       |            |          |     |      |
       [phar]       [high]     [phar] [Lab] [high]
```

Epenthesis of coronals follows from syllabification positing the simplest possible structure (a bare root node). Spell-out of Coronal is by default. Because coronals and velars may be unspecified for place of articulation, Rice (1996) predicts languages which have epenthetic coronal consonants and languages which have epenthetic velar consonants.

A possible account under Rice's proposal for underspecified place of articulation of the distribution asymmetries concerning nasals is as follows:

(60) a. Koyukon and Finnish: /n/ may occur in the rhyme because it is unspecified for place, i.e. rhymes do not need (independent) place

b. Kissi: /n/ may not occur in rhyme because it is unspecified for place, i.e. rhymes need place specifications

c. Japanese: /ŋ/ may occur in rhyme because it is unspecified for place (see 60a)

d. English and German: /ŋ/ may not occur in onset because it is unspecified for place, i.e. onsets require place of articulation.

I will now consider the consequences of Rice's proposal for consonant weakening processes. The facts that we have to account for can be summarised as below:

(61) a.	Finnish Gradation:	k:	--->	k;	k	---> y/w/0
b.	Welsh Lenition:	k	--->	k;	k	---> 0
c.	Irish Lenition:				t (d)	---> h (ɣ)

For Finnish and Welsh, I follow Rice (1996) in her proposal that velar consonants may be underlyingly underspecified for place features. When default coronal insertion fails to apply, the underspecified segment is interpreted as a velar consonant:

```
(62) Finnish and Welsh: labial stop     coronal stop    velar stop
                        A0 Amax         A0 Amax         A0 Amax
                         \ /             \ /             \ /
                        C-Place         C-Place         C-place
                           |               |
                         Labial          Coronal
```

The examples that Rice mentions to support her proposal for an underspecified C-place node, all involve stops. I claim here that this is not an accident. Complete obstruction in the oral cavity implies that there is a place of obstruction, i.e. 'closure implies place'. In a theory of aperture positions, this means that where there is an A0 position there is a place of articulation. Representations with an empty C-place node under A0 either (i) trigger default coronal insertion, (ii) or are phonetically interpreted as velar stops, or (iii) get their place specification from an adjacent consonant (in most languages this consonant is an oral or nasal stop, i.e. a segment with an A0 node).

In contrast, release in the oral cavity does not imply a place of articulation. Representations with an empty C-place node which lack an A0 position do therefore not trigger default coronal insertion and may not be phonetically interpretable. Consider in this respect that when the epenthetic consonant is a stop (e.g. example 56 for Axininca Campa), the coronal default rule applies, but when the epenthetic consonant is a continuant (e.g. examples 11a-d for Dutch), the place of articulation depends on the surrounding vowels and if no place of articulation is assigned to the continuant in question, it is not realised (as in, e.g., Dutch *chaos* [xa:ɔs] without a glide after the vowel *a*).

When the segments in (62) undergo Finnish consonant gradation (i.e. the loss of an A0-position) the representations in (63) result. In examples (49a) and (49b), the vowel which follows the resulting Amax position supplies a place feature for that position, but in (49c) the vowel *a* does not supply a place feature and the Amax position is not phonetically interpreted.

```
(63)Finnish:     labial glide     coronal glide     not interpretable
                    Amax             Amax              Amax
                     |                |                 |
                   C-Place          C-Place           C-Place
                     |                |
                   Labial           Coronal
```

The representations that result from Welsh initial lenition are given in (64):

(64) Welsh:

labial fricative	coronal fricative	not interpretable
A_{rel}	A_{rel}	A_{rel}
|	|	|
C-Place	C-Place	C-Place
|	|	
Labial	Coronal	

In Irish, coronal stops are unspecified for place of articulation. In non-lenition environments the feature Coronal is added by default:

(65) Irish /t/ and /d/:

	default:
A_0 A_{rel}	A_0 A_{rel}
\ /	\ /
C-Place	C-Place
	|
	Coronal

Under lenition, the initial aperture position for closure is deleted and what remains in Irish is an aperture position for consonantal release without place specifications. Without vocal cord vibration this results in the sound /h/ and with vocal cord vibration the phonetic interpretation is that of a velar voiced fricative:

(66) Irish /h/: / Irish /γ/:

Irish /h/:	Irish /γ/:
A_{rel}	A_{rel}
|	/ |
C-Place	[voice] C-Place

A consequence of the claim that coronal default insertion only takes place with A0-positions is that surface coronal stops may alternate with velar continuants, but surface velar stops may not alternate with coronal continuants. This prediction seems to be empirically correct.

To summarise the discussion above, (67) schematically represents the effects of consonant weakening of segments underspecified for place:

(67)

		default Coronal	weakened counterpart	place of articulation of weakened counterpart
Finnish	/k/	no	Amax	specification for place from vowel, or deleted
Welsh	/k/	no	Arel	specification for place from vowel, or deleted
Irish	/t/	yes	Arel	laryngeal interpretation
	/d/	yes	Arel-[voice]	velar interpretation

4 Conclusion

The things that happen to consonants in morphological environments are (i) weakening (loss of A-position), (ii) strengthening (addition of A-position), and (iii) nasalization (SV affixation). The chain-like changes in (1) and (2a) of the introduction can be schematically represented as follows:

(68) Changes in the manner of articulation

```
long plosive <---> short plosive <---> fricative/approximant <---> zero
A0 A0 Arel         A0 Arel             Arel                        Ø
```

(69) Changes in the manner of voicing

```
voiceless consonant ---> prenasalized/voiced consonant ---> nasal stop
A0 Arel                  A0 Arel,  or  A0 Arel              A0 Arel
                         |             |                     \  /
                         SV            SV                     SV
```

As a result of consonant weakening, there may be a change of place of articulation for short stops which are underlyingly unspecified for a place feature. Stops which undergo weakening and a change in their place articulation are either coronal or velar short stops. After weakening they are (i) realised as continuants without an oral place of articulation (i.e. as a laryngeal continuant), or (ii) they are realised as a continuant which shares its place of articulation with an adjacent vowel, or (iii) they are deleted. To explain this distribution I have argued that the feature Coronal may be added as a default feature in the case of stops, but not in the case of continuants.

References

Arnott, D.W. (1970), *The Nominal and Verbal Systems of Fula*, Oxford University Press, Oxford.

Ball, Martin J. and Nicole Müller (1992), *Mutation in Welsh*, Routledge, London and New York.

Chomsky, Noam and Morris Halle (1968), *The Sound Pattern of English*, Harper & Row, New York.

Davis, Stuart (1991), "Coronals and the phonotactics of nonadjacent consonants in English", in Paradis & Prunet.

Grijzenhout, Janet (1995), *Irish Consonant Mutation and Phonological Theory*, PhD dissertation, Led, Utrecht University, Utrecht.

Kiparsky, Paul (1992), "Blocking in non-derived environments", in Hargus, S and E. Kaisse (eds), *Studies in Lexical Phonology*, Academic Press, San Diego.

Lieber, Rochelle (1987), *An Integrated Theory of Autosegmental Processes*, State University of New York Press, Albany.

McCarthy, John (1988), "Feature Geometry and Dependency: A Review", *Phonetica* 43, 84-108.

Ní Chiosáin, M. (1991), *Topics in the Phonology of Irish*, PhD dissertation, University of Massachusetts, GLSA Publications, Amherst.

Paradis, C. (1992), *Lexical Phonology and Morphology: The Nominal Classes in Fula*, Garland Publishing, New York.

Paradis, C. and J.-F. Prunet (1991), "Introduction: Asymmetry and Visibility in Consonant Articulations", *Phonetics and Phonology Volume 2: The Special Status of Coronals:* 1-28.

Press, Margret L. (1979), *Chemehuevi: A Grammar and Lexicon*, University of California Press.

Rice, Keren D. (1993), "A Reexamination of the feature [sonorant]: the status of 'sonorant obstruents' ", in Grey Thomason, Sarah (ed.), *Language* 69, 308-344.

Rice, Keren D. (1996), "Default Variability: The Coronal-Velar Relationship", *NLLT* 14, 493-543.

Sapir, Edward (1930), "Southern Paiute, A Shoshonean Language", in *Proceedings of the American Academy of Arts and Sciences*, Vol. 65, No.1.

Skousen, Royal (1972), "On Capturing Regularities", in Peranteau, Paul M. et al., *Papers from the Eighth Regional Meeting Chicago Linguistic Society*, 567-577.

Steriade, Donca (1993), "Closure, Release and Nasal Contours", in Huffman, Marie K. and Rena A. Krakow (eds), *Nasals, Nasalization, and the Velum* (Phonetics and Phonology 5), Academic Press, San Diego, 401-470.

Steriade, Donca (1994), "Complex Onsets as Single Segments: The Mazateco Pattern", in Cole, Jennifer & Charles Kisseberth (eds), *Perspectives in Phonology; Proceedings of the University of Illinois Phonology Conference*, CSLI, Stanford, 203-291.

Vainikka, Anne (1988), "Finnish Consonant Gradation Unified", ms., University of Massachusetts, Amherst.

Wetzels, W.L.M. (1991), "Contrastive and Allophonic Properties of Brazilian Portuguese Vowels", Wanner, D and D.A. Kibbee (eds), *New Analyses in Romance Linguistics*, John Benjamins, Amsterdam and Philadelphia.

Willis, Penny (1986), *The Initial Consonant Mutations in Welsh and Breton*, Indiana University Linguistics Club.

This paper was written when the author was a visiting scholar at the Zentrum für Allgemeine Sprachwissenschaft in Berlin from 1 August until 1 November 1996. Many thanks to Tracy Alan Hall, Ursula Kleinhenz, Sylvia Löhken, Tine Mooshammer, and Marzena Rochon for discussions on life, work, and aperture positions.

In non-mutated contexts, word-initial /θ:, f, θ, l ,r/ are restricted to borrowings (Ball and Müller 1992:78-80).

It is possible that the phonetically long voiceless fricatives /f:/ and /θ:/ are also characterised by the aperture position for laryngeal release which merges with the aperture position for release of the fricative in non-mutation environments. This implies that under lenition they undergo the same process of deletion of the initial aperture position as plosives and sonorants and are realised as the corresponding short fricatives /f/ and /θ/. There is no data available to me that either supports or contradicts this assumption and I therefore leave it open for future research.

Another consonant weakening process in Welsh, known as 'spirantization', turns stem-initial voiceless aspirated stops into voiceless fricatives under certain morphological conditions. I propose here that this process is best analysed as deletion of two stem-initial aperture positions, i.e. the aperture positions for laryngeal release and closure, respectively. Spirantization in Welsh does not apply to unaspirated stops, because deletion of two aperture positions would result in deletion of unaspirated segments and, hence, in an ill-formed empty onset.

The historical development of retroflex consonants in Indo-Aryan*

T. A. Hall
Zentrum für Allgemeine Sprachwissenschaft

1 Introduction

Retroflex consonants in Old Indo-Aryan (i.e. Sanskrit) arose in at least two separate historical stages: (a) the emergence of the retroflex sibilant ṣ (= IPA [ʂ]), followed by (b) the emergence of retroflex stops *ṭ, ḍ, ṭʰ, ḍʰ* (= IPA [ʈ, ɖ, ʈʰ, ɖʱ] (see, for example, Misra 1967). The general assumption is that the sibilant that was the output of (a) developed from Indo-Iranian **š* (IPA [ʃ]), which in turn was derived from Indo-European **s* (i.e. Indo-European **s* → Indo-Iranian **š* → Old Indo-Aryan *ṣ*). If this is the correct sequence of events then the Indo-Iranian sibilant **š* underwent a context-free change to *ṣ* in Sanskrit (i.e. š → ṣ).

Although most previous studies of Sanskrit historical phonology presuppose this context-free change no satisfactory explanation for its occurrence has been proposed. The present study is an attempt to explain why *š* became *ṣ* in the development of Indo-Aryan. The major point I make is that the context-free change š → ṣ was triggered as a repair strategy in order to avoid a phonemic inventory that is otherwise unattested in natural languages.

This article is organized as follows. §2 is devoted to a discussion of the phonetics of the places of articulation that are relevant in my discussion of the Sanskrit historical phonology and to cross-linguistic generalizations regarding possible sibilant contrasts. §3 provides background information on the Sanskrit and Indo-Iranian consonant systems. In §4 I discuss the sources of the retroflex sibilant ṣ in Sanskrit, concentrating on the context-free sound change referred to above. My conclusions are summarized in §5.

2 Phonetics and phonology of postalveolar consonants

In the analysis I posit below for Sanskrit I make reference to the following three places of articulation: retroflex, palatoalveolar, and alveolopalatal. In the present section I discuss the phonetics of sounds produced at these three places and their phonological patterning.

* The material contained in this article was presented at the annual meeting of the Deutsche Gesellschaft für Sprachwissenschaft in Freiburg in March, 1996. I would like to thank the audience for several instructive comments and Haike Jacobs and Sylvia Löhken for important criticisms of an earlier written version.

2.1 Postalveolar places of articulation

The chart in (1) contains IPA symbols for voiced and voiceless stops and fricatives and nasals at seven places of articulation:

(1)	dental	alveolar	retroflex	palatoalveolar	alveolopalatal	palatal	velar
		t, d	ʈ, ɖ		c, ɟ		k, g
		n	ɳ		ɲ		ŋ
	θ, ð	s, z	ʂ, ʐ	ʃ, ʒ	ɕ, ʑ	ç, ʝ	x, ɣ

The most noticeable difference between (1) and most standard IPA charts (e.g. Ladefoged 1990) is that in the former one the three segments [c, ɟ, ɲ] are classified as "alveolopalatal" and not as "palatal". This reclassification is justified by the phonetic fact that sounds like [c, ɟ, ɲ] bear a closer affinity to the fricatives [ɕ, ʑ], as opposed to true palatal fricatives [ç, ʝ], both in terms of place of articulation and the articulator involved in their production. See Recasens (1990: 272), Keating (1991: 36), and Hall (1997: §1) for further discussion on this point.

I employ "postalveolar" here and below as a cover term for the three places of articulation "retroflex", "palatoalveolar", and "alveolopalatal". Retroflex sounds are articulated with the tongue tip or the underside of the tongue, i.e. they are apical or sublaminal. In contrast, sounds produced in the palatoalveolar and the alveolopalatal places utilize the tongue blade and are therefore laminal.[1] The term "alveolopalatal" (sometimes referred to as "prepalatal") describes the place of articulation in the postalveolar region between the palatoalveolar and the palatal places where fricatives and affricates like [ɕ, ʑ, tɕ, dʑ] in languages like Polish are produced (see Pullum & Ladusaw 1986: 31). Alveolopalatals, like palatoalveolars, are always laminal.

2.2 Inventory generalizations

Many languages, especially those indiginous to India and Australia, contrast a retroflex and an alveolopalatal noncontinuant.[2] Postalveolar fricative systems can contrast either a retroflex sound with a palatoalveolar, or a retroflex with an alveolopalatal. Representative examples of occurring postalveolar noncontinuant and fricative contrasts are given in (2)(a) and (b) respectively:

---

[1] Phoneticians sometimes assume that palatoalveolar sounds can be apical (e.g. Catford 1988: 90-91). I analyze "apical palatoalveolars" as phonologically retroflex. This view derives support from the fact that both sets of sounds are phonetically similar (i.e. both are postalveolar and apical) and that no language contrasts the two (see also Maddieson 1984, and Hume 1992).

[2] The generalizations in this section are based primarily on Maddieson (1984) and Hall (1997).

(2) Possible postalveolar contrasts

(a) Noncontinuants

/ʈ, c/ Pitta-Pitta (Dixon 1980), Tamil (Christdas 1988)

/ɳ, ɲ/ Pitta-Pitta (Dixon 1980), Tamil (Christdas 1988)

(b) Fricatives

/ʂ, ʃ/ Tolowa (Bright 1964), Toda (Emeneau 1984, Shalev et al 1993), Basque (Hualde 1991)

/ʂ, ɕ/ Pekingese (Ladefoged & Wu 1984)

What all of the systems in (2) have in common is that one of the sounds is apical and the other laminal.

One generalization that I assume to be exceptionless is summarized in (3):

(3) No language can contrast palatoalveolars and alveolopalatals.

The generalization in (3) appears to be an absolute universal because no attested examples exist with surface oppositions of these two types (see Maddieson 1984, and Hall 1996). According to Stevens (1989) both palatoalveolars and alveolopalatals lie in an acoustically stable area and are thus not expected to contribute to phonemic contrasts. This gap makes phonetic sense for noncontinuants if palatoalveolar stops and nasals are nonoccurring segment types (see Lahiri & Blumstein 1984 for discussion).[3] The analysis I propose for Sanskrit in §4 below presupposes that fricatives also obey (3), that is, contrasts like the ones in (4) are nonoccurring:

(4) Nonoccurring contrasts:

/ʃ, ɕ/, /ʒ, ʑ/ etc.

I account formally for the nonexistent systems in (4) by positing that a single set of distinctive features describes both palatoalveolars and alveolopalatals. Following standard approaches to distinctive features (e.g. Chomsky & Halle 1968, Keating 1988) I analyze palatoalveolar and alveolopalatal segments as in (5):

---

3 According to Mohanan & Mohanan (1984), Malayalam has a palatoalveolar nasal. However, this language does not violate (3) because this sound does not contrast with another laminal postalveolar nasal.

(5)	s	ʂ	ʃ, ɕ
[coronal]	+	+	+
[anterior]	+	-	-
[distributed]	-	-	+

The feature [distributed] accounts for the distinction between apical and laminal sounds: [-distributed] is interpreted to mean apical (or sublaminal) and [+distributed] laminal. Since palatoalveolars and alveolopalatals are identical phonologically whetever feature(s) that distinguish the two catergories on the phonetic surface are assigned in the phonetic component. This means that [+coronal, -anterior, +distributed] (in addition to [-voice, +continuant]) is phonetically realized in some languages as [ʃ] and in other languages as [ɕ].

Since retroflexes are apical and both palatoalveolars and alveolopalatals are laminal, the statement in (3) implies (6):

(6) If a language contrasts two postalveolar sounds then one will be apical and the other laminal.

Note that the languages in (2) obey (6). Several languages have been discussed in the literature that appear to violate (3) and (6), all of which will be argued below to be only apparent counterexamples.

In some of the languages that seem to contradict (3) and (6) the two segments also differ in terms of some manner feature, or some secondary articulation. Since (3) and (6) are intended to describe oppositions between alveolopalatals and palatoalveolars that agree in all other features these examples do not contradict my claim that palatoalveolars and alveolopalatals are identical phonological entities. For example, Ladefoged (1964) lists eight examples in his survey of sixty two West African languages with palatoalveolars and alveolopalatals, but the two sounds always differ in terms of a secondary articulation, e.g. Twi contrasts the voiceless alveolopalatal fricative /ɕ/ and the voiceless *labialized* palatoalveolar fricative /$ʃ^w$/. Another similar example is Swedish, which is sometimes said to contrast [ɕ] and [ʃ] (see Campbell 1991: 1289). However, Ladefoged & Maddieson (1996), citing Lindblad (1983), note that the Swedish "palatoalveolar" is multiply articulated. This segment is highly rounded, labiodental, velar or velarized and the source of friction is between the lower lip and the upper teeth. In addition, the tongue body is raised and retracted towards the velum to form a fairly narrow constriction. Thus, Swedish "[ʃ]" is more accurately [ɧʷ]. What this means is that the alleged contrast in Swedish does not involve [ʃ, ɕ], but [ɧʷ, ɕ].

Other examples of languages with apparent palatoalveolar vs. alveolopalatal contrasts involve oppositions between *retroflexes* and alveolopalatals. This is certainly the case in Polish, which is often erroneously referred to as a language with surface contrasts between "[ʃ]" and [ɕ]. (See, for example, Spencer 1986, and Dogil 1990, who employ these IPA symbols). However, Polish "[ʃ]" is phonetically very different from English [ʃ], which is a "true" palatoalveolar because it is produced with the tongue blade.[4] Specifically, Polish "palatoalveolars" are pronounced with a shorter length of constriction (and with velarization). Because the Polish sounds are much closer phonetically to "true" postalveolar apical sibilants, many linguists correctly refer to them as retroflex (e.g. Keating 1991: 35-36, Hume 1992: 104ff).[5]

One language with reliable articulatory data known to me in which a true palatoalveolar [ʃ] and a true alveolopalatal [ɕ] exist on the surface is Nantong Chinese (Ao 1993). However, as Ao (1993: 49-53) points out, [ʃ] and [ɕ] never contrast in Nantong, neither in the underlying representation, nor on the phonetic surface. Nantong Chinese therefore does not violate (3) and (6) because both of these generalizations hold for sounds that contrast.

Occasionally one encounters descriptions of languages that are said to contrast two postalveolar laminal sibilants within the same series, but these analyses often lack the phonetic description necessary to determine whether or not (3) and (6) are really falsified. For example, several Uralic languages are said to have oppositions between palatoalveolars and alveolopalatals (see Wurzel 1975: 175; and Veenker 1987, although neither linguist uses the terms "palatoalveolar" and "alveolopalatal"). Rédei (1975) shows that Syranisch has four postalveolar sibilants he transcribes as " š ", "s' ", " ž ", and "z' ". Both *s'* and *z'* are described as a palatalized š and ž respectively (p. 103), which implies that they are laminal. If *š, ž* are true palatoalveolars then this language would falsify (6), but Rédei does not discuss the phonetics of *š, ž* in enough detail to determine this. Similar systems are reported for Livonian (Viitso 1975), Moksha Mordvinian (Feoktistov 1984; Veenker 1987: 39) and Erza Mordvinian (Rédei 1984; Veenker 1987: 38). Unfortunately, no mention is made in any of these sources whether or not the nonpalatalized *š* and *ž* are true (laminal) palatoalveolars, or if they are apical (and hence retroflex [ʂ, ʐ]).

4 See also Jones & Ward's (1969: 134) description of the difference between English [ʃ] and Russian "[ʃ]".

5 See also Chomsky & Halle (1968: 314), who analyze Polish "palatoalvoelars" as [-distributed].

Should the examples cited in the previous paragraph (or other ones unknown to me) turn out to involve true contrasts between palatoalveolars vs. alveolopalatals then the generalizations in (3) and (6) will have to be demoted to markedness statements. That is, a system with /ʂ, ɕ/ would be less marked than one with /ʃ, ɕ/. Importantly, the analysis I propose in the following section for Sanskrit will hold regardless of whether or not oppositions of the type /ʃ/ vs. /ɕ/ are impossible, as expressed in (3), or simply highly marked.

3 Old Indo-Aryan

Old Indo-Aryan[6] (represented below as Sanskrit) had the following inventory of consonants and glides:[7]

(7)

p	t	ṭ	c	k	
p^h	t^h	ṭ^h	c^h	k^h	
b	d	ḍ	j	g	
b^h	d^h	ḍ^h	j^h	g^h	
φ	s	ṣ	ś	x	h
m	n	ṇ	ñ	ŋ	
v	l	r	y		

(7) contains the phonetic symbols traditionally employed in Indo-Aryan historical phonology. In that system aspirated and murmured stops are reflected uniformly, with the raised [^{h}]. The four segments *c, c^h, j, j^h* were probably stops as opposed to affricates (i.e. IPA [c, c^h, ɟ, ɟ^ɦ]; see Allen 1953: 52), *ṭ, ḍ, ṭ^h, ḍ^h, ṣ, ṇ, r* were retroflex (i.e.

---

6 Old Indo-Aryan (henceforth OIA) corresponds roughly to the period 1500 BC-600 BC (Chatterji 1927: 17, Masica 1991: 51) and is considered to be the earliest stage of Indo-Aryan. OIA is one branch of Indo-Iranian, the others being Nuristani (which is sometimes referred to by the derogatory term "Kafiri", see Masica 1991: 465) and the reconstructed language Old Iranian. The latter language subsequently split into the two (attested) languages Avestan and Old Persian. See Bartholomae (1883), Brugmann (1897a, b), Ghatage (1962), and Mayrhofer (1989) for a discussion of the historical development of the Indo-Iranian daughter languages.

7 The system of consonants and glides in (7) corresponds to both Early (i.e. Vedic) and later (i.e. Classical) Sanskrit. The six places of articulation in (7) are labial, dental, retroflex, alveolopalatal, velar, and glottal. Traditional descriptions of Sanskrit consonants include Whitney (1889), Wackernagel (1896), Allen (1953), Burrow (1955), and Thumb (1958).

Vedic Sanskrit also had the retroflex laterals ḷ and ḷh, which were the intervocalic realization of *ḍ* and *ḍh* respectively (Masica 1991: 161). Sanskrit *φ, x* and *ñ* are usually analyzed as allophones of other segments because of their limited and hence predictable distribution (see Cho 1990: 162)

IPA [ʈ, ɖ, $ʈ^h$, $ɖ^ɦ$, ʂ, ɳ, ɻ]),[8] and *ś* was a voiceless postalveolar laminal sibilant, which I assume to be equivalent to IPA [ɕ].[9]

At an earlier stage in Indo-Iranian prehistory the language probably had the consonants and glides in (8):

(8)	p	t	c	k		
	p^h	t^h		k^h		
	b	d	j	g		
	b^h	d^h	$ɟ^h$	g^h		
		s	ś			h
	m	n				
		l				
		r				
	w		y			

The inventory in (8) is similar to the reconstructed inventory for Proto-Indo-Iranian posited by Ghatage (1962: 83), and Misra (1967: 42). The major difference is that Ghatage's and Misra's reconstructions also include the palatoalveolar affricates *č, ǰ, $ǰ^h$* (which derived from Indo-European alveolopalatals **k̂, ĝ, $ĝ^h$* respectively). In (8) I assume that the sound change that converted *č* to *ś*, and merged *ǰ, $ǰ^h$* with *j, h* respectively had already occurred. The latter sound changes as well as alternate chronologies will be discussed in greater detail below.

A comparison of (7) and (8) reveals that there were no retroflex sounds in the latter language. The question I consider below is how and why retroflex sounds arose historically in Indo-Aryan. I demonstrate below that an answer to this question can be tied in with the generalizations in (3) and (6) above that govern synchronic systems.

---

8 For typographical reasons Sanskrit retroflex *r* is not usually transcribed as *ṛ* because the latter symbol represents the syllabic *r*, which also existed in the language.

9 Since most sources agree that *ś* was a voiceless postalveolar laminal sibilant (see below), then this sound was either IPA [ʃ] or IPA [ɕ]. (*ś* could not have been a true palatal fricative (i.e. [ç]), pace Chatterji (1960: 76), because this sound is not a sibilant). Phonological evidence can be adduced that *ś* was IPA [ɕ]: In Sanskrit dentals became alveolopalatal before alveolopalatals (e.g. ut+cariti → uccariti 'rises'; Cho 1990: 66). Importantly, Sanskrit *ś* surfaces in the same environment , e.g. tatas + ca → tataśca 'and then'. If *ś* and *c* have the same place of articulation then they are both alveolopalatal (recall (1)).

Linguists who assume that Sanskrit *ś* was IPA [ʃ] include Allen (1953: 20) and (probably) Whitney (1889: 22), who describes *ś* as "the usual and normal sh-sound". Thumb (1958: 205) states that *ś* is "a palatalized *š* (German sch), close to German *ç*", a description that comes close to IPA [ɕ]. That few scholars (if any) have stated explicitly that Sanskrit *ś* was IPA [ɕ] can be attributed to the fact that they were unaware that the IPA table distinguishes two postalveolar laminal fricatives.

The emergence of retroflex consonants in Sanskrit is often ascribed to areal influences from Dravidian (e.g. Bloch 1930: 731-733, Emeneau 1954: 284). While there is no questioning the fact that such borrowings took place in the development of Sanskrit, most instances of Sanskrit retroflex segments can be shown to have been the result of sound changes internal to Indo-Aryan[10] (see Wackernagel 1896, Brugmann 1897a, b, Burrow 1955: 96, Thumb 1958: 281-282, and Misra 1967: 64-65). For example, most instances of Sanskrit *ḍ, ḍʰ* derive from a sequence of retroflex sibilant plus corresponding dental stop. This can be illustrated with the data in (9) (from Misra 1967: 68-69):[11]

(9)	*Sanskrit*		*pre-Sanskrit*	*gloss*
	ni:ḍa	<	*niṣda	'nest'
	mi:ḍha	<	*miṣdha	'reward'
	voḍhum	<	*vaṣdhum	'to carry'
	leḍhi	<	*laṣdi	'licks'

In the examples in (9) the retroflexion features assimilated to the following dental stop, thereby producing *ḍ, ḍʰ*. These segments subsequently became phonemic in the pre-Vedic period when the preceding retroflex sibilant deleted.[12]

[10] Two examples of Dravidian loanwords in Sanskrit containing *ṭ* are *kuṭila* 'crooked', and *kuṭi* 'hut, cottage' (Burrow 1955: 97). Some loanwords with *ṭ, ḍ* are of uncertain origin (see Masica 1991: 157-158), e.g. *kuṭumba* 'household', *ḍimbha* 'newborn child', and *ta:ḍa* 'blow'. Loanwords containing *ṇ* were particularly common in *ṇḍ* clusters (Masica 1991: 160), e.g. *tuṇḍa* 'beak'. See Burrow (1945, 1946) for additional examples.

[11] The general assumption is that pre-Sanskrit *ṣ* (in (9)) had two allophones, *ṣ* and *ẓ*, where the latter sound surfaced only before voiced stops (Misra 1967: 65). Hence, the retroflex sibilant in the pre-Sanskrit forms in (9) was probably *ẓ* phonetically.

The reconstructed forms for the four words in the pre-Sanskrit column in (9) (as well as others of the same structure) are also presupposed by Burrow (1955: 93-94) and Masica (1991: 157). One argument that the retroflex stop in the Sanskrit words in (9) derives from an earlier sibilant is that the cognates in related non-Indo Aryan languages often have a sibilant, e.g. English *nest*, Avestan *mižda* 'reward'.

[12] Many instances of Sanskrit *ṭ, ṭʰ, ṇ* can also be shown to have Indo-European roots. For example, *ṭ* derives from a sequence of ṣ + s or ś + s in final position (Burrow 1955: 96). For another internal source of *ṭ* see below. Sanskrit *ṇ* derived from *n* after a syllabic or nonsyllabic *r*, or *ṣ* anywhere in the word when a vowel, glide or nasal immediately follow and a dental, retroflex, or alveolopalatal does not intervene (Masica 1991: 160), e.g. the *ṇ* in *brahamaṇa:* 'Brahaman'.

There are also many attested examples of "spontaneous" retroflexion, whereby original dentals became retroflex regardless of the environment, i.e. the context-free change t, d, tʰ, dʰ, n, l, s → ṭ, ḍ, ṭʰ, ḍʰ, ṇ, ḷ, ṣ occurred; see Burrow (1971). Two examples of Sanskrit words containing retroflex consonants that underwent spontaneous retroflexion are *sthu:ṇa* 'column' (cf. Avestan *stu:na*), and *aṭ* 'to wander', which was earlier *at* (see Burrow 1955: 97).

Pre-Sanskrit *ṣ* had a similar internal history in the sense that it derived both from Indo-European sounds, namely IE **s* after *r, u, k, i,* and IE voiceless alveolopalatal stop before *t* (see below for discussion).

The analysis described in the preceding paragraph presupposes that the retroflex obstruents in (7) came about in (at least) two sequential stages: [13]

(10)(a) the development of the retroflex sibilant *ṣ*

(b) the development of the retroflex stops *ḍ, ḍʰ, ṭ, ṭʰ*

Since *ḍ ḍʰ* in examples like the ones in (9) were triggered by an adjacent *ṣ,* the latter segment clearly arose prior to the former two.

In the following paragraphs I concentrate on the emergence of the retroflex sibilant ṣ (i.e. (10)(a)) and show that this development can only be understood within the context of the inventory generalizations discussed in §2.

4 The development of Sanskrit *ṣ*

Since I argue below that there is a connection between the emergence of Sanskrit *ṣ* and the segment *ś* in the inventory in (8), I begin this section by tracing the development of the latter sound.

The Indo-European alveolopalatal stops **k̂, ĝ, ĝʰ* not followed by *t* became *ś, j, h* respectively in Sanskrit.[14, 15] These changes are usually assumed because of comparative data like the ones in (11) (from Burrow 1955: 72), where Latin represents the centum languages, and Sanskrit and Avestan the satem languages:

(11)	*Latin*	*Sanskrit*	*Avestan*	*gloss*
	centum	śatám	satəm	'hundred'
	genu	ja:nu	za:nu	'knee'
	hiems	hima	zima	'snow'

13 Misra (1967: 73) argues that *ṭʰ* was the last among the retroflex stops to become phonemic; hence, (10)(b) is not intended to imply that *ḍ, ḍʰ, ṭ, ṭʰ* all entered the language simultaneously. The important point is that there was a pre-Sanskrit stage in which the only retroflex segment present was *ṣ*.

14 This development is characteristic of the so-called satem languages. See Wackernagel (1896: 227: 229), Brugmann (1897a: 556), Bloomfield (1911), Edgerton (1946: 6-7), Thumb (1958: 286-288), and Allen (1978) for traditional descriptions of these changes. Solta (1965) questions the basic centum vs. satem division. See Tischler (1990) for a defense of the traditional theory.

15 Another source of Sanskrit *ś* is *s* before *c* (recall note 9), as in *tatas + ca* [tataśca] 'and then' (Cho 1990: 66). IE voiceless alveolopalatal stops after *s* surfaced in Sanskrit as *cʰ* in Sanskrit words like *cha:ya:* 'shadow' (Misra 1967: 53-54)

The initial consonant in the words in (10) derives from IE $^*\hat{k}, \hat{g}, \hat{g}^h$ respectively.

Most investigators have argued that IE $^*\hat{k}, \hat{g}, \hat{g}^h$ underwent an affrication stage before surfacing in Sanskrit as *ś, j, h* (e. g. Morgenstierne 1945: 225-233; Burrow 1955: 73; and Misra 1967: 26-27). These developments are summarized in (11):[16]

(12) IE $^*\hat{k}, \hat{g}, \hat{g}^h \rightarrow$ č, $\check{j}, \check{j}^h \rightarrow$ Sanskrit ś, j, h

Let us assume that immediately prior to the emergence of *ṣ*, the alveolopalatal sibilant *ś* had already entered the language, i.e. *č* had become *ś*, a stage in which the language had the inventory of consonants and glides in (8) above. One might alternatively argue that *ś* arose at a later stage. I consider this possibility below and show that this chronology is also compatible with my analysis. However, for the remainder of this section, I assume that prior to the emergence of *ṣ*, the alveolopalatal sibilant *ś* was in the language.

The general consensus is that Sanskrit *ṣ* had two historical sources: (a) IE **s* after *r, u, k, i* unless an *r* follows, and (b) IE voiceless alveolopalatal stop before *t*. Consider first the data in (13) (from Burrow 1955: 79), which illustrate development (a):[17]

(13)	*IE*	*Sanskrit*	*Avestan*	*gloss*
	*s	víṣa	viša	'poison'
	*s	śúṣka	huška	'dry'
	*s	dákṣiṇa	dašina	'right hand'

The comparative method demands that the retroflex sibilant in these and similar Sanskrit words derive from IE **s* because the corresponding etymons in other Indo-European daughter languages contain *s* (e.g. Sanskrit *mu:ṣ* vs. English *mouse*). The Avestan forms in (13) illustrate that IE **s* after *r, u, k, i* surfaced as palatoalveolar *š* (i.e. IPA [ʃ]) in the Iranian branch of Indo-Iranian.[18]

---

16 Burrow (1955: 73) states that the affrication stage affected all satem languages. Brugmann (1897a: 242ff.) assumes that IE $^*\hat{k}$ changed directly into *ś* in Indo-Iranian.

17 The Sanskrit data in (13) have generated a vast literature over the past hundred years. For historical analyses the reader is referred to Whitney (1889: 180-185), Wackernagel (1896: 230-235), Brugmann (1897b: 728ff.), and Thumb (1958: 305). Synchronic treatments include Zwicky (1970), Vennemann (1974), and Cho (1990: 85-89).

18 This generalization also holds for Old Persian, e.g. Sanskrit *dhṛṣṇoti* 'dares' vs. Old Persian *adaršnauš* 'he dared'.

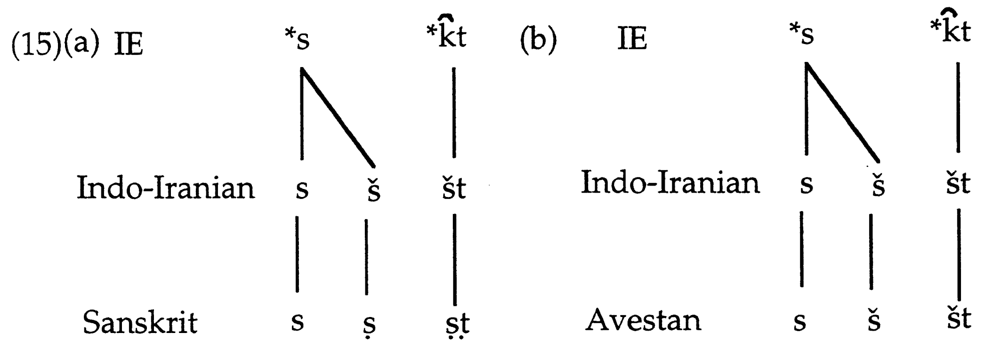
With some qualifications (see below) retroflex sounds are basically nonexistent in Middle and New Iranian dialects (see Gray 1902: 136 and Schmitt 1989: §4). Hence, in New Iranian languages, e.g. Farsi (Jensen 1931, Lambton 1961, Boyle 1966, Majidi 1986) and Kurdish (Mackenzie 1961) the postalveolar sibilants are palatoalveolar, as opposed to retroflex. (One phonetic study to my knowledge (i.e.

The second source of Sanskrit *ṣ* is IE **k̑* before *t*. Thus, consider the data in (14), which have been taken from Misra (1967: 30):

(14) *IE*	*Sanskrit*	*Avestan*	*gloss*
*k̑t	aṣṭau-	ašta-	'eight'
*k̑t	vaṣṭi	vašti	'wishes' 3rd. sg. pres. ind.

As illustrated in (14), Sanskrit ṣṭ in these and similar words corresponds to Avestan *št* (see also Bloomfield 1911, and Burrow 1955: 96-97).

The general assumption in Sanskrit historical phonology - which will be defended below - is that IE **s* and **k̑* in examples like the ones in (13) and (14) underwent an intermediate shift to *š*, prior to its emergence in Sanskrit as *ṣ* (e.g. Wackernagel 1896: 230, Brugmann 1897b: 638, 728, Misra 1967: 28-30, and Mayrhofer 1989: 8). Assuming that the changes from IE **s* after *r, u, k, i* and IE **k̑t* to palatoalveolar occurred sometime during the Indo-Iranian period, the development of these Indo-European sounds into Sanskrit and Avestan is illustrated below in (15)(a) and (b) respectively:



The palatoalveolar sibilants that emerged during the Indo-Iranian period remained palatolavoelar in Avestan (and Old Persian). In contrast, *š* changed to *ṣ* in Sanskrit.

Smirova & Ejubi 1985: 96) confirms that the postalveolar sibilants in Kurdish are palatoalveolar, as opposed to retroflex).

A detailed description of the phonology of the New Iranian languages is contained in Schmitt (1989). In contrast to the generalization in the preceding paragraph, several modern East Iranian languages have retroflex consonants, e.g. the Southwest dialects of Pashto, Munǰi, Yidgha, Wakhi (see Penzl 1955, Skjærvø 1989a: 371, 1989b). However, the general consensus is that these sounds were not inherited from Indo-Iranian, but that they instead arose at a later stage in the development of the respective language. For example, Geiger (1894: 217) and Morgenstierne (1927: 77-79) show that voiceless retroflex sibilant in Pashto (a sound symbolized by both authors as "ṣ̌") corresponds to Avestan sr (e.g. Pashto *ṣ̌a* 'good' Avestan *srao;* Pashto *ṣ̌na* 'hip bone', Avestan *sraoni* 'hip'). See also Skjærvø (1989a, b), who agrees that retroflex segments in Pashto (and in the other East Iranian languages) are a relatively late development.

Three arguments can be adduced for the intermediate stage in (15) with the palatoalveolar sibilant *š*. First, the development of IE **s* into *š* after *r, u, k, i* occurred not only in Avestan and Old Persian, but also in Slavic (where **š* surfaced as the velar fricative [x]) and to a limited extent in Baltic (see Martinet 1951, Andersen 1968). If there were indeed an early stage of pre-Indo-Iranian history when this language family and Slavic shared this common development (see Burrow 1955: 18; 79-80; Thumb 1958: 305) then the comparative method demands that the output of this change be *š* and not *ṣ*. What this means is that the sibilant in the Avestan words in (13) and (14) represents the more ancient sound than the *ṣ* in the corresponding Sanskrit forms and that *ṣ* arose out of *š* in Indo-Aryan. Second, and more importantly, the change from IE **k̑* to *š* (with an intermediate affrication stage to *č*) before t makes more sense phonetically than the alternative, which would have IE **k̑t* convert directly into *ṣṭ* in Sanskrit (and to *št* in Avestan). The reason the development from IE *k̑t → čt → št (→ ṣṭ) is more plausible than IE *k̑t → ṣṭ is that the alveolopalatal stop, the palatoalveolar affricate and the palatoalveolar sibilant are all postalveolar and laminal. Hence, these changes from *k̑t → čt → št simply involve manner features alone, i.e. stop → affricate → fricative. If sound change is gradual then these developments make more sense than a discrete change that involves manner *and* place features, i. e. *k̑t → čt → ṣṭ. Third, the change from palatoalveolar affricate to the corresponding fricative (e.g. č → š) is attested both diachronically (e.g. the development of Old French *č* into New French *š;* Hock 1986: 130) and synchronically (*č* becomes *š* word-finally and before stops, nasals and laterals in Luiseño; Munro & Benson 1973). In contrast, the alternative development (i.e. č → ṣ) is unattested.

The Indo-Iranian developments in (15) (which are reflected in the Avestan forms in (13) and (14)) are expressed in (16)(a) and (b): [19]

(16)(a) IE *s → š / r, u, k, i ___
(b) IE *k̑ → š/ __ t

Again, some linguists have assumed that (16)(a) predated Indo-Iranian because a similar sound change converting IE **s* into *š* occurred in Slavic. The sound change in (16)(b) is traditionally viewed as one that added the phoneme *š* in Indo-Iranian because *š* contrasted with *s* (< IE **s*) before *t* (Misra 1967: 30). This can be illustrated with the Indo-European word **Hesty* 'is' 3rd. sg. pres. ind., which surfaced as *astiy*

[19] IE **z* became *ž* in the same environment as (15)(a). Like its voiceless counterpart, *ž* also became retroflex, but was never phonemicized because it was later deleted (see Misra 1967: 67).

in Old Persian, and *asti* in Sanskrit vs. Sanskrit *vaṣṭi* 'wishes' 3 sg. pres. ind., Avestan *vašti* (< IE **wekty*).

One question that has vexed historical linguists for decades is why IE *$\hat{k}$ became *ṣ* in Sanskrit only in *$\hat{k}$t clusters and *ś* elsewhere (see Bloomfield 1911 for discussion). (That IE *$\hat{k}$ in the data in (14) could not have become *ś* by (16)(b) is obvious, otherwise the output would have merged with the *ś* segments that derived from IE *$\hat{k}$ in other positions). I assume a stage in IE prehistory in which š̌ was an allophone of *ś* in the sense that the former sound only occurred before *t* and the latter in all other positions (see also Burrow 1955: 90).

The sequence of events summarized in (15) presupposes that the change from palatoalveolar to retroflex was a later Indo-Aryan development (for a similar view see Wackernagel 1896: 165, 229, Brugmann 1897a,b: 728, Burrow 1955: 90; 95-96, Misra 1967: 28-29; 65, and Mayrhofer 1989: 8). Thus, if OIA ṣ derived from Indo-Iranian š, then the following context-free sound change must have occurred sometime in the pre-Vedic period:

(17) š → ṣ

In contrast to the assimilatory change responsible for the development of the retroflex stops *ḍ*, *ḍ*^{h} in (9) above, (17) was context-free. What is more, (17) was classically Neogrammarian in the sense that it was exceptionless; that is, every palatoalveolar shifted to retroflex.

In view of the fact that (17) has no obvious phonetic motivation, and that the output was a more marked segment phonologically than its input (see Maddieson 1984: 44-45 who shows that [ʃ] is much more common in synchronic systems than [ʂ]) the obvious question to ask is why this sound change would occur at all.

One could appeal to the existence of Dravidian loanwords with retroflex consonants in them as an explanation. On these same lines one might contend that there must be a connection between such loanwords and the change in (17) because the related languages Avestan and Old Persian had neither (Misra 1967: 63). Although intuitively plausible, the Dravidian loanword hypothesis suffers from several discrepancies. First, the earliest non Indo-Aryan loan words with retroflex consonants in them contained retroflex stops and nasals but apparently no examples are attested with retroflex sibilants (see Burrow 1945, 1946).[20] Hence, the puzzle is

[20] Burrow (1945) lists four Dravidian loans in Sanskrit which contain *ṣ* (e.g. *maṣi* 'ink, lampblack', p. 10). However, none of the cognates in the Dravidian languages contains [ṣ], e.g. Sanskrit *maṣi* is cognate with the Dravidian language Kui *ma:si* 'dirt'. The *s* in this and similar borrowed Dravidian

why the context free change in (17) would affect the dental sibilant and not the dental stops or nasal. Second, the Dravidian loanword hypothesis cannot account for the fact that (17) was exceptionless.

Hock (1986: 79) claims that retroflexion can develop in a nonassimilatory (i.e. context-free) way, tentatively citing the development of Latin ll to [ḍḍ] in Sicilian and Sardinian dialects as an example. The geminate retroflex stop [ḍḍ] was preceded by a stage with the palatoalveolar (in his terminology "palatal") affricate [ʤ]. The retroflex affricate [ḍẓ] developed out of [ʤ] by a context-free sound change because "retroflex is a possible variant of palatal articulation in the sibilants...". By this Hock means that palatoalveolar sibilants can be pronounced either with the tongue tip up, or down, where he regards the former pronunciation as retroflex.

There are good reasons for rejecting such an explanation for (17). Although the "tip-up" pronunciation of English palatoalveolars, whereby the tongue tip is raised to the dental/alveolar region, is possible (cf. Ladefoged & Maddieson 1996: 149-150), the tip does not make contact *behind* the alveolar ridge, which is how retroflex consonants are produced. Thus, Hock's claim that retroflex is an optional pronunciation of palatoalveolars is false.[21]

The explanation I offer below for the sound change in (17) can only be understood by considering the system of sibilants in the respective stages. Since the two developments in (16)(a) and (b) predated the one in (17), the following inventory change occurred:

(18) stage 1: /s, ś/ → stage 2: /s, š, ś/ → stage 3: /s, ṣ, ś/

words apparently underwent spontaneous retroflexion (recall note 12) and converted to ṣ. That none of Burrow's examples of Dravidian loanwords in Sanskrit with ṣ contained ṣ in the cognate Dravidian languages is hardly surprising in view of the fact that Dravidian languages typically contain no sibilants at all (Zvelebil 1990: 8) and that Proto-Dravidian has been reconstructed without any such sounds (Zvelebil 1970: 76).

[21] In contrast, there may be good reasons to believe that retroflex r is an optional pronunciation of the alveolar *r*. For approximant *r* sounds (IPA [ɹ] = alveolar and [ɻ] = retroflex) the tongue does not come as close to the alveolar ridge as it does for dental/alveolar stops and fricatives like [t, d, s, z]; since the tongue tip is not inhibited in any way, the curling back of the tongue tip behind the alveolar ridge (i.e. [ɻ]) is a conceivable variant pronunciation for [ɹ]. This might account for the development of the Indo-European *r*, which was presumably alveolar, to retroflex in Sanskrit. This change, like the one in (17), was context-free, since all *r*'s became retroflex. (That Sanskrit *r* was retroflex is uncontroversial because this segment caused a following dental *n* to become retroflex; recall note 12. See also Whitney 1889: 47, who notes that Sanskrit *r* is retroflex, i.e. "lingual" in his terminology. Since this sound was not trilled, it was most likely an approximant.) I assume that a similar context free change of a dental *r* to a retroflex *r* occurred in American English.

Stage 1 in (18) corresponds to the system in (8) above for Indo-Iranian, and stage 3 to the Sanskrit system in (7). The palatoalveolar sibilant *š* entered the language by (16)(a) and (b) later in the Indo-Iranian period, thereby producing the system at stage 2, with the three sibilants /s, š, ś/.

Recall from §2 above that no language is attested with the phonemic system of sibilants at stage 2 in (18) (i. e. /š , ś/ = IPA /ʃ, ɕ/). Given this generalization, the transition from stage 2 to stage 3 (by (17)) went into effect in order to eliminate a nonoccurring system and to bring it in line with the generalization in (6). This explanation also accounts for the fact that the output of (17) was a more marked segment than the input. Specifically, an increase in segmental markedness is tolerated in order to alleviate the violation to (3), which, as an absolute universal, takes precedence. Consider now a concrete example. The Sanskrit word *viśati* 'settles' (3 sg. pres. ind.) (from Misra 1967: 66) contained *ś* as early as stage 1, a stage when the *š* in the Avestan (i.e. Indo-Iranian) word *viša* 'poison' (see (13)) was still *s*. When **visa* became *viša* at stage 2 (via (16)(a)) the sound change in (17) was triggered because *ś* and *š* contrasted. [22]

If (3) is a true absolute universal as opposed to a strong cross-linguistic tendency then stage 2 in (18) never really corresponded to a synchronic stage in the language. What this means is that the two developments in (16)(a) and (b), caused (17) to enter the language at the same time. Should (6) prove to be not an absolute universal, but instead a statement reflecting markedness, then this would imply that stage 2 in (18) did correspond to a synchronic system which was eliminated because it was highly unstable.

A question of equal importance is why (17) was exceptionless. The general assumption is that sound changes that are not phonetically motivated (e. g. by being context free) exhibit lexical diffusion effects (see Chen & Wang 1975, Labov 1981, Kiparsky 1988). Indeed, (17) can be contrasted with context-free sound changes in other languages, or the spontaneous retroflexion of dentals referred to above that occurred in Indo-Aryan, that were lexically gradual. One could presumably argue that (17) *was* lexically gradual but that we cannot know this because this sound change was so ancient that the exceptions that used to exist were gradually eliminated and therefore never surfaced, even in the earliest Vedic texts. In light of

---

22 Note that the explanation offered above for (17) is not a classic push chain shift (in the sense of Martinet 1981). In a true push chain shift there is a direct causality between two rules (i) and (ii), where (i) A → B, and (ii) B → C (and A is not C). The reason the Sanskrit developments do not constitute a push chain shift is that rule (17) above (which would correspond to (ii)) was not triggered by the rules in (16) (=rule (i)), but instead by the existence of *ś*. However, Martinet (1981: 55) apparently does allow for independent phonemes to trigger sound changes.

the absence of data, this position is difficult to falsify; however, I contend that (17) must have been exceptionless because all š segments had to be expunged from the language at stage 2 in order to eliminate of the contrasts between š and ś. It is also important to compare (17) with the spontaneous retroflexion of dentals, which was also context-free. Burrow (1971: 559) notes that the latter change affected primarily Old Indo-Aryan (as opposed to Middle Indo-Aryan). Hence, spontaneous retroflexion was also a very ancient sound change - although admittedly not as ancient as (17) - and yet, it was riddled with exceptions.

Other questions pertaining to the transition form stage 2 to stage 3 can be raised at this point. First, if stage 2 in (18) is impossible, then the violation to (3) could presumably be repaired in some other way. In other words, the palatoalveolar produced by (16)(a) and (b) did not necessarily have to become retroflex; it could have merged with one of the other sounds in the language, e.g. *s* or *ś*, or it could have become an entirely new sound. While I cannot say whether or not (17) is the unmarked repair strategy languages employ in order to eliminate violations of (3) that arise diachronically, it is worth noting that palatoalveolars and retroflexes are very similar phonologically, i.e. both are [+coronal, -anterior] according to many theories of distinctive features and only differ in terms of [distributed] (recall (5)). Thus, (17) involves the change of a single feature. In contrast, were š to become some other sound, such as *s*, then more than one feature would have to change. This fact therefore might have tipped the scales in favor of converting š into ṣ (as opposed to some other sound).[23]

While my analysis presupposes that the existence of ś was instrumental in the emergence of ṣ, one need not necessarily assume as I have in (8) that ś was the first postalveolar sibilant to enter Indo-Iranian. There are in fact three logical chronologies: (i) ś was in the language before ṣ (and hence š), (ii) ṣ (and hence š) was in the language before ś, and (iii) š and ś entered the language at the same time. In the preceding paragraphs I assumed (i), but, as I show below, (ii) is compatible with my analysis as well. (Since my treatment follows from the two chronologies in (i) and (ii), (iii) is equally possible). Let us now consider (ii).

Assuming that Indo-Iranian had /s, š/ and that the future ś was still the affricate č, then the emergence of ṣ has a similar explanation. When č became ś via (12), this

[23] Should the featural explanation be correct then this would imply that a more likely repair strategy at stage 2 in (18) would be the merger of š with ś. The reason is that š and ś are phonologically identical in terms of features (recall (5)). At this point I am unaware of languages in which š merges with ś diachronically.

change then caused *š* to become retroflex by (17). This sequence of events is summarized in (19):

(19) stage 1: /s, š/ → stage 2: /s, š, ś/ → stage 3: /s, ṣ, ś/

Carlton (1990: 96-97) discusses similar facts from Slavic prehistory that suggest that in that language the emergence of *x* (< *š*) (via (16)(a)) preceded the change from IE *$\hat{k}$ to *s*. As mentioned above, IE **s* surfaces in Slavic as the voiceless velar fricative *x* after *r, u, k, i* and before a vowel. Importantly, only IE **s* became *x*, and not the *s* that emerged from *ś* (< IE *$\hat{k}$). This therefore implies that the Slavic equivalent of (16)(a) was older than the spirantization of IE *$\hat{k}$.

In Avestan the facts superficially suggest a similar chronology. For example, the *s* in the Avestan word for "settlement" *vis* (Sanskrit *víś*) derives from IE *$\hat{k}$, but this *s* did not become *š* via (16)(a). However, there is an alternative explanation that presupposes the opposite ordering: Suppose that IE *$\hat{k}$ first became *ś* in all satem languages (or alternatively, only in Indo-Iranian) and this development preceded (16)(a). This *ś* then converted into s in Avestan after (16)(a) had become inactive in the grammar. In fact, there is an additional argument for the alternative sequence of events just described. Stage 2 in (18) (and in (19)) probably occurred at a point later on in the Indo-Iranian era when OIA and Old Iranian were dialects of the same language. OIA dealt with the violation to (3) at this stage by implementing (17), but the illicit inventory also existed in Old Iranian. How was the violation to (6) reconciled in that language? Recall from (11) that Indo-European voiceless alveolopalatal stops not followed by *t* surface regularly as *ś* in Sanskrit, but as *s* in Avestan. One might assume that the development of IE *$\hat{k}$ into *s* in the latter language is an arbitrary fact of Avestan, but if there was an intermediate stage to *ś* (< IE *$\hat{k}$) then a systematic explanation emerges. If IE *$\hat{k}$ became first *ś* in Indo-Iranian, and then this sound shifted to *s* in Avestan, the latter change can be seen as the same kind of context-free change in (17) that brought Sanskrit into conformance with (6). Significantly, the intermediate stage with *ś* derives phonetic support. If sound change is gradual, then the change from a voiceless alveolopalatal stop to a voiceless alveolopalatal fricative (with an intermediate affrication stage) only involves the change of a manner feature, i.e. stop → affricate → fricative.

5. Conclusion

In the preceding paragraphs I have offered an explanation for the context-free sound change in the development of Indo-Aryan that converted all palatoalveolar sibilants into retroflex sibilants. My claim is that this change was triggered by a sibilant opposition that is otherwise unattested cross-linguistically and that this is the only explanation that accounts for why the change was both context-free and exceptionless.

An obvious question to ask at this point is whether or not retroflex consonants in other languages have a similar historical development. One possible example is Polish. According to Stieber (1973: 55) fifteenth century Polish contrasted /š, ž/ and /ś, ź/, which would be problematic for (3) and (6) as absolute universals if both sets of sounds were [-anterior, +distributed]. Significantly, Stieber (1973: 64) notes that /š, ž/ became "dispalatalized" but that this did not take place before the fifteenth century. If "dispalatalization" involved the change from true palatoalveolar to retroflex (i.e. (17)) and if this change did not occur until *after* /ś, ź/ entered the Polish language in the fifteenth century then the latter change might have been the cause for the former one.

Whether or not the Polish retroflex sibilants, or the retroflex sibilants in other languages, have a historical development that is parallel to the equivalent Sanskrit sounds are questions I leave open for further study.

References

Allen, W. S. (1953). Phonetics in Ancient India. London: Oxford University Press.

Allen, W. S. (1978). The PIE velar series: Neogrammarian and other solutions in the light of attested parallels. Transactions of the philological society 1978: 78-110.

Andersen, H. (1968). IE *s after i, u, r, k in Baltic and Slavic. Acta Linguistica Hafniensia 11.2: 171-190.

Ao, B. X. (1993). Phonetics and phonology of Nantong Chinese. Ph.D. dissertation. Ohio State University.

Bartholomae, C. (1883). Handbuch der altiranischen Dialekte. Leipzig: Breitkopf & Härtel.

Bloch, J. (1930). Some problems of Indo-Aryan philology. Bulletin of the School of Oriental and African Languages 5.4: 719-756.

Bloomfield, L. (1911). Indo-European palatals in Sanskrit. American Journal of Philology 32: 36-57.

Boyle, J. A. (1966). Grammar of modern Persian. Wiesbaden: Otto Harrassowitz.

Bright, J. O. (1964). The phonology of Smith River Athapaskan (Tolowa). International Journal of American Linguistics 30: 101-107.

Brugmann, K. (1897a). Vergleichende Laut-, Stammbildungs- und Flexionslehre der indogermanischen Sprachen. Volume 1. Second edition. Strassburg: K. J. Trübner.

Brugmann, K. (1897b). Vergleichende Laut-, Stammbildungs- und Flexionslehre der indogermanischen Sprachen. Volume 2. Second edition. Strassburg: K. J. Trübner.

Burrow, T. (1945). Some Dravidian words in Sanskrit. Transactions of the Philological Society: 79-120.

Burrow, T. (1946). Loan words in Sanskrit. Transactions of the Philological Society: 1-30.

Burrow, T. (1955). The Sanskrit language. London: Faber & Faber.

Burrow, T. (1971). Spontaneous cerebrals in Sanskrit. Bulletin of the School of Oriental and African Languages 34: 538-559.

Campbell, G. L. (1991). Compendium of the world's languages. London: Routledge.

Carlton, T. R. (1990). Introduction to the phonological history of the Slavic languages. Columbus: Slavica.

Catford, J. C. (1988). A practical introduction to phonetics. Oxford: Clarendon Press.

Chatterjee, S. K. (1927). The origin and development of the Bengali language. Calcutta: Calcutta University Press.

Chatterjee, S. K. (1960). The pronunciation of Sanskrit. Indian Linguistics 21: 61-82.

Chen, M. Y. & Wang, W. S.-Y. (1975). Sound change: activation and implementation. Language 51: 255-281.

Cho, Y.-M. Y. (1990). The parameters of consonantal assimilation. Ph.D. dissertation: Stanford University.

Chomsky, N. & M. Halle (1968). The sound pattern of English. New York: Harper & Row.

Christdas, P. (1988). The phonology and morphology of Tamil. Ph.D. dissertation. Cornell University.

Dixon, R. M. W. (1980). The Languages of Australia. Cambridge: Cambridge University Press.

Dogil, G. (1990). Hissing and hushing fricatives: A comment on non-anterior spirants in Polish. Unpublished ms. Universität Bielefeld.

Edgerton, F. (1946). Sanskrit historical phonology. New Haven: American Oriental Society.

Emeneau, M. B. (1954). Linguistic prehistory of India. Proceedings of the American Philosophical society 98: 282-292.

Emeneau, M. B. (1984). Toda grammar and texts. Philadelphia: American Philosophical Society.

Feoktistov, A. P. (1984). Phonologische Analyse des Mokša-Mordwinischen Konsonantismus. In: P. Hajdú & L. Honti 1984: 201-207.

Geiger, W. (1894). Etymologie und Lautlehre des Afghanischen.

Ghatage, A. M. (1962). Historical linguistics and Indo-Aryan languages. Bombay: Bombay University Press.

Gray, L. H. (1902). Indo-Iranian phonology. New York: Columbia University Press.

Hajdú, P. & L. Honti (eds.) (1984). Studien zur phonologischen Beschreibung uralischer Sprachen. Budapest: Bibliotheca Uralica.

Hall, T. A. (1997). The phonology of coronals. to appear: Amsterdam: Benjamins.

Heidolph, K. E. et al (eds). (1975). Linguistische Studien. Symposion. Phonologische Analyse der uralischen Sprachen. Berlin, 17.-20. September 1974. Berlin: Akademie Verlag.

Hock, H. H. (1986). Principles of historical linguistics. Berlin: Mouton.

Hualde, J. I. (1991). Basque phonology. London: Routledge.

Hume, E. (1992). Front vowels, coronal consonants and their interaction in nonlinear phonology. Ph.D. dissertation: Cornell University.

Jensen, H. (1931). Neupersische Grammatik. Heidelberg: Carl Winter.

Jones, D. & D. Ward (1969). The phonetics of Russian. Cambridge: Cambridge University Press.

Keating, P. A. (1988). A survey of phonological features. Bloomington: Indiana University Linguistic Club.

Keating, P. (1991). Coronal places of articulation. In: Paradis & Prunet 1991: 29-48.

Kiparsky, P. (1988). Phonological change. In: F. J. Newmeyer (ed.) Linguistics: The Cambridge survey. volume 1. linguistic theory: foundations: 363-415. Cambridge: Cambridge University Press.

Labov, W. (1981). Resolving the neogrammarian controversy. Language 57: 267-308.

Ladefoged, P. (1964). A phonetic study of West African languages. Cambridge: Cambridge University Press.

Ladefoged, P. (1990). The revised international phonetic alphabet. Language 66: 550-552.

Ladefoged, P. & I. Maddieson (1996). The sounds of the worlds languages. Oxford: Blackwell.

Ladefoged, P. & Z. Wu (1984). Places of articulation of Pekingese fricatives and affricates. Journal of Phonetics 12: 267-278.

Lahiri, A. & S. Blumstein (1984). A re-evaluation of the feature coronal. Journal of Phonetics 12.2: 133-146.

Lambton, A. K. S. (1961). Persian grammar. Cambridge: Cambridge University Press.

Lindblad, P. (1983). Svenskans sje- och tje-ljud i ett allmänfonetiskt perspektiv. [Some Swedish sibilants]. Travaux de l'Institut de Linguistique de Lund XVI.

Mackenzie, D. N. (1961). Kurdish dialect studies I. Oxford: Oxford University Press.

Maddieson, I. (1984). Patterns of sounds. Cambridge: Cambridge University Press.

Majidi, M.-R. (1986). Strukturelle Grammatik des Neupersischen (Farsi). Volume I. Phonologie. Hamburg: Helmut Buske Verlag.

Martinet, A. (1951). Concerning some Slavic and Aryan reflexes of IE s. Word 7.2: 91-95.

Martinet, A. (1981). Sprachökonomie und Lautwandel. Eine Abhandlung über die diachronische Phonologie. Stuttgart: Klett-Cotta. (Translation of: Martinet, A. (1955): Economie des changements phonétiques. Bern: Francke).

Masica, C. P. (1991). The Indo-Aryan languages. Cambridge: Cambridge University Press.

Mayrhofer, M. (1989). Vorgeschichte. In: Schmitt 1989: 4-24.

Misra, B. G. (1967). Historical phonology of modern standard Hindi: Proto-Indo European to the present. Ph.D. dissertation. Cornell University.

Mohanan, K. P. & T. Mohanan (1984). Lexical phonology of the consonant system in Malayalam. Linguistic Inquiry 15: 575-602.

Morgenstierne, G. (1927). An etymological vocabulary of Pashto. Oslo: I Kommisjon Hos Jakob Dybwsd.

Morgenstierne, G. (1945). Indo-European k in Kafiri. Norsk Tidsskrift for Sprogvidenskap 13: 225-238.

Munro, P. & J. Benson (1973). Reduplication and rule ordering in Luiseño. International Journal of American Linguistics 39: 15: 21.

Paradis, C. & J.-F. Prunet (eds.) (1991). Phonetics and phonology. The special status of coronals. Internal and external evidence. New York: Academic Press.

Penzl, H. (1955). A grammar of Pashto. Washington: American Council of Learned Societies.

Pullum, G. K. & W. A. Ladusaw (1986). Phonetic symbol guide. Chicago: University of Chicago Press.

Recasens, D. (1990). The articulatory characteristics of palatal consonants. Journal of Phonetics 18: 267-280.

Rédei, K. (1975). Phonologische Beschreibung des Syrjänischen. In: K. E. Heidolph et al 1975: 103-116.

Rédei, K. (1984). Phonologische Analyse des Erza-Mordwinischen. In: P. Hajdú & L. Honti 1984: 209-230.

Schmitt, R. (ed.) (1989). Compendium Linguarum Iranicarum. Wiesbaden: Dr. Ludwig Reichert Verlag.

Shalev, M., P. Ladefoged, & P. Bhaskararao (1993). Phonetics of Toda. UCLA Working Papers in Phonetics 64: 89-125.

Skjærvø, P. O. (1989a). Modern East Iranian languages. In: Schmitt 1989: 370-383.

Skjærvø, P. O. (1989b). Pashto. In: Schmitt 1989: 384-410.

Smirnova, I. A. & K. P. Ejubi (1985). Fonetika kurdskogo jazika. Leningrad: Nauka.

Solta, G. R. (1965). Palatalisierung und Labialisierung. Indogermanische Forschungen 70: 276-315.

Spencer, A. (1986). A non-linear analysis of vowel-zero alternations in Polish. Journal of Linguistics 22: 249-280.

Stevens, K. (1989). On the quantal nature of speech. Journal of Phonetics 17: 3-45.

Stieber, Z. (1973). A historical phonology of the Polish language. Heidelberg: Carl Winter.

Thumb, A. (1958). Handbuch des Sanskrit. Heidelberg: Carl Winter Universitätsverlag.

Tischler, J. (1990). Hundert Jahre kentum-satem Theorie. Indogermanische Forschungen 95: 63-98.

Veenker, W. (1987). Vorschlag für ein phonologisches Beschreibungsmodell der uralischen Sprachen. In: K. Rédei 1987: 7-46.

Vennemann, T. (1974). Sanskrit ruki and the concept of a natural class. Linguistics 130: 91-97.

Viitso, T.-R. (1975). Outlines of Livonian phonology. In: K. E. Heidolph et al 1975: 85-103.

Wackernagel, J. (1896) Altindische Grammatik. Band I. Lautlehre. [1957, Göttingen: Vandenhoek & Ruprecht].

Whitney, W. D. (1889). Sanskrit grammar. Cambridge: Harvard University Press.

Wurzel, W. U. (1975). Die phonologischen Merkmale der uralischen Sprachen. In: K. E. Heidolph et al 1975: 167-183.

Zvelebil, K. (1970). Comparative Dravidian phonology. The Hague: Mouton.

Zvelebil, K. (1990). Dravidian linguistics. An introduction. Pondicherry: Pondicherry Institute of Linguistics and Culture.

Zwicky, A (1970). Greek-letter variables and the Sanskrit ruki class. Linguistic Inquiry 1: 549-554.

OPTIMALITY THEORY AND PHONOLOGICAL CHANGE

Haike Jacobs

French Department
Nijmegen University/Free University Amsterdam
CLS/HIL

1. Introduction

In generative phonology it has been, since Kiparsky (1968), a standard practice to account for sound change by means of rule addition, rule simplification, rule reordering and rule loss. Given that the phonologiocal rule as such no longer exists in recently proposed constraint-based theories of phonology, such as Optimality Theory (cf. Prince and Smolensky (1993), McCarthy and Prince (1993a and b), the question arises how sound change can be accounted for in these theories.

In this paper[1] we will address this issue. It will be claimed, as is to be expected, that in Optimality Theory, sound change can straightforwardly be accounted for by constraint reranking[2]. This will be illustrated by an example of sound change in the historical phonology of French. It involves the loss of the possibility to phonologically encliticize an unstressed object pronoun to a preceding stressed element (phonological enclisis). The formal account of this change relies on a reordering of Alignment constraints.

This paper purports to show not only that sound change can be analyzed as constraint-reranking (as already mentioned an expected result), but also that the analyses proposed are not thwarted by the same drawbacks of previous derivational nonlinear accounts. Finally, we will raise the issue of how to account for markedness and unmarkedness in sound change.

This study is organized as follows. In section 2, we will discuss the first example, the loss of phonological enclisis. We will point out some of the problematic aspects of previous phonological and syntactic accounts. Section 3 presents an OT-analysis making use of Alignment constraints. It will be argued that this sound change can be more adequately accounted for by a constraint-reranking. More importantly, the proposed analysis will be shown to have some attractive consequences for syntactic accounts of similar phenomena. Finally, in section 4, we will summarize and discuss the main results of the proposed analyses.

---

1 **This is a slightly revised and expanded version of a paper presented at the ABRALIN1 conference which is going to appear in the proceedings of that conference. During my stay as a Gastwissenschaftler at the Zentrum für Allgemeine Sprachwissenschaft in Berlin I had the opportunity to discuss the facts and analyses presented here in a very stimulating academic environment. I would like to thank the ZAS for inviting me, and, in particular, T.A. Hall, Ursula Kleinhenz, Sylvia Löhken, and Gustav Wurzel for their warmhearted hospitality.**

2 **Löhken (1996) presents a very detailed account of sound change and Optimality Theory in the historical phonology of German.**

2. The loss of phonological enclisis in French

In Old French, monosyllabic unstressed function words could be pronounced either as part of the word that preceded them (enclisis[3]; for instance, *jol vi* 'I saw him') or as part of the word that followed them (proclisis; for instance, *jo l'aim* 'I love him/her'). In the evolution from Gallo-Romance to Old French, the possibility of encliticizing disappeared[4]. Traditional scholars such as Kukenheim (1971) have argued that the loss of enclisis was caused by a change in the rhythmic structure of the language. Classical Latin had initial stress and a descending rhythm, which was replaced by an ascending rhythm and final stress in the evolution from Gallo-Romance to Old French. The idea of a relation between phonological enclisis and strong initial stress has been advanced more recently by Adams (1987) in her study of null-subjects and Verb-second effects in Old French. Adams (1987) not only attributes the loss of enclisis to the above-mentioned change, but also considers this change to have been the cause for the cliticization of subject pronouns and the loss of Verb-second phenomena.

Neither Kukenheim nor Adams formalized their insights. Adams states that "the loss of enclisis was part of a process by which all elements in the phrase gave up their individual accent to that of the final tonic syllable; it thus points to a fundamental change in accentuation. As long as *je*, for example, in *jes avrai* remained an independent form with its own accent, *les* could cliticize to it." Adams (1987:165) presents the two grammars, repeated below as (1a) and (1b), for Gallo-Romance and Old French in order to clarify the difference in constituent structure.

(1a)

Jé	les	avrái
Jé+les		avrái
Jés		avrái

(1b)

Je	les	avrái
Je	les +	avrái
Je +	les +	avrái

3 **Enclisis was optional in Old French. In Modern French only a few lexicalized remnants can be observed, such as, the contraction of *de + le > du, de + les > des, à + le > au* and *à + les > aux*. These forms can be analyzed along the lines of Zwicky (1987) as syntactic allomorphy or along the lines of Hayes (1990) as precompiled phrasal allomorphy. The optional character of Old French enclisis seems to exclude such an analysis. The reader is referred to Jacobs (1993) for a more detailed account and to Evers (1994) for a partially lexicalized approach.**

4 **Clitics are considered following Garde (1968:70-72) as basically stressless and therefore have to be integrated into prosodic words. A distinction has to be made between phonological and syntactical proclisis and enclisis. An unstressed object pronoun in preverbal position can syntactically be proclitic on the verb, but phonologically be enclitic on a preceding stressed element. This is most clearly shown by cases such as Old French *jot vi* 'I saw you', where the phonological enclisis is clear by the reduction and integration of the object pronoun into the preceding stressed subject pronoun, but where syntactically the object pronoun is proclitic on the verb.**

Recent research on prosodic phonology has led to the development of a number of prosodic theories which all postulate a level of representation that is not necessarily isomorphic with syntactic structure and that mediates between the phonological and syntactic components of a grammar. These theories make it possible to give a more formal expression of the insights provided by scholars such as Kukenheim and Adams. Moreover, as will be shown, they not only do that, but also can add to our understanding of why the changes took place in the way they did. This section is organized as follows: in section 2.1, we will concentrate mainly on the prosodic conditions, that is the domain of application of enclisis and proclisis, and, on the evolutionary change from enclisis to proclisis. We will first discuss the prosodic theories of Selkirk and Shen (1990) and Nespor and Vogel (1986) which both allow for a more formal account of enclisis and proclisis as well as for the above-mentioned evolutionary change. Next, we will examine in section 2.2 the predictions made by and the problematical aspects of both theories by looking at Brazilian and European Portuguese. After that, section 2.3 briefly discusses how syntactic theories try to explain partially the same phenomena as prosodic theories.

2.1 Prosodic theories

Selkirk and Shen's (1990) Edge-based theory of the syntax-phonology mapping allows for the construction of two prosodic constituents: the Prosodic Word (PW) and the Major Phrase. The mapping of syntactic structure to prosodic structure is defined by the algorithm in (2).

(2) Syntax-Phonology Mapping (Selkirk and Shen, 1990:319)

For each category Cn of the prosodic structure of a language
there is a two-part parameter of the form

Cn: {RIGHT/LEFT;Xm}
Where Xm is a category type in X-bar theory

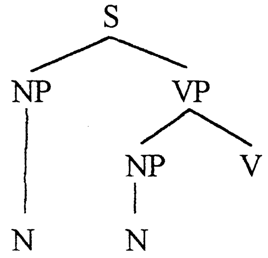
For each language it must be specified whether the right- or the left-edge of syntactic categories in the syntax-phonology mapping is used. For the Gallo-Romance syntax-phonology mapping rule constructing Prosodic Words, the choice of the left-edge of syntactic categories yields domains in which the preverbal clitic object pronoun is separated from the verb although syntactically being dependent on it. The object pronoun is thus phonologically enclitic, but syntactically proclitic.
Furthermore, if for Old French the parameter is reset to right-edges, one obtains a domain in which the preverbal clitic object pronoun is no longer separated from the verb, but together with a preceding non-lexical item incorporated within the same domain as the verb. The Gallo-Romance and Old French Prosodic Word rules can be stated as in (3).

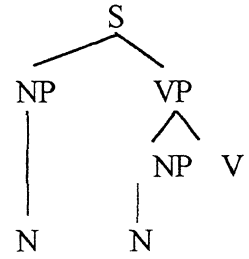
(3) a) Gallo-Romance Prosodic Word rule
Prosodic Word: {Left, Lex0}

b) Old French Prosodic Word rule
Prosodic Word: {Right,Lex0}

The different prosodic constituent structures made possible by the different parameter settings in (3) are listed in (4), and, are entirely consistent with the different constituent structures assumed by Adams (1987) in (1) above.

(4a) Parameter 2 is set to LEFT (4b) Parameter 2 is set to RIGHT






```
PW-Left                  [          PW-Right              ]
   (          )     ( )                (        )

    jo        le       vi                   jo       le    vi
```

In (4a) the left-edge is used for the syntax-phonology mapping and in (4b) the right-edge. The syntactic representation in (4) is greatly simplified for ease of exposition. Both subject and object pronouns have been represented as NP's. The exact syntactic nature, however, does not crucially alter the prosodic constituent structure. Selkirk and Shen (1990) allow a variation in (2) according to which all syntactic categories or only lexical syntacic categories count for purposes of domain construction. Given that pronouns are functional and not lexical instances of syntactic categories whatever their label, the same prosodic constituents as in (4) will obtain as long as we specify that lexical X0's are choosen for the PW-rules.
It is clear that the evolution from enclisis to proclisis can straightforwardly be described as a change in the edge-parameter setting for the construction of Prosodic Words from (3a) to (3b). In (4a), the parameter setting LEFT automatically induces a PW boundary at the beginning of the utterance, hence the domains (*jo le*) and (*vi*). *Jo* and *vi* being the heads of different PW's both receive stress and, therefore, *le* is able to encliticize on the preceding stressed subject pronoun. Given that stress is final in structure (4b), the subject pronoun is no longer stressed and consequently a following clitic (unstressed) object pronoun can no longer cliticize onto it. Furthermore, the cliticization of subject pronouns themselves also follows quite naturally from the change in the edge-parameter setting.
Moreover, given that the edges that are relevant for the syntax-phonology mapping correspond to the location of word and phrasal stress (final in Old and Modern French; initial in Classical Latin), it seems logical to try and connect these facts in a more principled way. Jacobs (1993) has proposed the principle in (5).

(5) The syntax-phonology mapping parameter is set to RIGHT in a language with final and to LEFT in a language with initial stress.

The edge-based theory of the syntax-phonology mapping thus seems to provide a straightforward explanation for the loss of enlcisis. The change from a descending rhythm to an ascending rhythm (from Classical Latin phrase initial stress to Gallo-Romance and Old French phrase final stress) can formally be described as a change in the edge-parameter setting for the construction of Prosodic Words. The different prosodic constituent structure that resulted from this switch can be claimed to be responsible for the loss of enclisis (cf. Jacobs 1993, for a more detailed account).

There is, however, another way of accounting for the facts which is based on the prosodic theory of Nespor and Vogel (1986) and which makes different empirical predictions. In the prosodic theory advocated by Nespor and Vogel a constituent is proposed which mediates between the phonological word and the phonological phrase and which, at first sight, seems to be the ideal candidate for defining the domain of application of clitic phonology: the clitic group. The construction of the clitic group (C) groups together a host and its clitics according to the algorithm in (6).

(6) Clitic group formation

I C-domain

The domain of C consists of a W (Phonological Word) containing an independent (i.e. a nonclitic) word plus any adjacent W's containing

a. a DCL, or
b. a CL such that there is no possible host with which it shares more category memberships.

II C-construction

Join into an n-ary branching C all W's inluded in a string delimited by the definition of the domain of C

Nespor and Vogel distinguish between directional clitics (DCL), such as, for instance, the Greek possessives (Nespor and Vogel 1986:153) which always attach to a host in one specific direction regardless of the syntactic configuration and clitics *tout court* which attach to that host with which they share most syntactic category memberships.[5] Given that object pronouns can attach to a host on the right as well as on the left (compare *jol vi* and *jo l'aim*), they must not be considered DCL's, but CL's. However, given that the object pronoun will always share more syntactic category membership with the following verb, and, given that *jo* -being able to occur in isolation and being able to be separated from the verb by an adverb- is not a clitic itself, but an independent word, the algorithm in (6) will yield for the proclisis and enclisis examples above the same C-domain division. This is illustrated in (7).

(7)
```
C       C       C       C
|      /|       |      /|
W   W   W       W   W   W
|   |   |       |   |   |
jo  le  vi      jo  le  aim
jol     vi      jo      l'aim
```

It should also be noticed that, because *jo* is not a clitic, it is not possible either to have one single clitic group consisting of *jo* + object pronoun + verb. Besides the clitic group, which is not able to define the proper domain for the application of proclisis and enclisis, Nespor and Vogel's theory contains the phonological phrase, which, as will be shown, can serve as the correct

[5] **Syntactic category membership can be defined as follows: X and Y share category membership in Z if Z dominates both X and Y.**

characterization of the domain of application of enclisis and proclisis.
The definition of the phonological phrase is given in (8) (cf. Nespor and Vogel, 1986).

(8) Phonological Phrase (PP)

Join into a PP any lexical head (X) with all items ont its non-recursive side within the maximal projection and with any other non lexical items on the same side.

Given that the word order in the evolution of French has changed from basically OV to VO (a change that took place at about the same time as the loss of enclisis[6]), the construction of PP's must have changed also according to (8). In the OV-period, the object pronoun will be separated from the verb (hence liable to encliticize onto a preceding stressed host), whereas in the VO period it will be grouped together with the verb into one single PP (hence liable to procliticize onto the verb). Enclisis and proclisis could then be thought of as ways of licensing clitics by incorporating them into the prosodic hierarchy, and, the loss of enclisis would follow as a natural consequence of the change in the PP-domain triggered by the syntactic word order (that is, please recall fn.6, head-initial to head-final) change.
In this section, we have discussed how the prosodic theories of Nespor and Vogel (1986) and of Selkirk and Shen (1990) allow for a description of the Old French clisis processes. In the next section, we will discuss the problematical aspects and empirical predictions of these theories.

2.2 **Problems and predictions**

There are a number of problematical aspects with both analyses presented in the previous section. We will only disuss the more important ones and refer for a more detailed account to Evers (1994).
First of all, both the Edge-based account and the PP-account seem to imply that proclisis was not possible until the edge-parameter was reset from LEFT to RIGHT or until the word order changed from OV to VO. This is so because the Gallo-Romance parameter setting (3'a) ({Lex0,Left}) will always induce a word boundary between a proclitic word and a following lexical X0. Therefore, the intermediate stage of the language where both enclisis and proclisis are possible (an example of this is given in (9) cf. Evers, 1994:15) is hard to describe.

(9) Enclisis and proclisis in Old French (12th century *Chanson de Roland*)

Fors ***s'en*** *eissirent li Sarrazins dedenz*
Sis *cumbatirent* ***al*** *bon vassal Rolant*

The Sarrasins who were inside went outside
and fought with the good knight Roland

---

[6] **It is more accurate to say that syntactic structures have changed from head-initial to head-final, thus from having the recursive side on the left of the head (Latin, Gallo-Romance) to having the recursive side on the right of the head (Old French, Modern French). For a more detailed account the reader is referred to Bichakjian (1988) and Bauer (1992).**

One either has to assume in the edge-based theory account that the two different rules of prosodic word construction (3a) and (3b) coexisted for some time or that after the switch from LEFT to RIGHT ((3a) to (3b)), initial stress subsisted for some time (cf. Jacobs, 1993). Given that this period must have lasted four centuries, since enclisis and proclisis coexisted from the earliest Old French documents on, neither assumption is very attractive. Mutatis mutandis, the same holds for the PP-construction in the OV-period. In conclusion, both prosodic theories to some extent provide a possible account for the evolutionary change from enclisis to proclisis, but face similar problems in defining the correct domains for the application of proclisis and enclisis in the Old French period.

Second, the two theories make *different empirical predictions. Whereas the analysis casted in the* edge-based theory predicts that enclisis can occur in languages that have initial stress, but cannot occur in languages that have final stress, the analysis presented in the Nespor and Vogel theory predicts enclisis to be possible in OV-languages, but not in VO-languages.

Let us briefly examine these predictions. One of the differences between European Portuguese (EP) and Brazilian Portuguese has been claimed to be the preference for enclisis in EP and for proclisis in BP. Brandão de Carvalho (1989) argues that EP object pronouns are always enclitic[7], whereas BP object pronouns are always proclitic. Some examples are given in (10).

(10a)EP	Eu vi-te ontem	'I saw you yesterday'
	Ele disse-te que..	'He told you that...'
	O gato apanhou-o	'The cat caught him'

(10b)BP	Eu te vi ontem
	Ele te disse que...
	O gato o apanhou

The cases in (10) are cases where pronominal placement is theoretically free, that is in independent or root clauses with an overt subject and no initial adverbial complement. According to Brandão de Carvalho (1989:407) the utterances in (10a) are "perfectly possible in Brazil; however, they are perceived as European and/or quite normative in BP speech." The same holds for the possibility of (10b) in EP, which "will generally be felt as 'Brazilianlike' or somewhat 'literary' in normal styles." Given that the basic word order in both varieties is identical, the predictions of the Nespor and Vogel account are not borne out by the facts. Moreover, the fact that both (10a) and (10b) are possible in both varieties makes a **strong** link between either initial stress or OV-word order and the existence of enclisis in a language hard to sustain. This does not mean of course that stress does not play a role in clisis processes. The function of stress can quite clearly be observed in the EP and BP contrastive examples in (11).

(11a)EP	Diga-me	(11b)	BP	Me diga
	De-me			Me de

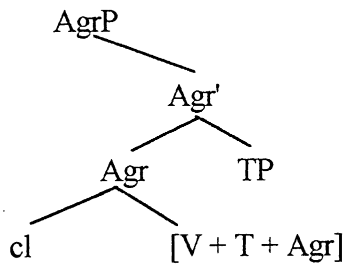
[7] Brandão de Carvalho (1989:409) also considers EP cases where the pronoun figures obligatory in preverbal position (subordinates, interrogative sentences, and utterances with an initial adverbial complement) as being phonologically enclitic. Thus, for instance, in cases such as '*Não te vi*' 'I didn't see you', or '*Já te dou*' 'I'll give (it) to you now', the object pronoun although in preverbal position is considered to be enclitic on the preceding stressed element. In BP, the object pronoun is considered to be syntactically as well as phonologically proclitic to the verb.

Brandão de Carvalho states that no EP utterance can begin with an unstressed pronoun[8], hence the enclitic forms in (11a). Given the absence of such a constraint in BP, the proclitic forms in (11b) which do have an unstressed pronoun in utterance initial position are only possible in BP.
In this subsection, we have discussed the problematical aspects and empirical predictions of the edge-based and Nespor and Vogel-based accounts of enclisis and proclisis. It has been argued that both analyses face problems in accounting for the Old French period in which both enclisis and proclisis were possible, and, that the predictions they make are not borne out by the facts of Brazilian and European Portuguese. In the next section, we will briefly discuss how syntactic proposals try to account for partially the same phenomena as phonological accounts.

2.3 **Syntactic accounts**

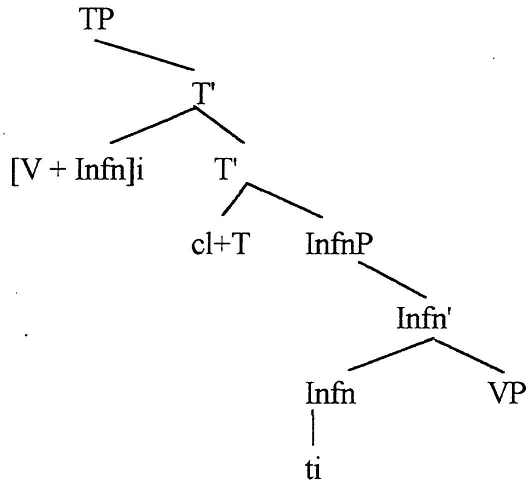
An overview of the literature on clitic placement in syntax cannot be but incomplete. We will therefore only briefly discuss a recent proposal by Madeira (1993). After discussing Kayne's (1991) account of clitic placement in Italian and Spanish, where the position of the clitic is related to the tensed/untensed nature of clauses (compare Italian *la guardano* 'they look at her' versus *guardar-la* 'to look at her'), Madeira examines EP where the position of the clitic in tensed clauses is not always preverbal.
Clitics are assumed to be base-generated as the head of a DP subcategorized for by the verb. The surface position is reached by movement into a higher functional head. For Italian and Spanish tensed clauses, the clitic is assumed to left-adjoin to the functional head where the verbal complex is found: AGR. In infinitival clauses, the clitic is moved to an "abstract T-node and movement of the verbal complex is past it to a position adjoined to T"'. Schematically this can be represented as in (12) (taken from Madeira (1993:157).

(12) a) Italian and Spanish tensed clauses



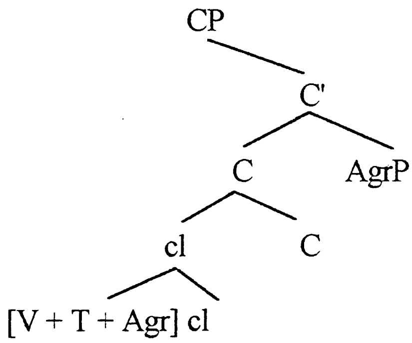
[8] **This constraint is traditionally known as the Tobler-Mussafia Law.**

b) Italian and Spanish infinitival clauses



Languages with a verb-clitic root order thus pose a problem for the account along the lines of (12). Madeira suggests that the clitic in these cases is moved into an empty C-node in root clauses. The verb also moves into C and is left-adjoined to the clitic, resulting in the order verb-clitic. This is illustrated in (13).

(13) Enclisis in root clauses



Movement of the verb into C is motivated by the fact that the clitic must syntactically be incorporated by the verb presumably in order to check case features. We will not discuss this matter in detail here. The reader is referred to Madeira (1993:161-162; especially fn.6). Madeira then generalizes her account to Italian, Spanish and European Portuguese by assuming that clitics move to the highest functional head with the restriction that movement to empty C is only possible in EP, but not in Italian and Spanish.
Movement of the clitic to the empty C in root clauses is taken to be the explanation for the fact that in EP clitics cannot occur in sentence-initial position. This is illustrated in (14).

(14) Telefonou-lhe o Paulo

*lhe telefonou o Paulo

Madeira (1993:173) explains this as follows. If the clitic is moved to the highest available functional head (the empty C) then "in order to satisfy the incorporation requirement, the verbal complex must move up, left-adjoining to the clitic in C." If the verbal complex does not attach to to the clitic, the representation is ruled out. However, if we recall the contrastive pairs in (11a) and (11b), it becomes clear that the ban on clitic-first is to some extent independent of syntactic incorporation requirements. Rather, it seems to be the case that the verbal complex adjoins to the LEFT in EP, but to the RIGHT in BP. Moreover, in Madeira's analysis of EP, the clitic surfaces to the left in clitic-verb sequences (where the verb is taken to be the head (12a)), but to the right in verb-clitic sequences (where the clitic is taken to be the head (13)). No syntactic motivation is given to support this asymmetry (cf. Madeira, 1993:162, fn.7).
In this section, we have briefly discussed a recent syntactic proposal for clitic placement. We have tried to demonstrate that contrasts such as the ones in (11) can observationally be described in terms of different directions of adjunction, but that no motivation for such a state of affairs can be provided. This of course does not mean that syntactic considerations do not play a role in clitic placement, but only that not all facts concerning clitic placement can adequately be handled syntactically[9].
In the next section we will present the outlines of an optimality-theoretic approach to proclisis and enclisis which is not thwarted by the same problems as the prosodic theories discussed above, and, which can also account for the aspects that were shown to be syntactically problematical.

3. An optimality approach

In Optimality theory (Prince and Smolensky 1993) phonology is thought of as a universal set of constraints which are hierarchically ranked on a language-specific basis. The relation between input and output is accounted for by two functions, GEN and H-EVAL, which respectively generate for each input all possible outputs and evaluate which output is optimal (cf. Prince and Smolensky 1993 for a more detailed account). Thus in Optimality theory the phonological rule as such no longer exists. Rather, starting from an input **all** possible outputs are generated and evaluated against the constraint-ranking of the language until the optimal output is found. The candidate which best satisfies the constraint hierarchy is evaluated as the optimal one. The role of phonological rules has thus been entirely subsumed by the constraint hierarchy (for more details see Prince and Smolensky 1993). In Optimality theory, constraints may be violated, depending on the ranking of other constraints. This then is a crucial difference between the way constraints have hitherto been conceived of and Optimality theory. The following example, taken from Prince and Smolensky (1993:29) should make this clear. Speaking in derivationalist terms, languages normally do not allow heavy syllables to be split by foot-construction rules. A principle of Syllabic Integrity, stating that foot-parsing may not dissect syllables, (cf. Prince (1976)) is

[9] Another phenomenon which is quite difficult to handle syntactically is the so-called mesoclisis in EP, as in, for instance, *visitar-te-emos* 'we will visit you' (cf. Madeira (1993:157) and van der Leeuw (1994) for an elegant account in Optimality theory).

assumed to guarantee this. It is therefore, that in Classical Latin, for instance, moraic trochee construction skips a light penultimate syllable, if the antepenultimate syllable is heavy. Now, in Tongan main stress falls on the penultimate mora of a word. However, unlike in Classical Latin, for instance, in a sequence /-CVVCV/, the VV sequence is split in two, yielding CV.(V.CV). Compare *húu* 'go in' (monosyllabic) versus *hu.ú.fi* 'open officially' (trisyllabic). In a rule-based approach, a rule of foot construction then necessarily has to violate a constraint (Syllabic Integrity) assumed to be universal.
Optimality theory offers a solution in terms of constraint domination. Two constraints are invoked by Prince and Smolensky (1993:28-29). One is EDGEMOST which states that the most prominent foot in the word is at the right edge, and the other one is ONS which states that every syllable has an onset. If the constraint EDGEMOST dominates ONS, the facts of Tongan will obtain. In (15) this is illustrated in a so-called constraint tableau. The F points to the optimal candidate, the * means a violation of a constraint, and the ! points to crucial constraint satisfaction failure.

(15) Candidates	Edgemost	ONS
F hu.(ú.fi)		*
(húu).fi	s!	

In (15) only candidates that are properly bracketed are considered. Other ill-parsed possible candidates will be ruled out by other constraints (cf. Prince and Smolensky (1993) for a more detailed account). If constraints in a rule-based theory can be conceived of as a sort of 'phonological customs inspection' (Kenstowicz (1994:531), where a violation is fatal, the constraints in Optimality theory are less rigid, where candidates are allowed to violate constraints, as long as they better satisfy higher-ranked constraints than other candidates.
McCarthy and Prince (1993 a and b) propose a unified theory (called Generalized Alignment) to account for the different ways in which constituent-edges are referred to in phonology and morphology. Basically a Generalized Alignment requirement means that an edge (R/L) of a prosodic or morphological constituent must coincide with an edge (R/L) of another prosodic or morphological constituent according to the general schema in (16).

(16) General schema for ALIGN

In ALIGN (GCat, GEdge, PCat, PEdge), the GEdge of any GCat must coincide with PEdge of some PCat, where

GCat = Grammatical category, among which are the morphological categories
MCat = Root, Stem, Morphological Word, Prefix, Suffix, etc.
PCat = Prosodic Category = μ, s, Ft, PrWd, PhPhrase, etc.
GEdge, PEdge = Left, Right

The general schema for Alignment can be understood, according to McCarthy and Prince (1993b:32) as "extending to word-internal constituency the edge-based theory of the syntax/phonology interface." The Alignment schema in (16) can thus be understood as defining part of the Morphology-Phonology interface.

Conversely, the syntax-phonology mapping parameter from (2) above can be defined in terms of Generalized Alignment, which allows for a definition of the Syntax-Phonology interface in a similar way as the Morphology-Phonology interface.
For Gallo-Romance, we reformulate the Prosodic Word rule (3a) as the alignment instruction (17).

(17) Align-Lex0-LEFT: Align (Lex0, L, PrWd, L)

According to the constraint in (17) any left-edge of a Lex0 should coincide with the left-edge of a Prosodic Word. The problems that both the prosodic theories discussed in section 2 were confronted with (accounting for both enclisis and proclisis at the same time) can now easily be solved. As a constraint (17) can be violated depending on the ranking of other constraints. The cases where it is violated in Old French are precisely the cases where we have proclisis.
Now in order to enforce proclisis, we need a constraint ranking that in the case of a vowel hiatus, will ensure the non-surfacing of the first vowel rather than the insertion of an epenthetic consonant. The constraints needed to guarantee this are motivated in Prince and Smolensky (1993:85-96). The constraints involved are ONS (syllables must have onsets), PARSE (underlying segments must be parsed) and FILL (syllable positions must be filled with underlying segments). In order to illustrate these constraints, let us consider an input string /V/. If ONS dominates the other two constraints, the relative ranking of PARSE and FILL yields the results in (18a) and (18b), where a dot represents a syllable boundary, an angled bracket unparsed material and ¨ an empty node.

(18a) /V/	ONS	PARSE	FILL
. V .	*!		
<V>		*!	
F .¨V.			*

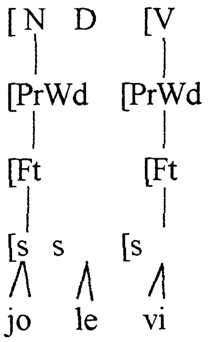
(18b) /V/	ONS	FILL	PARSE
. V .	*!		
F <V>			*
.¨V.		*!	

In (18a), where PARSE dominates FILL, the non-syllabification of input /V/ is less optimal than adding extra material (the epenthetic consonant represented by ¨). In other words, PARSE demands fully syllabified candidates, regardless of whether they contain extra material, whereas, FILL demands non-epenthetic forms, even if they contain unparsed material.
For Gallo-Romance and Old French proclisis, we need underparsing (deletion) of V's, hence the constraint order ONS>>FILL>>PARSE. In order to be effective these constraints must dominate the constraint in (17). The constraint ranking assumed thus far is given in (19).

(19) ONS>>FILL>>PARSE>>ALIGN(Lex0).[10]

[10]It should be noticed that proclisis did not apply to any V#V sequence. It only applied to a specific set of unstressed function words (articles, object pronouns, possessives and

Let us now return to enclisis. Two issues will be addressed here. One, why is it that the object pronoun occurs in preverbal position? Is this something that can be accounted for by the constraint hierarchy or is this an aspect that should be accounted for in terms of syntactic movement rules, and, thus should be handled in another component of the grammar? Two, how does the phonological enclisis (reduction of the unstressed object pronoun and enclisis onto the preceding stressed element) translate into Optimality theory?
Let us start with the second issue. In order to facilitate the discussion, let us first consider (20).

(20) 

In (20) we have given a partial representation where according to the constraint in (17), the left-edges of a Lex0 have been left-aligned with a PrWd. Also, in (20) we have partially represented in terms of alignment, the Prosodic Hierarchy Hypothesis (cf, among others, Selkirk (1980), Nespor and Vogel (1986)). The analysis that will be given depends on the underlying representation of the articles and object pronouns that were liable to encliticize in Old French (*le*, *les*, (articles and pronouns) and, occasionally, the unstressed pronouns *me*, *te*, *se*, and *en*). If it is assumed that in underlying representation these function words do not have a vowel nucleus (cf. Evers (1994)), the possible outputs *jol vi* and *jo le vi* can be accounted for by assuming that both forms are incorporated into the preceding word (the syllable to the preceding foot in the case of *jo le* and the segment *l* to the preceding syllable in the case of *jol*) in order to satisfy the constraint FtBIN (according to which feet must be binary under syllabic or moraic analysis, cf. Prince and Smolensky (1993) and McCarthy and Prince (1993b)). The choice between either *jol* or *jo le* can be made on the basis of an optional reranking of the relevant syllable-structure constraints (cf. Prince and Smolensky, 1993:85-96). For *jol*, the constraint -COD (syllables must not have a coda) must be dominated by PARSE, which in turn must dominate FILL^{nuc} (a nucleus node must be filled with underlying segments). For *jo le*, the constraint ranking must be the following: -COD>>PARSE>>FILL^{nuc}. This is illustrated in (21). The input is /jol/.

(21a) Candidates	PARSE	FILL^{nuc}	-COD
F .jol.			*
.jo<l>.	*!		
.jo.l¨		*!	

particles (cf. Einhorn, 1974:12). We will not work this out in more detail nor provide constraint tableaus motivating the ranking assumed.

(21b) Candidates	-COD	PARSE	FILL^{nuc}
.jol.	*!		
.jo<l>.		*!	
F .jo.l¨			*

The constraint FILL^{nuc} must be interpreted as follows. Given our assumption that the object pronoun *le* in underlying representation does not have a nucleus, providing one by GEN as in the last candidate in both (21a) and (21b) is a violation of Fill^{nuc}. If this form is the actual output, as it is in (21b), then that empty nucleus is interpreted as a schwa in the output (cf. Prince and Smolensky, 1993:50-51 for related discussion).

Summarizing thus far, we have presented the outlines of an optimality analysis of Gallo-Romance and Old French proclisis and enclisis. It has been argued that the problem of earlier prosodic theories to define the domains of application for enclisis and proclisis can be solved in Optimality theory, precisely because the syntax/phonology mapping defined in terms of alignment is less rigid in the sense that the constraints defining it can be violated by higher-ranked constraints, viz. the syllable structure constraints in the case of Old French. Furthermore, the predictions made by the two theories discussed in section 2 which were shown to be problematic are no longer made by the present Optimality-theoretic analysis.

Let us now address the question as to whether enclisis and proclisis in the syntactic sense of preverbal and postverbal position can also be handled along the lines of the Alignment-based syntax/phonology interface or should be left to the syntactic component. It should be made clear at the outset that this is a question which can not be answered easily and that a lot more research in this respect is needed.

In section 2.3, we have discussed Madeira's (1993) proposal to account for the ban on clitic-first in EP. We have argued that given the contrastive pairs in (11), the ban on clitic-first is very likely to be independent of syntactic incorporation requirements. Rather, it seemed that the verbal complex adjoins to the LEFT of a clitic in C in EP, but to the RIGHT in BP. What we would like to propose in order to account for the difference between EP and BP is that the clitic and the verb are adjoined to the relevant syntactic node by syntactic movement rules, but that the order in which they surface follows as a consequence of the Alignment instruction for the syntax/phonology mapping. We then no longer need to have unmotivated left- or right adjunction of the verbal complex in C, but can make the surface position of the clitic follow as a rather straightforward result of the syntax/phonology mapping.

In order to account for the contrastive EP and BP pairs in (11), we assume for EP a high-ranked constraint that demands a left-alignment of Lex0's with PrWd's (ALIGN-Lex0-L), whereas for BP the opposite ranking is assumed: a high-ranked constraint is assumed that demands right-alignment of Lex0's with PrWd's (ALIGN-Lex0-R), as informally represented in (22). This is in general, what seems to be required to motivate a constraint: adduce examples of other languages where the same constraint or exactly the opposite constraint ranking is necessary.

(22) EP ALIGN-Lex0-L >> ALIGN-Lex0-R

BP ALIGN-Lex0-R >> ALIGN-Lexo-L

To further strengthen this idea, let us first try to work out some more EP examples. Following Brandão de Carvalho (1989), we assume the forms in (23) to represent the unmarked state of affairs with respect to pronoun-placement in EP. Brandão de Carvalho (1989:407) distinguishes between cases where pronoun placement is theoretically free (23a-d), but where EP "shows a greater propensity to 'enclisis', that is to postverbal position" and cases (23e-g) where pronoun placement is not free (please recall fn.7).

(23) unmarked a) Eu vi-te ontem "I saw you yesterday"
marked b) Eu-te vi ontem id.

unmarked c) O gato apanhou-o "The cat caught him"
marked d) O gato o apanhou id.

obligatory e) Diga-me "Tell me"
position f) Não te vi "I did not see you"
g) Quem me vê? "Who sees me?"

As far as syntactic movement is concerned, we follow the analysis of Kayne (1991) from (12a) and (12b) above, in which it is assumed that clitics are adjoined to the functional head where the verbal complex is found. Now, if we want to get the position of the clitic with respect to the verbal complex for free as a result of alignment, then, besides the constraints demanding left-alignment and right-alignment of Lex0 with PrWd's in EP, one more alignment constraint is needed that demands that a PrWd must begin with a foot: ALIGN-PrWd (PrWd, L, Ft, L). This constraint has been extensively motivated in OT-literature (cf. among others Prince and McCarthy (1993)).
It should be recalled from the discussion of Madeira's proposal that the order verb-clitic only occurs when the C-position is empty. When the C-position is filled the order is clitic-verb (interrogative sentences (23g) and utterances with an initial adverbial complement (23f)). If it is assumed now that in these cases the filled C-position does not constitute a lexical category (that is non-lexical items occur in it), then in, for instance, in (23f) the pre-position of *te* offers two prosodic grouping possibilities. Either it is grouped with the verb in a PrWd or it is grouped in a PrWd with the preceding *não*. Grouping it together with the verb as (*não*) (*te vi*) would both violate the constraint ALIGN-Lex0-L (*te* not being a lexical category) and the constraint ALIGN-PrWd (the second PrWd (*te vi*) does not begin with a foot). The grouping of *te* with preceding *não*, as (*não te*)(*vi*), would neither violate ALIGN-Lex0-L nor ALIGN-PrWd (as a matter of fact both *não* and *vi* are stressed, cf. fn.7), but only a violation of ALIGN-Lex0-R. In this way we can provide an account for the obligatory preposition and obligatory phonological enclisis in EP in cases like (23f-g). This is illustrated in tableau (24) for (23f).

(24) *não te vi*	AL-Lex0-L	AL-PrWd	AL-Lex0-R
F (*não te*)(*vi*)			*
(*não*) (*te vi*)	*!	*	
(*não*) (*te*)(*vi*)		*!	*

In order to further determine the correct hierarchical ranking of the Alignment constraints, let us consider (23e). In (23e), the preposition of *me* would yield only one possibility: grouping it together with the verb. This grouping would violate both Align-PrWd and Align-Lex0-L. Both constraints must therefore in EP dominate Align-Lex0-R, in order to account for the obligatory postposition in (23e). Please recall that, as mentioned above, for BP a high-ranked constraint ALIGN-Lex0-R is assumed. This will account for the obligatory preposition in *me diga* cases in BP.

The postposition of *te* or *o* as in (23a and c) violates ALIGN-Lex0-R, but does not violate the higher-ranked constraints ALIGN-Lex0-L and ALIGN-PrWd. Preposing it, again, gives two possibilities. If the pronoun is grouped together with the verb (phonological proclisis, that is, as *(Eu)(te vi)* and *(O gato)(o apanhou)*), this entails a violation of ALIGN-Lex0-L and Align-PrWd. However, the grouping of the pronoun together with the preceding subject (phonological enclisis, that is, as *(Eu te)(vi)* and *(O gato o)(apanhou)*) would only violate ALIGN-Lex0-R, and, because Align-PrWd and Align-Lex0-L dominate Align-Lex0-R, is the optimal output when the pronoun occurs in preverbal position. It thus follows straightforwardly from the constraints and the constraint ranking assumed that pronouns in EP, whether they appear pre- or postverbally, are always phonologically enclitic. It should be observed, however, that the marked status of (23b) and (23d) in EP does not follow directly from the constraint hierarchy. A possible way to account for it, might be to assume that, all else being equal, grouping takes place preferably with the head from which it receives case.

In conclusion, the fact that both (23a/b) and (23c/d) are possible in EP, but **only** if they are phonologically enclitic, follows thus straightforwardly from the constraint ranking. The analysis presented here thus nicely captures the different surface order of the verb and the clitic in BP, which was shown to be problematical in a syntactic account. Moreover, it also straightforwardly formalizes, as mentioned, Brandão de Carvalho's insight that pronouns in EP, whether they appear pre- or postverbally, are always phonologically enclitic.

Conversely, because ALIGN-Lex0-R is high-ranked in BP, the preferred groupings for BP are the ones involving phonological proclisis: *(Eu)(te vi)* and *(O gato)(o apanhou)*. In this section we have presented the outlines of an analysis in Optimality theory of enclisis and proclisis. It has been argued that Optimality theory is better equipped to handle the domain aspects of the enclisis and proclisis phenomena discussed here. The problem of identifying the correct domains for the application of enclisis and proclisis which the prosodic theories discussed in section 2 were confronted with, has vanished in the present analysis. Furthermore, no empirically unmotivated predictions are made. Finally, we have argued that extending the Generalized Alignment theory to the syntax/phonology mapping, opens up a new perspective to account for phenomena that hitherto have proven reluctant to a purely syntactic account.

4. **Summary and discussion: sound change and markedness**

In this paper, we have discussed an example of sound change in the historical phonology of French, where we have shown not only that Optimality Theory by constraint reranking can adequately account for sound change, but also that the proposed analyses are not thwarted by drawbacks of previous nonlinear derivational accounts.

There is, however, one aspect that needs to be briefly discussed. In section 2, in the discussion of the edge-based analysis of the loss of enclisis, we briefly discussed the tentative principle (7), in which a link was made between the edge relevant for the edge-parameter and the edge where phrasal stress is located. For the sake of the argumentation, we will side-step for the moment the empirical adequacy fo the principle in (7), and concentrate on the role of principles such as these in Optimality Theory. Until the advent of constraint-based theories, such as Optimality Theory, it has been a common practice among phonologists to add to the Theory of Markedness by formulating principles such as the one in (7). The obvious question that arises now, of course, is how such principles can translate into Optimality Theory. In other words, does Optimality Theory allow for a Theory of Markedness based on similar principles, or should the unmarked state of affairs in one way or another follow as a natural result from the hierarchy? The latter position seems unmotivated as there is nothing inherent in the formalism of constraint-hierarchies that would prefer one ranking above the other. Also, each constraint itself, because it is considered to be part of Universal Grammar, is already in a sense a statement about markedness. Nevertheless, one would like to be able to express more formally that some rankings are more natural or less marked than others. Therefore, for the case at hand, we could, if the first position is adopted, and, assuming that phrase-initial and phrase-final stress can be accounted for by edge-aligning feet with Prosodic Words, propose a principle like (24).

(24) *If, in a language, the constraint ALIGN (PrWd, L, Ft, L) is higher ranked than the constraint ALIGN (PrWd, R, Ft, R), then, the constraint ALIGN (LEX0, L, PrWd, L) is, in the unmarked case, higher ranked than the constraint ALIGN (LEX0, R, PrWd, R).*

It is, however, easy to see that this way of dealing with markedness can easily become very complicated. For instance, if we want to express the generalization (cf. Hayes (1993)) that certain languages (that is those that have iambic stress rules) tend to have quantitative rules such as rhythmic vowel lengthening, consonant gemination, vowel reduction and vowel deletion, whereas other languages (those having trochaic stress rules) do not, this would become, if expressed in a similar way as (24), very intricate. It is clear that especially in this area a lot of future research is needed.

References

Adams, M. 1987. *Old French, Null Subjects and Verb Second Phenomena.* Ph.d.Diss., University of California, Los Angeles.

Bauer, B. 1992. *Du latin au français: le passage d'une langue SOV à une langue SVO.* Ph.d.Diss., Nijmegen University, Nijmegen.

Bichakjian, B. 1988. *Evolution in Language.* Ann Arbor, Kuroma.

Brandão de Carvalho, J. 1989. "Phonological conditions on Portuguese clitic place ment: evidence for stress and rhythmical patterns," *Linguistics*, 27:405-436.

Einhorn, E. 1974. *Old French.* Cambridge, Cambridge University Press.

Evers, V. 1994. *Where are Old French enclisis and proclisis?* MA-thesis, Nijmegen University.

Garde, P. 1968. *L'Accent.* Paris, Presses Universitaires de France.

Hayes, B. 1990. "Precompiled Phrasal Phonology," *The Phonology-Syntax Connection*, ed. by S. Inkelas and D. Zec, 85-108. Chicago, University of Chicago Press.

Jacobs, H. 1993. "The Phonology of Enclisis and Proclisis in Gallo-Romance and Old French," *Linguistic Perspectives on the Romance Languages*, ed. by W. Ashby, M. Mithun, G. Perissinotto and E. Raposo, 149-163. Amsterdam/Philadelphia, John Benjamins.

Jacobs, H. 1994. "How optimal is Italian stress?," *Linguistics in the Netherlands 1994.*

ed. by R. Bok-Bennema and C. Cremers, 61-70. Amsterdam/Philadelphia, John Benjamins.

Kayne, R. 1991. "Romance Clitics, Verb Movement, and PRO," *Linguistic Inquiry*, 22:647-686.

Kenstowicz, M. *Phonology in Generative Grammar.* Cambridge MA/Oxford, Blackwell, London.

Kukenheim, L. 1971. "Le rôle de la prosodie dans l'histoire de la langue française," *Mélanges de philologie romane Jean Boutière*, vol. 1, ed. by I. Cluzel and F. Pirot, 317-331. Liège, Editions Soledi.

Leeuw, F. van der. 1994. "Cliticization, Stress and Phonological Words in European Portuguese: An Optimal(ity) Approach," Ms. University of Amsterdam.

Löhken, S. 1996. *Entwicklung der deutschen Wortprosodie.* Ph.D.Diss., Berlin.

Madeira, A. M. 1993. "Clitic-second in European Portuguese," *Probus*, 5:155-174.

McCarthy, J. and A. Prince. 1993a. *Prosodic Morphology I*, Technical Report#3 of the Rutgers Center for Cognitive Science, Rutgers University.

McCarthy, J. and A. Prince. 1993b. "Generalized Alignment." Ms.

Nespor, M. and I. Vogel. 1986. *Prosodic Phonology.* Dordrecht, Foris Publications.

Prince, A. and P. Smolensky, 1993. *Optimality Theory. Constraint Interaction in Generative Grammar.* Technical Report#2 of the Rutgers Center for Cognitive Science, Rutgers University.

Selkirk, E. 1980. "Prosodic Domains in Phonology: Sanskrit Revisited," *Juncture*, ed. by M. Aronoff and M-L. Kean, 107-129. Saratoga, Anma Libri.

Selkirk, E. and T. Shen. 1990. "Prosodic Domains in Shanghai Chinese," *The Phonology-Syntax Connection*, ed. by S. Inkelas and D. Zec, 313-337. Chicago, Chicago University Press.

Zwicky, A. 1987. "French Prepositions: No peeking," *Phonology Yearbook* 4, 211-227.

Department of French
Nijmegen University/Free University Amsterdam
CLS/HIL
Erasmusplein 1
6500 HD Nijmegen
Netherlands E-mail: h.jacobs@let.kun.nl

Evidence for the p-center syllable-nucleus-onset correspondence hypothesis

Peter M. Janker
Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung, Jägerstraße 10/11, D-10117 Berlin
janker@fas.ag-berlin.mpg.de

Abstract

The question of whether or not the vowel onset is the parameter responsible for the rhythmical alignment in timing speech utterances still remains unclear. The experiment described here supports the hypothesis of a congruence between the mental representation of the syllable-nucleus-onset and the moment of occurrence or so-called p-center. Utterances consisting of phonologically identical but physically distinct syllable rhyme and different syllable initial consonance were presented in a synchronous tapping experiment. The results show that the vowel onset estimates the location of the tapping position and the differences between the stimuli as good as other models, suggesting a clear correspondence between p-center and vowel onset, thereby strongly supporting the hypothesis.

Introduction

The prosodic aspects of speech have gained more and more interest within recent years. Although the main interest was concerned with intonation patterns and tonal aspects the question of which parameters influence the rhythmic structure of utterances was not totally neglected. Since Terhardt/Schütte [1], Morton et al. [2] and Marcus [3] it is known that an acoustic (speech-) signal, like a syllable, is not perceived by its physical onset but somewhat later in time depending on its properties and/or physical make-up. Since then various investigations [4-13] have been undertaken and models proposed [3, 4, 10, 12, 14] to estimate this so-called 'Ereigniszeitpunkt', 'moment of occurrence or p-center' and to define the parameters influencing it. A single parameter could not be found, but it seems that the duration and/or physical make-up of the initial consonance of a syllable or the transition between consonance and vowel is of more importance than the syllable rhyme or coda.

But the question of which parameters influence the rhythmic structure of utterances was not only discussed in the light of the p-center/'Ereigniszeitpunkt' approach but also addressed by others, such as Allen [15], Rapp [16], Hollister [17] and Meyer [18]. Looking for physically measurable equivalents of the rhythm

beat, syllable beat, 'Silbenschlag' etc. these researchers asked subjects to tap[1] synchronously while listening to or speaking lyric / poetic utterances. They all agreed on that the 'beat' is somewhat near the end of the syllable initial consonance and the beginning of the following vowel as Meyer already stated in his 1898 paper "Über den Takt".
In some of my later work [9, 19-21] I argued for the hypothesis of a congruence of the p-center and the mental representation of the syllable-nucleus-onset and pointed out that the mismatch between tapping location and vowel onset found in these investigations might be due to not taking into account the phenomenon of anticipation occurring in tapping experiments (see 'Why anticipation occurs' below).

Experimental Design

The experiment described here is undertaken to test the hypothesis that a congruence exists between the mental representation of the syllable-nucleus-onset and the p-center. In discussing results of former experiments I also suggested, that – besides neglecting the anticipation phenomenon – somewhat misleading results might occur due to the laboratory situation and the fact that some of the stimuli used in tapping and adjustment experiments might be somehow artificial lacking any compensational variations common in normal utterances, as they are, for the purpose of the experiment, varied in only one parameter and kept constant in the others. This experiment uses only segments of naturally spoken utterances as stimulus material.

Stimuli

The set of stimuli used here consists of eleven words with #CVC structure and two interjections [pstʰ, sːtʰ]. As the initial part of a syllable, the syllable head, has the main influence on the perception of the p-center position and the rest of the syllable has a lesser influence, the stimuli produced have different initial consonants [ʔ, p, f, k, h, kʰ, l, m, pʰ, ʁ, tʰ] and a constant syllable rhyme [astʰ], phonotactically the same but physically due to the natural generation different with all the compensational information of a natural utterance.
The words were well pronounced (explicitly demonstrated) in focus position within the frame sentence <Ich habe das Wort _____ gesagt.> (I said the word _____.)[2] and recorded in a soundproof studio using an Electro Voice 631B microphone and a DAT recorder. The interjections were produced as if the speaker wants to get someones attention [sːtʰ] or causes someone to be quiet [pstʰ]. The re-

1) For clarification: Alignment is the caused relation between two events (here sequences). Adjustment is an alignment achieved by changing the relation of two presented event strings, tapping an alignment achieved by performing an action (the tap) to a presented event string. The adjustment or tapping position (alignment position) is this alignment expressed in relation to an arbitrarily chosen origin. The location of the adjustment or tapping position is this measurement expressed in relation to other properties. Adjustment or tapping position are NOT the p-center position, rather they are induced by the perception of the p-center thus reflecting any change in the location of the p-center position.

2) N.b. that the position in the German phrase is not sentence final.

cording was transmitted to a Macintosh using Digidesign Audiomedia II (SP/DIF input), downsampled to 20 kHz and segmented using Signalyze on the Macintosh. The signal segments of interest were cut out of the sound stream and saved in a separate file for presentation. Figures 1 to 3 show as examples the

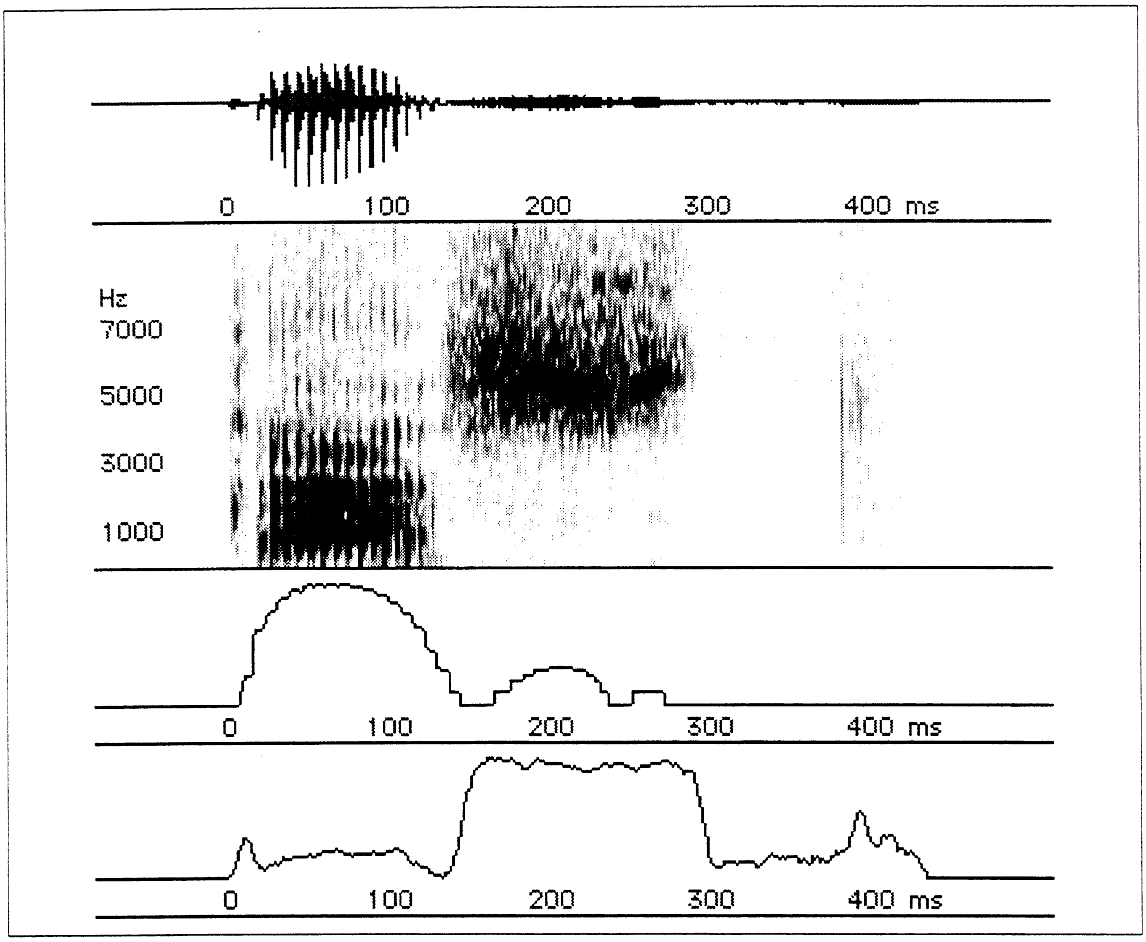


Figure 1: *<gast> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <gast> with an overall duration of 405 ms.*

shortest stimulus [kastʰ] with a duration of 405 ms, the longest stimulus [ʁastʰ] with an overall duration of 567 ms and the interjection [sːtʰ] with a duration of 416 ms. The figures show the amplitude as well as a wide band sonagram (10 kHz, pre-emphasis, 300 Hz window), the amplitude envelope (30 ms window) and the zero crossing (10 ms window) for the respective stimuli. An artificial 5 ms, 1 kHz tone burst perceived as a click signal was used as control stimulus.

Subjects

29 subjects (19 female, 10 male) took part in a series of experiments including this experiment. None of them had participated in experiments on rhythm perception before. To become familiar with the computer, the stimulus presentation program and the data acquisition procedure every subject had a 10 minute introduction using a click signal, the sound [pst] and the word <schwimmst> [ʃβɪmst] as stimulus material.

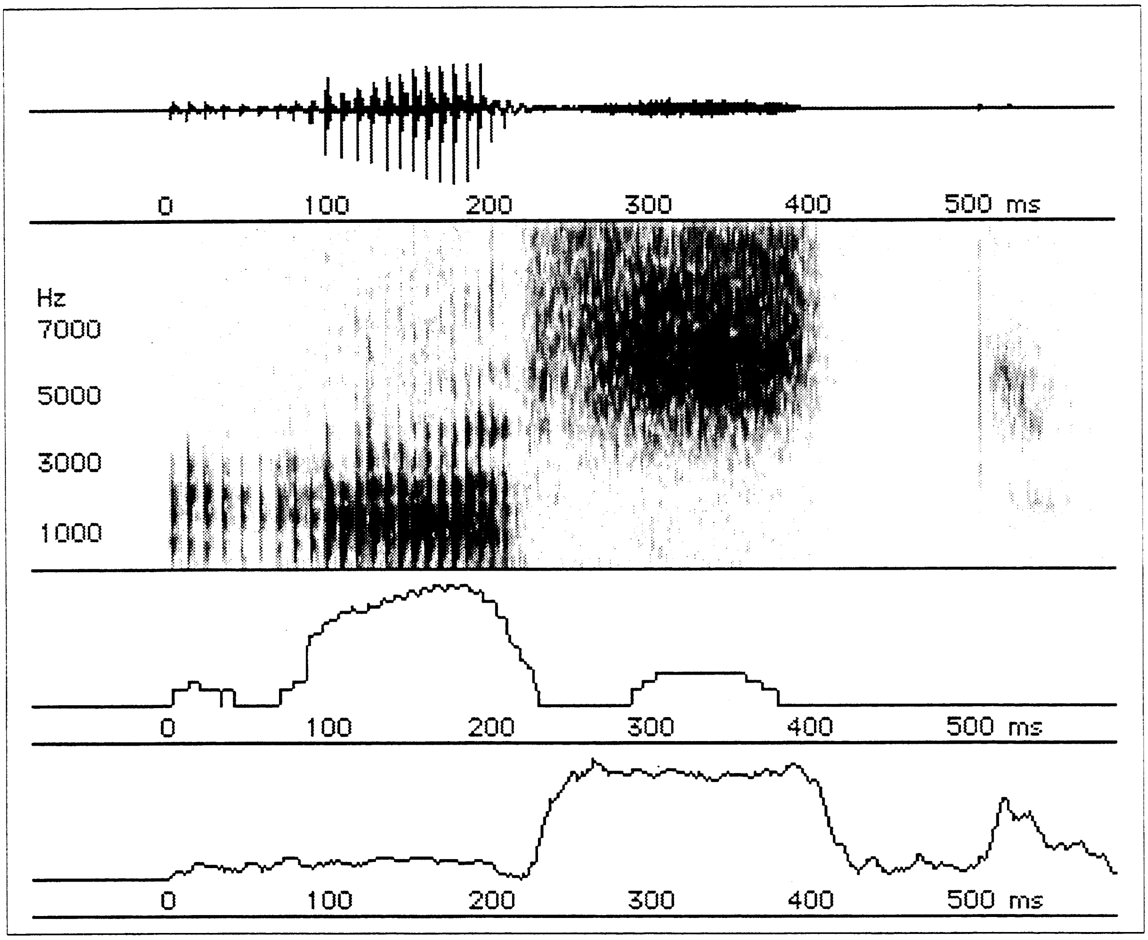


Figure 2: <rast> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <rast> with an overall duration of 567 ms.

Method

The stimuli were presented using Sennheiser HD 480 II headphones under computer control (Compac Deskpro 486/33, Data Translation DT2821SE DA-converter, Behringer PEQ 305 filter, Sony F535R) with 20 kHz sample rate and low-pass filtered at 6 kHz (24 dB/oct). The subjects had to perform a synchronisation task by tapping to sequences of binaurally presented stimuli. A sequence consisted of 15 repetitions of the same stimulus with an inter stimulus interval of 700 ms. The interval between sequences was 1400 ms. The stimulus sequences were randomized and grouped in blocks of 10. The subjects started the presentation of the next block by pressing the return key. A sequence was repeatedly presented as long as the subject did not start to tap. Each stimulus sequence was given four times with at least two different intermediate sequences. To register the taps a 5 x 10 cm capacitive sensory field was used. For analysis the taps to the first three and the last two presentations within a sequence were omitted (leaving 16,240 taps). Figure 4 gives a schematic description of the measurement procedure; a more detailed description of the experimental design and measurement procedure can be found in [6].

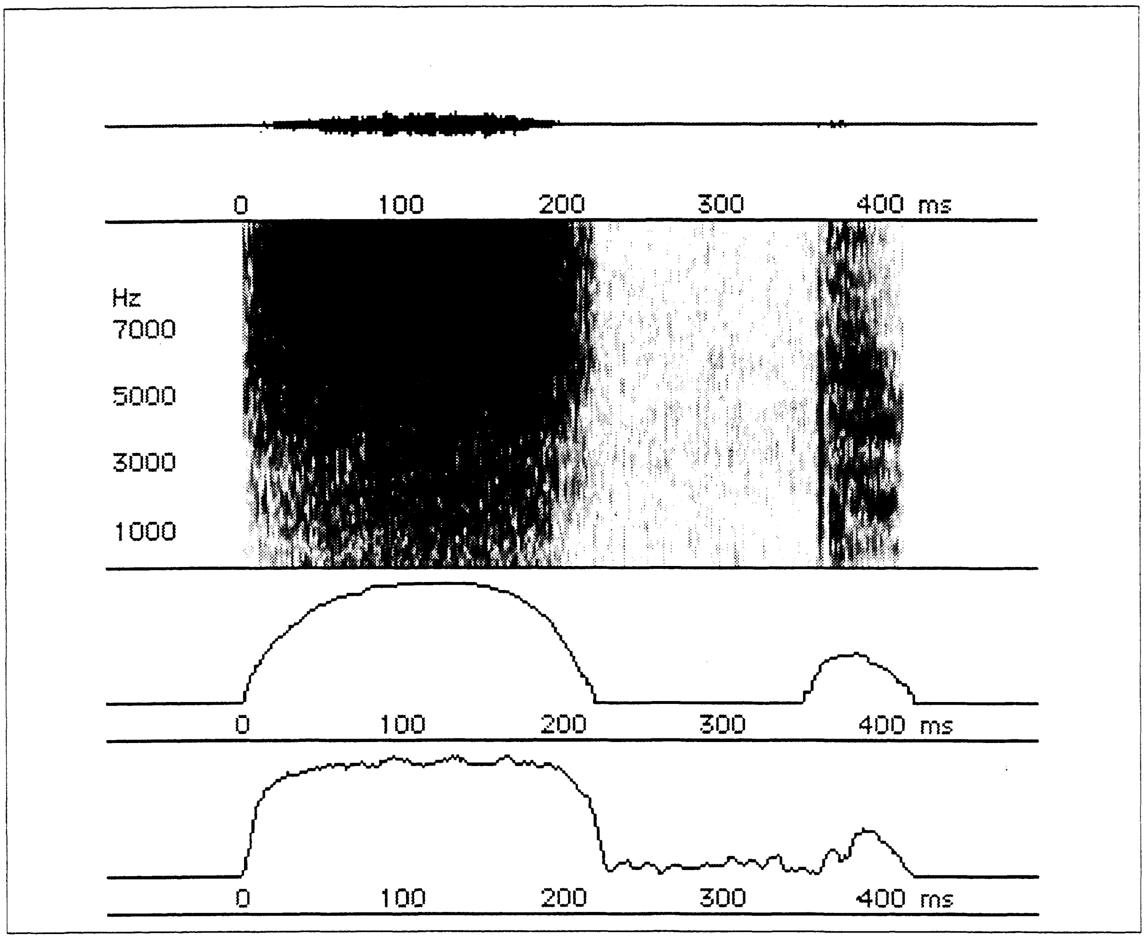


Figure 3: *<sst> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <sst> with an overall duration of 416 ms.*

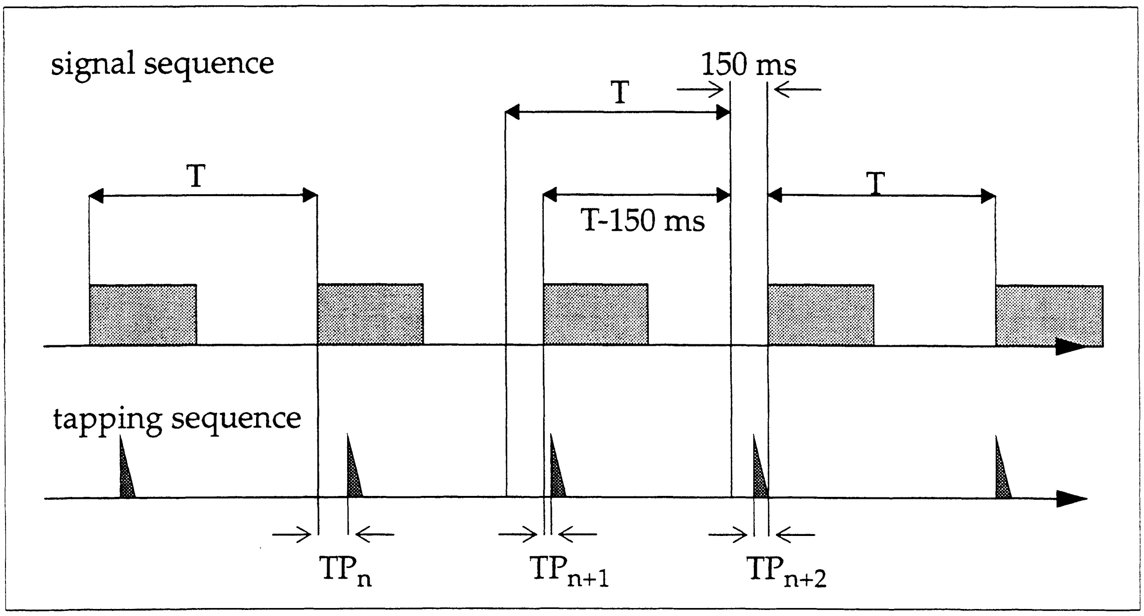


Figure 4: *Schematic description of tapping position measurement with tapping position TP_n in relation to the signal onset of the respective stimulus presentation (T = inter stimulus interval, TP_n = tapping position for the nth stimulus presentation, 150 ms offset for technical reasons)*

RESULTS AND DISCUSSION

As the initial consonants of the stimuli presented had been different in duration and the distribution of spectral energy the repeated measurement design analysis gives a significant effect ($F = 99.79$, $p < .001$) for the factor stimulus as well as for the factor subject ($F = 18.05$, $p < .001$) with a significant interaction ($F = 1.49$, $p < .001$), as four subjects do not show a significant stimulus effect (post hoc Scheffé (.01)).
In tapping experiments the subjects have to perform three different tasks simultaneously. First to adopt a uniform rhythm given by an event string of identical stimulus repetitions, second to act to these events by tapping synchronously to the events and third to judge whether or not they succeeded in doing so. Therefore, a large difference in subject responses is typical for tapping experiments. To give an impression of the variability found in experiments like this, figure 5

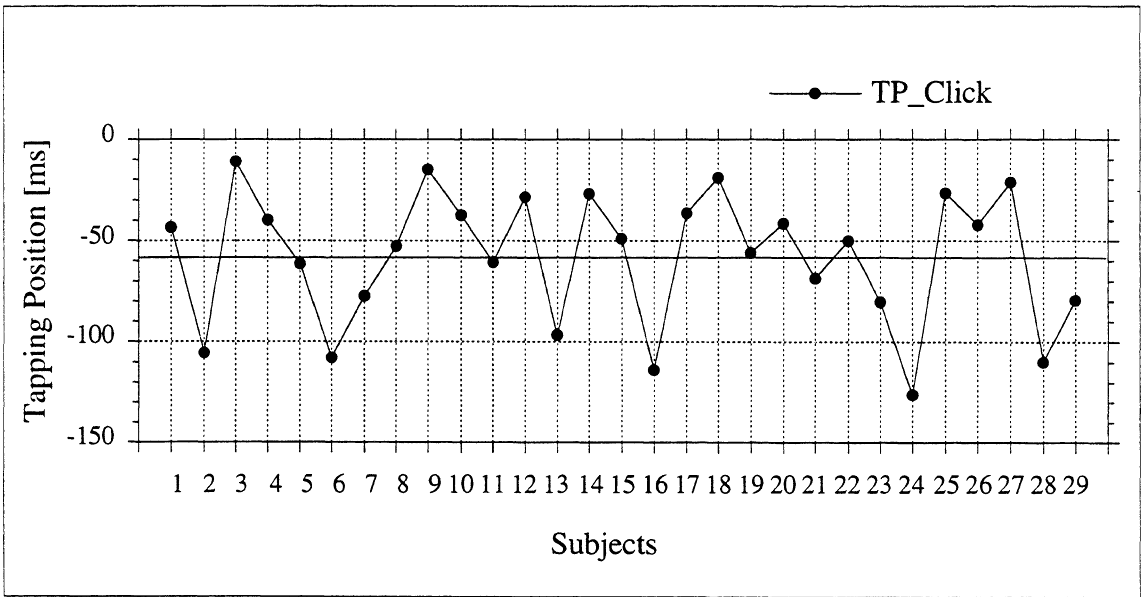


Figure 5: *Mean tapping positions TP for the control stimulus (click signal) for all subjects with indicated overall mean (solid line)*

shows the tapping positions for the control stimulus, a click signal of 5 ms duration, split by subject. Pooled over all subjects the taps precede the click signal by about 58 ms (SD 46.5 ms) which is in good agreement with findings in my former experiments and values found elsewhere. This effect of anticipation, known since Dunlop [22], was recently investigated in more detail by Radil et al. [23], Aschersleben/Prinz [24], and Gehrke [25].

Why anticipation occurs

One should not neglect the amount of estimations necessary to synchronise an action, the tapping of a hand, to a perceivable forthcoming acoustic event (in more detail Mates [30]). Figure 6 schematizes the timing relations of the involved contributors and gives an impression of what kind of estimates are necessary for a synchronized tapping task and why anticipation occurs.
The external world is only discoverable through our senses and the mental representations of sensible events. In figure 6 the external world is represented by the

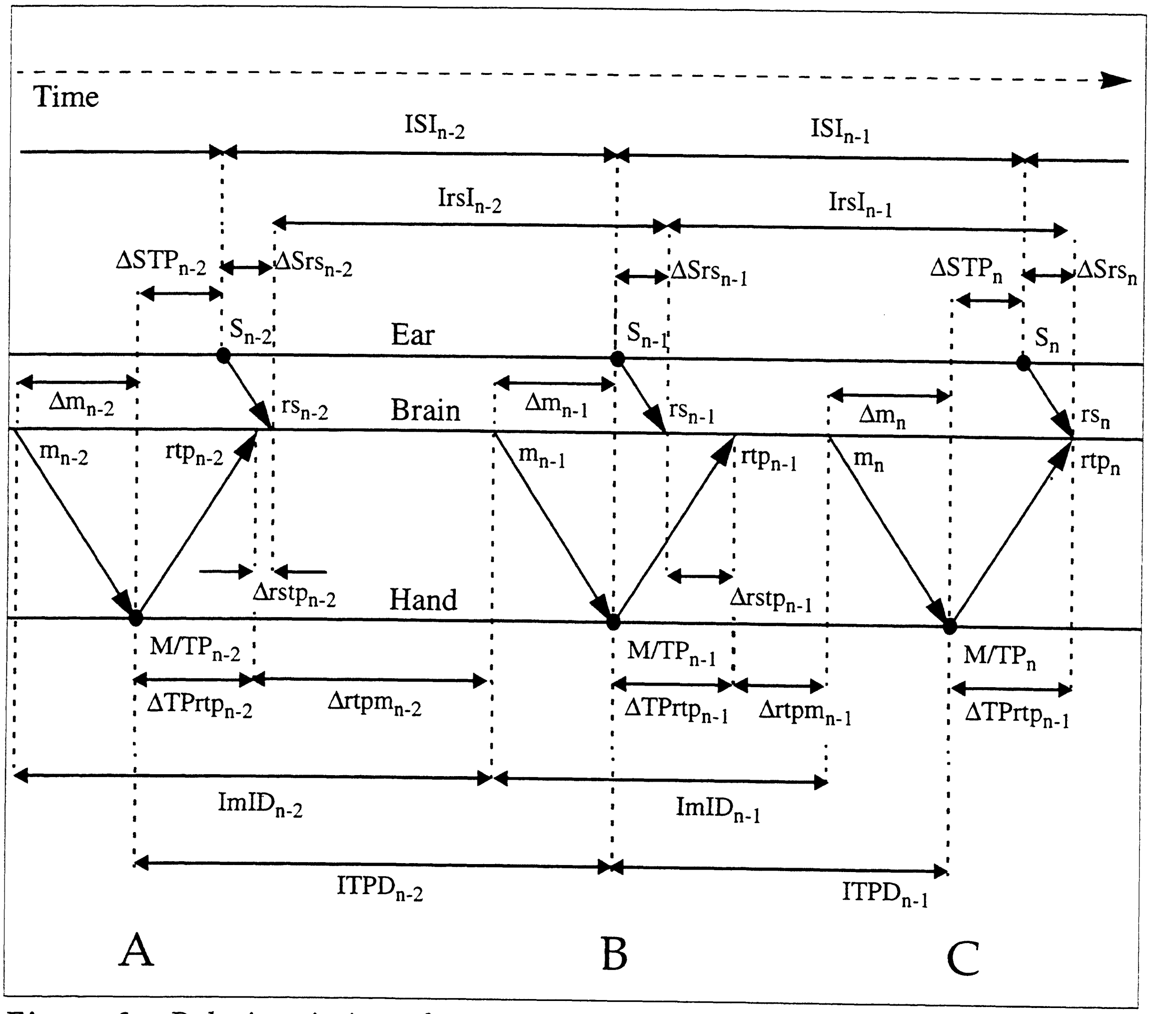


Figure 6: *Relative timing of external events (signal S; movement M; tap TP) and internal representations (rs, m, rtp) with estimated intervals (Δm, IrsI, ΔSrs, ΔTPrtp, Δrtpm, Δrstp, ImID), inter-tap distance(ITPD) and observable anticipation (ΔSTP).*

time event axes of ear and hand, the internal world by the time event axes of the brain (solid continuous lines). The ear-line is depicted closer to the brain, since the afferent information of the ear about a speech signal (S) reaches the brain earlier than the information of the hand about a tapping action (TP), as the ear is closer to the brain than the hand.

As we know subjects do not try to produce a similar time interval in between two events, but try to produce equally timed p-center alignments. Hence, they have to identify the inter-stimulus interval (ISI) of the signal sequence from the interval of the mental representations of successive perceived signals (IrsI). Furthermore to have a coincidence of a signal S and a tap TP (as in figure 6 case B) in the external world the subjects have to estimate the intervals ΔSrs (between the speech signal S and its representation rs), ΔTPrtp (between the tap TP and its representation rtp), Δrstp (the difference between these intervals) and Δm (between the mental representation of a movement instruction m and the actual movement M of the finger), to decide on the interval Δrtpm (between the last tap representation and the representation of the next movement instruction m) for the right moment in time to give the next movement instruction. A coincidence of signal S and tap TP in the external world then occurs, if the subject is able to keep the interval Δrstp between the mental representation rs of the signal S and the feedback rtp of the tap TP, derived from the estimated intervals, constant at the magnitude of the subject specific difference ΔSrs minus ΔTPrtp. But the real world coincidence task does not only involve a lot of estimations, it also has to be done without any reliable feedback, as in the artificial situation of a tapping experiment an external validation of the coordination (like in tennis: hit or miss) of whether or not the task was performed successfully, is due to the experimental implications not present, and internal information for validation other than the estimates themselves is not available.

Real world events involving human action seem to coincide only if external information about the success of the performance is available and the performer able to compensate for the different time delays of the afferent and efferent information.

On the other hand, if we suppose that the task is accomplished by a coincidence of the mental representations rs and rtp of the signal and the tap (as in figure 6 case C) the subjects have only to make sure that there is no difference in time between the two representations. This can be done by adopting the IrsI interval as timing interval (ImID) for the movement instruction adjusted in a way that Δrstp gets minimized. This minimization process still involves a lot of estimation (Mates [30]) but no external information is necessary for validation; the task is successfully performed if the two event representations can no longer be differentiated in time. As a result the tap naturally has to precede the signal and a real world mismatch ΔSTP, the anticipation of the signal by the tapping subject, can be observed.

Normalization

Therefore, the anticipation measured above was also used to compensate for the effect of anticipation. That is, the tapping alignments of the subjects have been normalized by adding their individual anticipation[1], as measured in relation to the click signal, to the raw TP-values. This normalization reduces subject differences with respect to anticipation only, it does not alter any other individual differences like i.e. the amount of intra-subject variability, it especially does not change a possible stimulus effect. It also shifts the location of the overall mean tapping position by the amount of the mean anticipation into the positive direction with respect to the stimulus onset, that is further into the stimulus. Figure 7 gives the distribution of tapping responses for all subjects before and after normalization.

Out of the 29 subjects three seem to be unable to perform the task as intended. They show a much larger variation than the other subjects (figure 7 is somewhat misleading in that respect showing the responses pooled over all stimuli but the click). Figure 8 gives the 75% and 90% percentiles of the variances per stimulus for all subjects. There is a noticeable difference in the amount of variability for the three marked subjects in relation to all others. Their 75% percentiles are larger than the 80% percentile of the 90% percentiles of the other subjects. Furthermore

1) This does not assume that the 'individual' anticipation is an unalterable constant throughout the individuals entire life-span, but merely that it remains constant for the time of the experiment. If someone would participate in several experiments there would be an 'individual' anticipation value for every experiment.

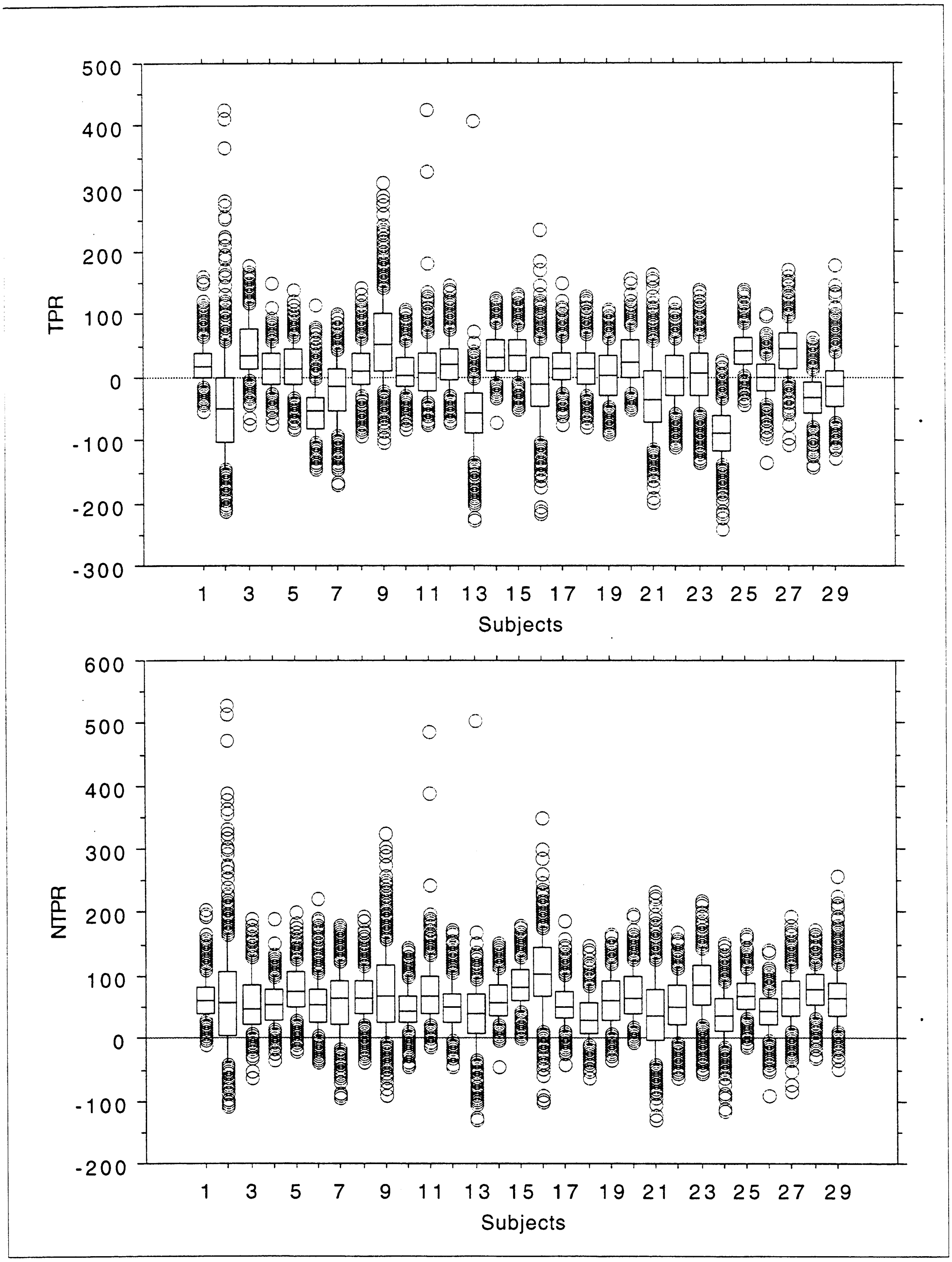


Figure 7: Box plot of the tapping positions TPR and normalized tapping positions NTPR with median (middle line), inter quartile range (box), indicated 10% and 90% percentiles and outliers.

they show significant differences (Scheffé (.01)) to other subjects in their inter quartile range, being not significantly different from one another, with the remaining 26 also not showing any significant difference in their inter quartile ranges. Figure 9 shows the 75% and 90% percentiles of the inter quartile range per

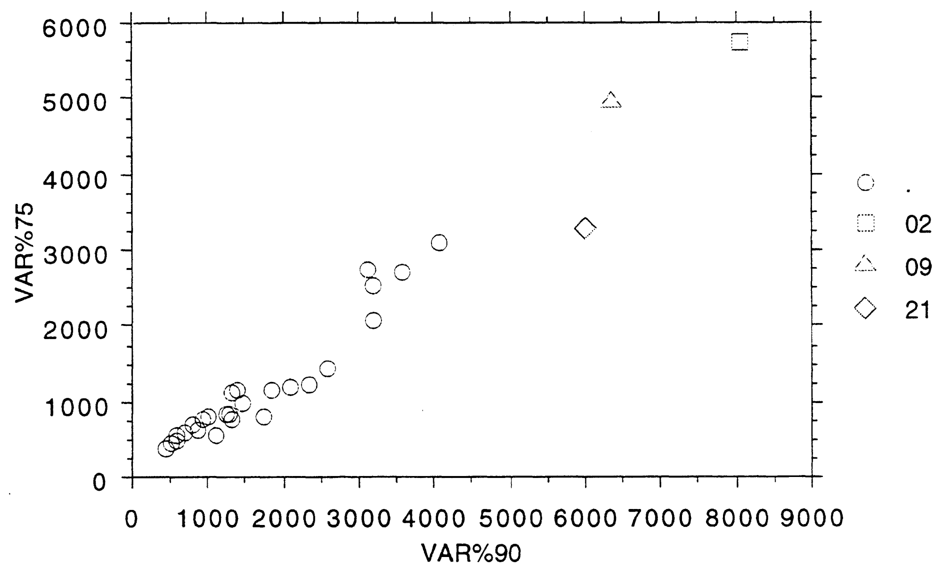


Figure 8: *75% and 90% percentiles of the variances per stimulus of the normalized tapping positions NTP for all subjects with three subjects marked showing a noticeable variability discrepancy.*

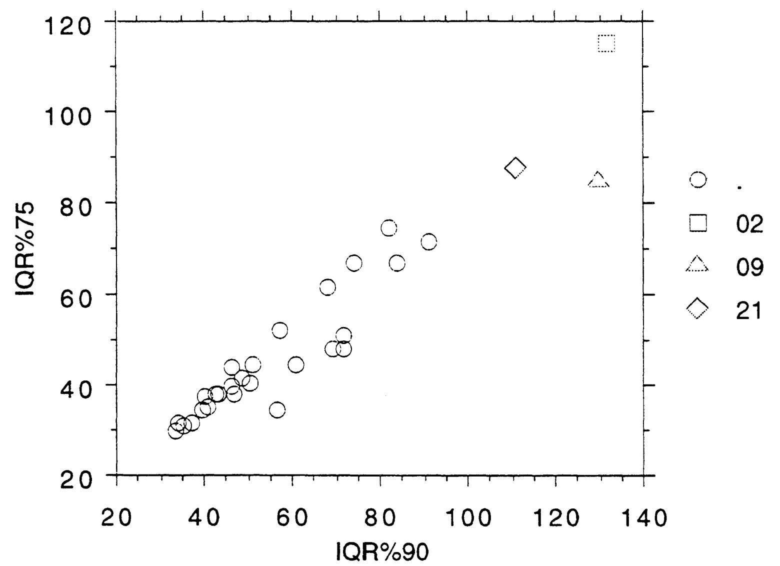


Figure 9: *75% and 90% percentiles of the inter quartile range per stimulus of the normalized tapping positions NTP for all subjects with three subjects marked showing a noticeable variability discrepancy.*

stimulus for all subjects, figure 10 the box plot of the inter quartile range. For further analysis the three marked subjects were omitted. Visual inspection seems to suggest the exclusion of one or two other subjects as well, but they did not come up in the analysis as being significantly different and remained therefore in the data set.

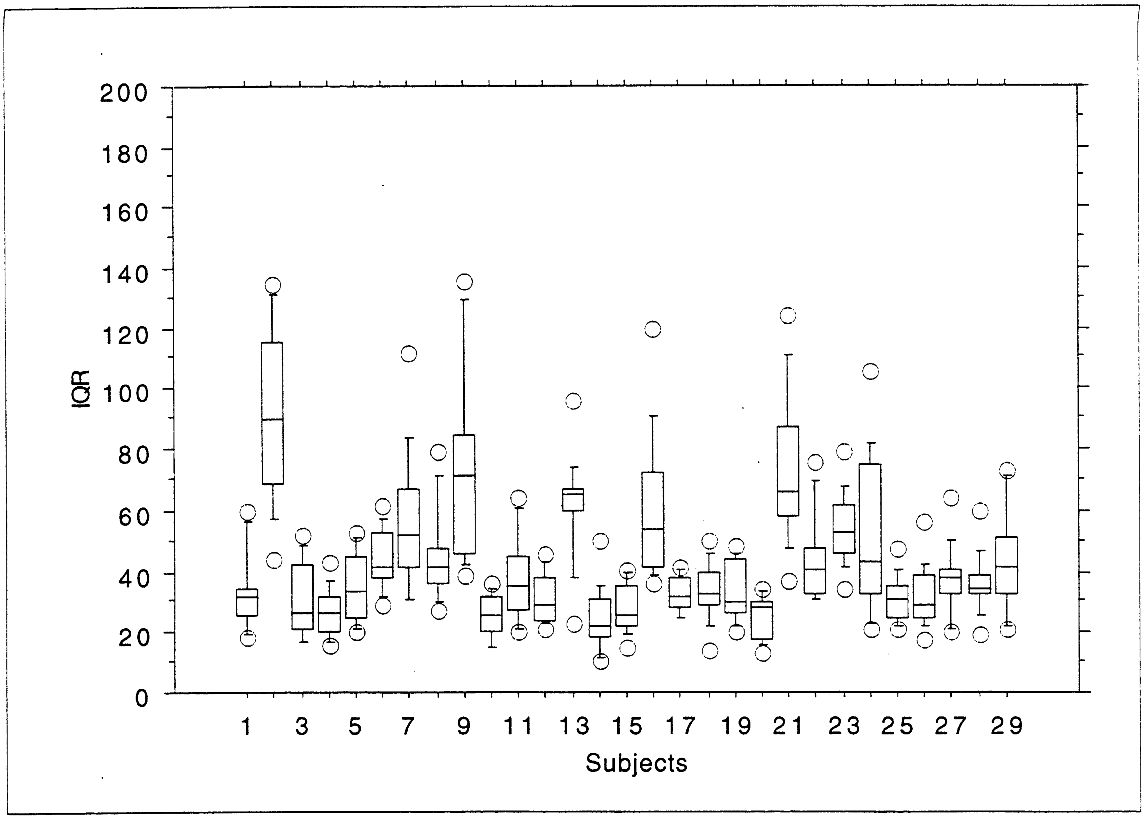


Figure 10: *Box plot of the inter quartile ranges IQR per stimulus for all subjects with median (middle line), inter quartile range (box), indicated 10% and 90% percentiles and outliers.*

Stimulus effects

For the remaining 26 subjects the repeated measurement analysis still gives, as expected due to the differences in the syllable initial consonants, a significant effect ($F = 114.58$, $p < .001$ / for words only $F = 115.41$, $p < .001$) for the factor stimulus as well as for the factor subject ($F = 23.03$, $p < .001$ / for words only $F = 18.89$, $p < .001$) with a significant interaction ($F = 1.46$, $p < .001$ / for words only not significant), due to the four already mentioned subjects which do not show a significant stimulus effect for the post hoc Scheffé (.01) test. Although the syllable rhyme has been kept phonotactically constant and the initial consonance has the main impact on the p-center, the significant differences for the factor stimulus are caused by the respective stimulus as a whole since the actual realization of the syllable rhyme differs physically. The locations of the normalized tapping positions NTP (with indicated standard deviation in the positive direction) can be seen in figure 11 in relation to the duration of the initial consonants, the vowel as syllable nucleus and the coda (see table I below). As can be seen the mean tappings are located around the consonant to vowel transition somewhat before and after the segmented syllable-nucleus-onset with no obvious standard deviation irregularities (mean sd =37.02 ms). Taken into account that the depicted locations are normalized for anticipation, this is in good agreement with the findings of my former experiments and the data reported in most of the above mentioned investigations.

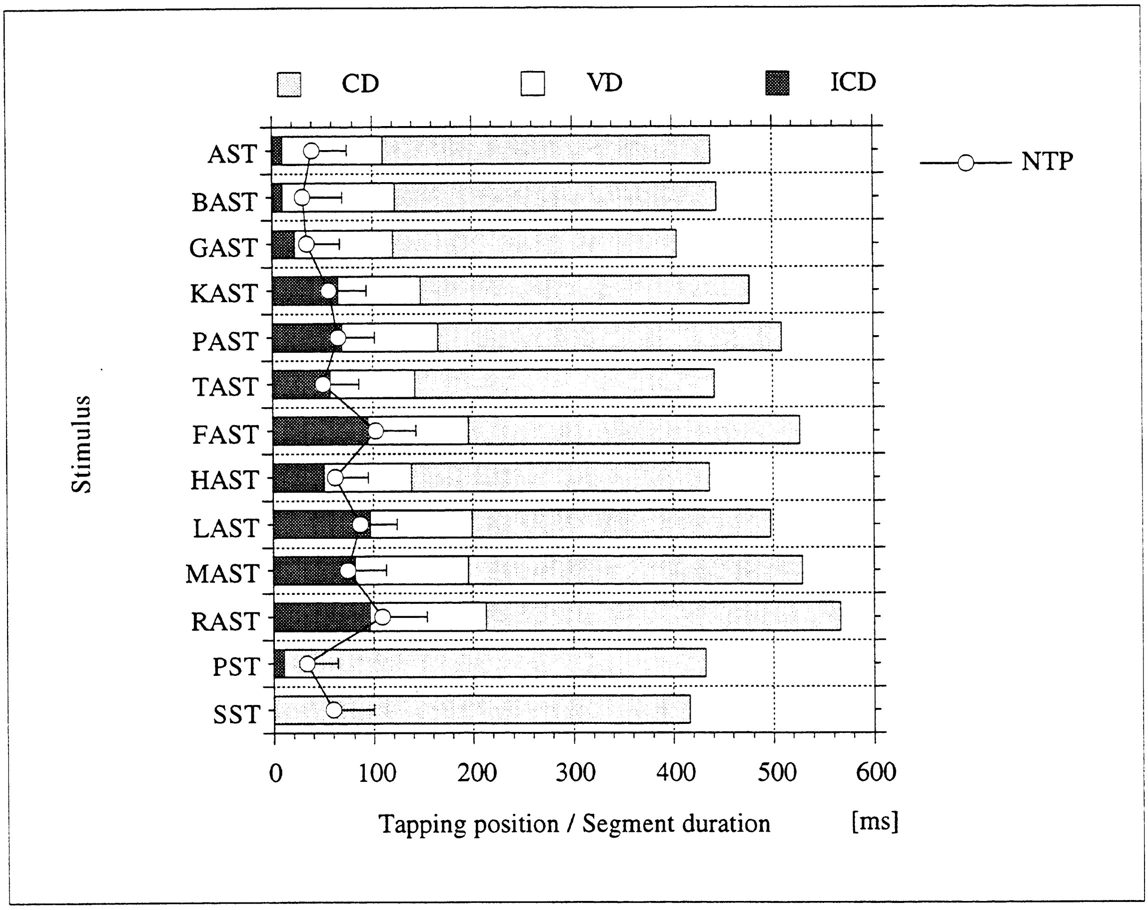


Figure 11: *Normalized tapping positions NTP for all stimuli, with indicated standard deviation in positive direction, in relation to the initial consonance duration (ICD = syllable-nucleus-onset), nucleus duration (VD) and coda duration (CD).*

Model estimates and correlations

To assess, how good the measured syllable-nucleus-onset as a predictor of the p-center alignment fits with the location of the normalized tapping positions and to compare it to other estimates of the p-center position for some of the above mentioned models their estimates and the corresponding correlations were computed. Table I gives the measured values as well as the model specific estimates for the models of Schütte (SHT), Marcus (MPC), Köhlmann (KLM), Howell (HWL) and Pompino-Marschall (BPM), table II the corresponding correlation matrix. The correlations between the normalized tapping positions and the model estimates as well as between the normalized tapping positions and the syllable-nucleus-onset are quite high (Fisher's r to z cor > .827, p < .0009). As expected the syllable nucleus, coda or rhyme do not correlate with either the tappings or any of the models, but the rhyme (nucleus + coda) correlates more highly with the coda (p < .0001) than with the nucleus (p = .0084) showing that the nucleus has to be more stable while the coda is more variable that is more open for compensational effects. Figures 12 and 13 show the correlations between the normalized tapping positions and the model estimates for the word stimuli as well as the syllable-nucleus-onset. The correlations are only calculated for the word stimuli as some of the models rely on entities not available with the interjections.

Table I: Measurements and model estimates

Normalized tapping positions NTP with standard deviation SD, measured duration of initial consonance ICD, nucleus VD, coda CD, syllable rhyme RM and estimated alignment values for the models of Marcus MPC, Pompino-Marschall BPM, Howell HWL, Schütte SHT, and Köhlmann KLM.

Stimulus	NTP	SD	ICD	VD	CD	RM	MPC[1]	BPM	HWL	SHT	KLM
AST	39.54	35.3	10	101	328	429	113.8	14.5	68.8	12.0	10.3
BAST	30.16	39.7	10	113	322	435	115.2	15.1	64.4	12.9	9.1
GAST	34.40	33.3	23	98	284	382	110.5	28.6	82.6	27.4	13.7
KAST	56.11	36.9	66	83	328	411	145.7	61.1	117.	60.0	34.0
PAST	65.28	36.5	70	96	343	439	155.2	64.2	117.5	73.3	35.3
TAST	50.44	35.3	58	85	299	384	133.7	51.0	109.3	60.4	33.6
FAST	102.70	40.2	102	94	331	425	172.5	104.6	152.6	106	47.0
HAST	62.36	32.2	59	80	298	378	132.9	48.9	104.8	60.2	26.1
LAST	87.03	36.8	90	109	298	407	160.3	77.9	147.2	46.4	58.1
MAST	74.12	38.2	82	113	334	447	165.1	74.6	137.4	40.8	51.7
RAST	108.30	45.2	96	117	354	471	180.2	110.2	174.3	98.3	58.2
PST	32.81	31.5	6	0	422	422	109.4	33.9	122.6	15.8	8.4
SST	59.78	40.1	0	0	416	416	104.0	69.2	123.4	41.3	39.6

1) Can be distinctly improved in relation to alignment location depending on a post hoc optimization of the so-called 'arbitrary' constant (see text)

Table II: Correlation matrix for measurements and estimates

Normalized tapping positions NTP for words, duration of initial consonance ICD, nucleus VD, coda CD, syllable rhyme RM, estimates of the models of Marcus MPC, Pompino-Marschall BPM, Howell HWL, Schütte SHT, Köhlmann KLM, p-values for the framed correlations: SHT/NTP p = .0008, SHT/ICD p = .0005, all others p < .0001

	NTP	ICD	MPC	BPM	HWL	SHT	KLM	VD	CD	RH
NTP	1.000	.932	.949	.969	.963	.828	.907	.243	.458	.446
ICD	.932	1.000	.957	.969	.969	.842	.952	.048	.345	.277
MPC	.949	.957	1.000	.970	.966	.805	.946	.262	.577	.543
BPM	.969	.969	.970	1.000	.986	.883	.927	.164	.456	.410
HWL	.963	.969	.966	.986	1.000	.819	.969	.215	.408	.397
SHT	.828	.842	.805	.883	.819	1.000	.688	-.180	.415	.230
KLM	.907	.952	.946	.927	.969	.688	1.000	.271	.347	.376
VD	.243	.048	.262	.164	.215	-.180	.271	1.000	.397	.731
CD	.458	.345	.577	.456	.408	.415	.347	.397	1.000	.916
RH	.446	.277	.543	.410	.397	.230	.376	.731	.916	1.000

The simple threshold model – compared to some of the other models – of Schütte, taking overall duration and amplitude envelope into account, shows at .828

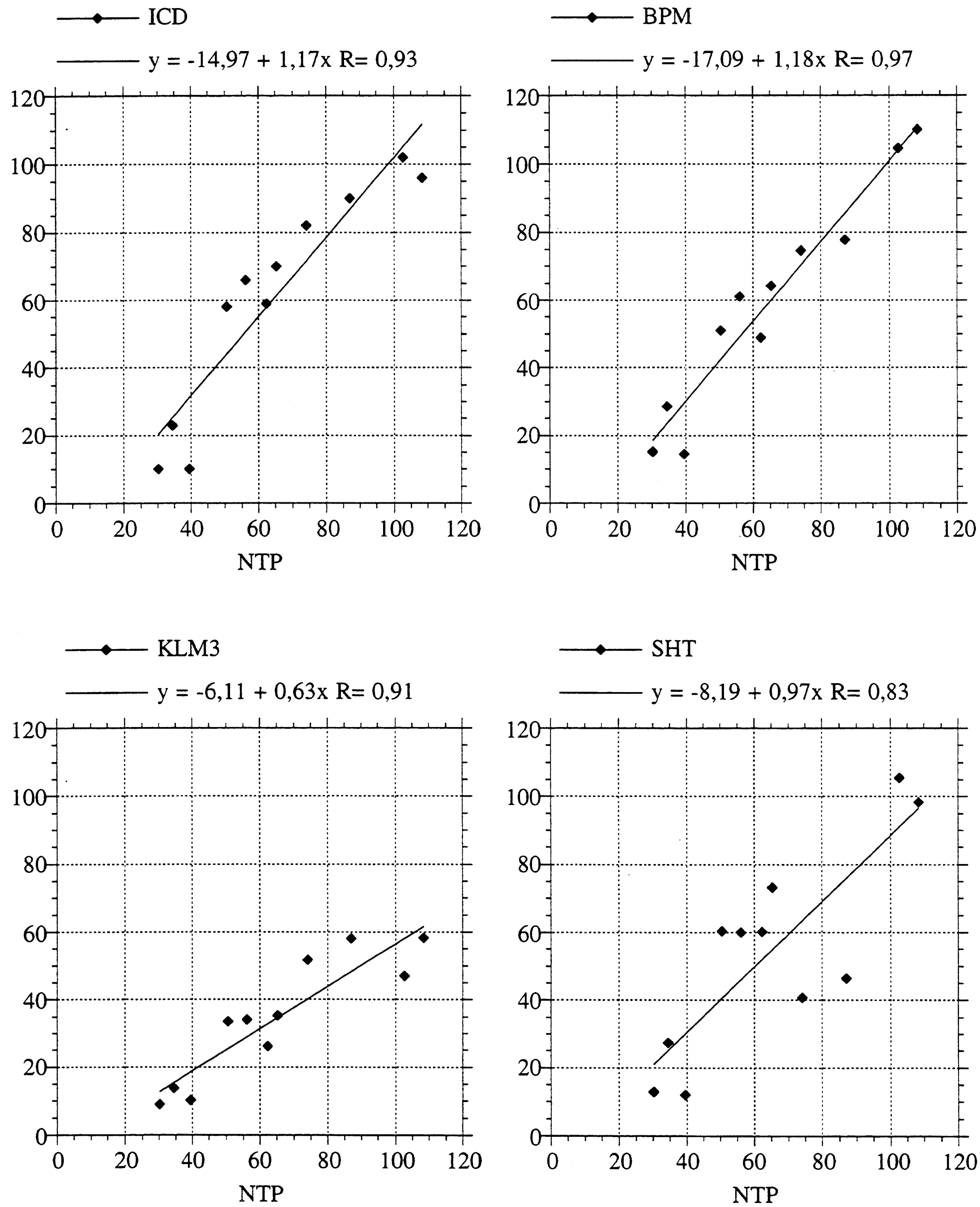


Figure 12: *Correlation and regression line of the normalized tapping positions NTP and the measured syllable-nucleus-onset ICD as well as the model-derived estimations of the alignment position of the word stimuli for the models of Pompino-Marschall (BPM), Köhlmann (KLM3) and Schütte (SHT).*

(p = .0008) the least but still a good correlation to the tappings. The correlation between the normalized tapping positions and the syllable-nucleus-onset is at .932 quite high, but not as good as the correlation of the original p-center model by

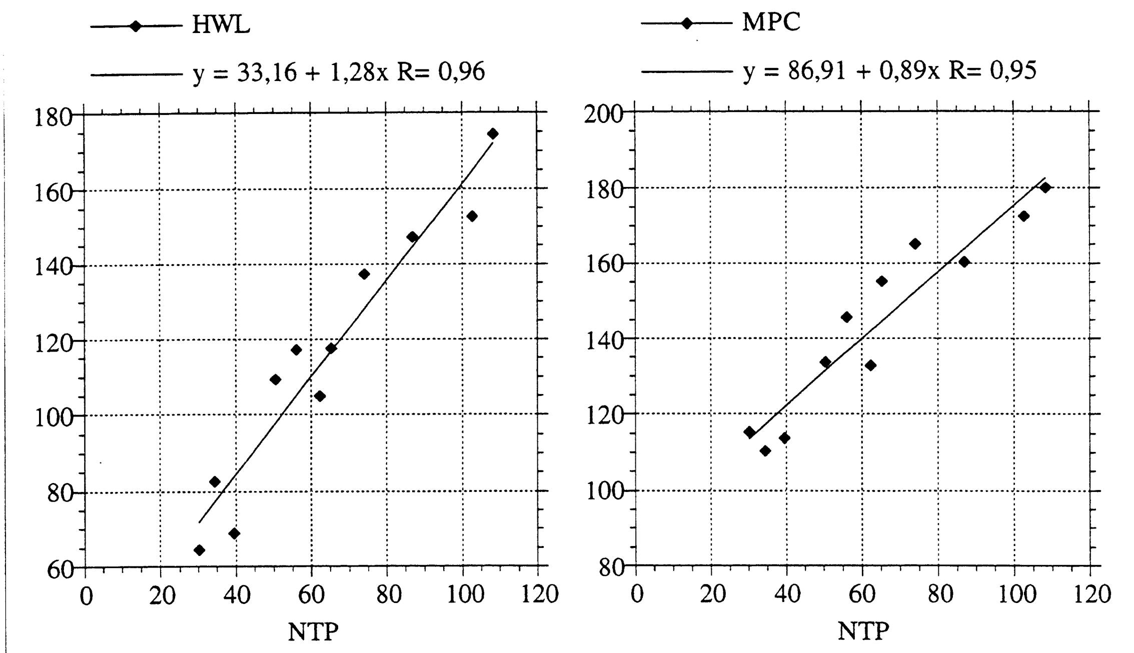


Figure 13: *Correlation and regression line of the normalized tapping positions NTP and the model-derived estimations of the alignment position of the word stimuli for the models of Howell (HWL) and Marcus (MPC) (note the difference in the y-axes values).*

Marcus, which weights the duration of the initial consonance and the rhyme. This is somewhat surprising since a correlation of the rhyme with the tapping positions (Fisher's r to z cor = .446, n.s.) could not be found.
The psycho-acoustic model of Köhlmann uses the pitch and the loudness contours of the acoustic signal to detect onset events, which are integrated to form an 'Ereigniszeitpunkt' (EZP, moment of occurrence) if closer in time to one another than 120 ms. This sometimes leads to more than one event (EZP) per stimulus. For the correlation analysis here – as this is done to assess the predictive ability of the p-center syllable-nucleus-onset hypothesis, not to discuss the problems of the Köhlmann model – if in doubt the events which give the better correlation are used. At .907 the correlation is not as good as the correlation of the nucleus onset.
The psycho-acoustic model of Pompino-Marschall gives at .969 the best correlation. This model calculates partial onset and offset events for rising and falling flanks of the loudness contour of the single critical bands and integrates them to a single 'syllable onset'.
At .963 the p-center estimates of Howells model – which can be interpreted for monosyllabic signals as the area bisect of the rectified signal amplitude envelope – also correlate extremely well.
Unfortunately, a high correlation as such is not sufficient to assess the predictive strength of a model and the reliability of the model estimates.
As can be seen in figure 14, which depicts the model estimates of the four best correlating models (table II) in relation to the categorical segments of the stimuli, most of the estimates do monitor the tendency of the differences of the normal-

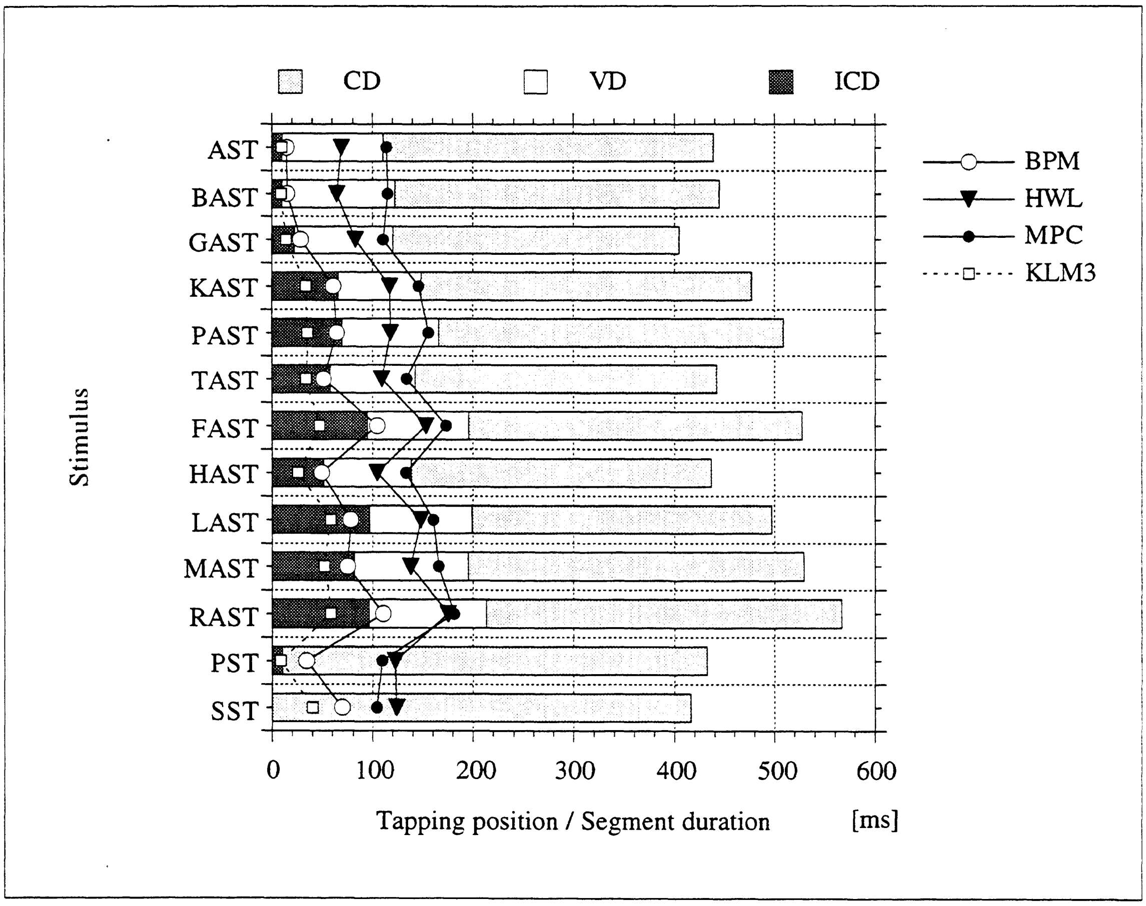


Figure 14: *The estimated alignment positions of all stimuli for the models of Pompino-Marschall (BPM), Howell (HWL), Marcus (MPC) and Köhlmann (KLM3) in relation to the segments of the stimuli (duration of ICD = initial consonance (syllable-nucleus-onset), VD = nucleus, CD = coda).*

ized tapping position quite well but only the BPM estimates are for all stimuli in the magnitude of the normalized tapping positions and hence close to their actual location. For some of the stimuli also the SHT and the KLM models give estimates in the magnitude of the tappings. The magnitude of the MPC estimates can be distinctly improved with respect to the place of the alignment, if one chooses an appropriate constant factor. The problem here simply is, that one needs to know the alignment position beforehand to choose the 'optimal' one (for this data cons. = 76.56). Marcus does not provide any, as the model is not intended to give absolute[1] p-center estimates, and the one I have found for other stimuli elsewhere (cons. = 17.02) wouldn't have helped with the data reported here.
Furthermore, as one might have already noticed and can be seen in figure 15, the correlation matrix (table II) also reveals that all of the models correlate at least as good as or even more highly with the syllable-nucleus-onset than with the normalized tapping positions. In this respect it seems, that the different models used to estimate the alignment position are just some very sophisticated methods to estimate the variation and trend of the syllable-nucleus-onset.

---

1) Marcus [11] states that the constant "is an arbitrary constant representing the fact that we are only determining *relative* (emphasis by Marcus) P-center locations of stimuli to one another." (p. 252f).

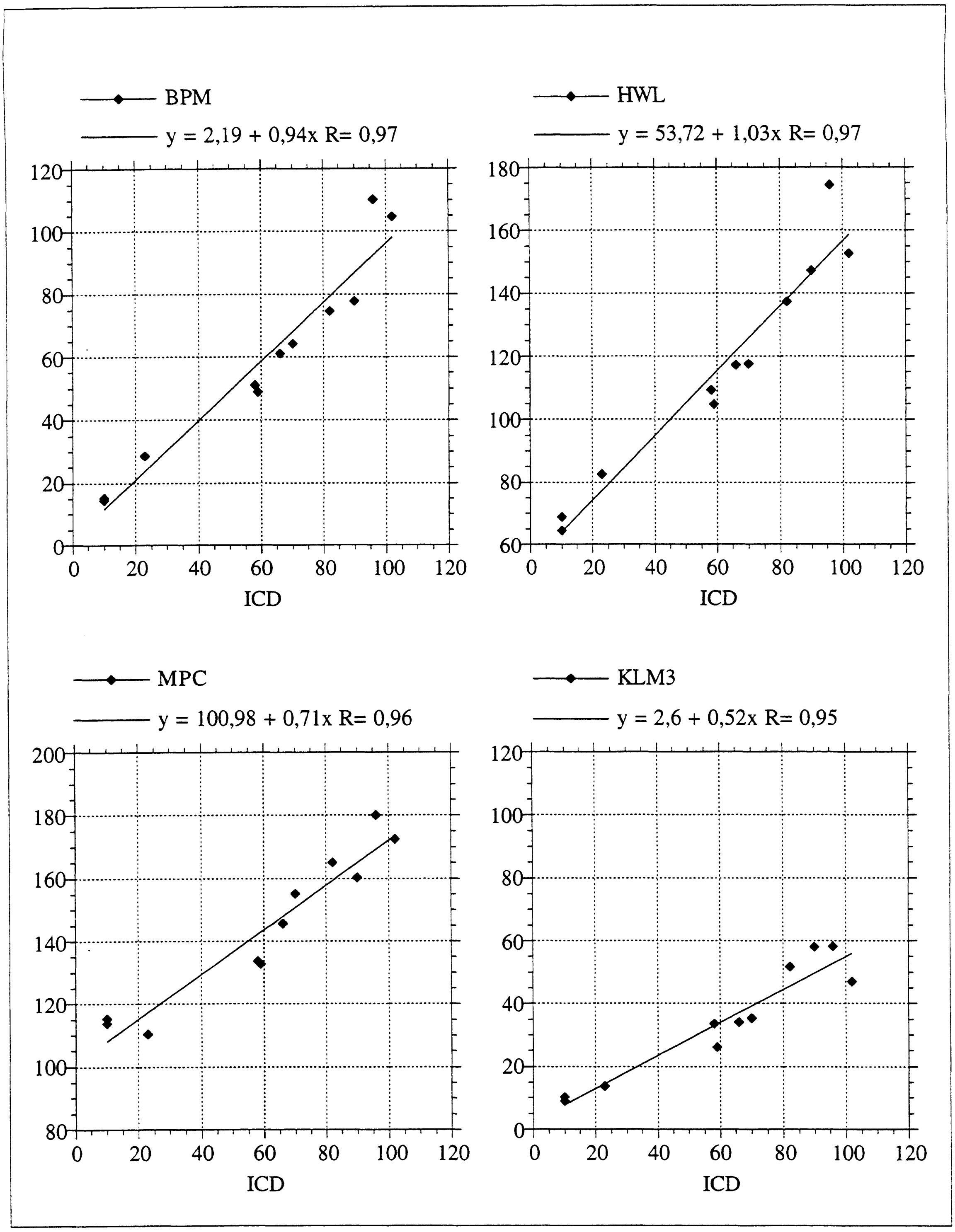


Figure 15: *Correlation of the measured syllable-nucleus-onset ICD and the model-derived estimations of the alignment position of the word stimuli for the models of Pompino-Marschall (BPM), Howell (HWL) Marcus (MPC) and Köhlmann (KLM3).*

On the magnitude of estimates

This result suggests that there is good evidence for the hypothesis of a p-center syllable-nucleus-onset correspondence but a complete match could not be found. The question therefore is, what might be the cause for this mismatch here and the mismatches found elsewhere.

• Firstly the anticipation phenomenon as described above.

The fact, that in most of the other investigations the location of the alignment was somewhat before the consonance-vowel transition was already mentioned above, and might be at least partly due to this phenomenon for which the data here has been corrected. As the method used here is quite straight forward but imprecise, there might be still some inaccuracy due to anticipation, but in my opinion not a substantial one.

• Secondly, and more important the measurements for the segment boundaries have a large influence on the match.

Some variability and uncertainty is introduced by the fact that, depending on the rules for segmentation in use, and the clarity of the signal, a difference of one or two glottal pulses in the location of the segment boarder in both directions can easily occur. Depending on the fundamental frequency the speaker is speaking with, that is, the period duration (70 Hz $\approx$ 14 ms/100 Hz $\approx$ 10 ms) of the signal, this can account for up to 50 ms difference for the measured onset in the data. Furthermore, whether this measured postulated signal syllable-nucleus-onset really represents the mental representation of the syllable-nucleus-onset with sufficient accuracy is also not completely clear. Unfortunately this kind of variability and uncertainty cannot be averted.

• Thirdly, the ability to perform the intended task successfully.

As already mentioned above in experiments with simultaneous tapping to a presented signal, subjects have to perform several tasks at once: recognizing the rhythm pattern, tapping this pattern by estimating the occurrence of the next events in time, judging whether or not they matched the presented event with the tapping action. Hence, quite large difference in subject responses can be expected. The consideration of misleading measurements caused by the first two aspects can partly be avoided as the inability of recognizing the rhythm pattern, estimating the next event or coordinating the tapping action manifests itself for the subjects concerned in a noticeable larger variability of the responses not between stimuli but within the respective stimulus.

The third aspect, judging the co-occurrence, is limited by the fusion and the order threshold, the ability to recognize differences between two successive events (such as Δrstp in figure 6 case A) and the order in which they occur. For auditory signals as such this order threshold (Pöppel [27], Steinbüchel/Pöppel [28]) is supposed to be at about 20 ms to 35 ms (fusion threshold 2 ms), depending on the measurement procedure (recently discussed in detail by Steffen [29]).

Allen [15] found in his data, where a click signal was superimposed on a speech signal, that a range of about 200 ms seems to be an interval within which his subjects still judged events as being simultaneous. Compared to the estimates of the models and hypotheses discussed here – all of them would suggest alignment relations within this distance from the location of the tapping positions and the consonance-vowel transitions – these findings are rather vague and can not be a basis to decide whether any of the models and hypotheses is more successful than the other in estimating the alignment relation.

Therefore, the order threshold of about 20 ms was used as a basis to define an uncertainty or indifference interval (II) of 40 ms within which the order of two events, i.e. tapping and presented syllable, cannot doubtlessly be told. The number of taps within this interval around the estimates of the various models was counted and fed into a contingency analysis (see table III). The observed fre-

Table III: Model specific indifference interval contributions

Contingency Analysis with observed frequencies,
expected values and percents of column totals for the indifference intervals of
the syllable-nucleus onset ICDII and the estimates of the models of Marcus[1] MPCII,
Pompino-Marschall BPMII, Howell HWLII, Schütte SHTII, Köhlmann KLMII
(Chi Square = 7315.171, p < .0001)

		ICDII	BPMII	KLMII	MPCII[1]	HWLII	SHTII	Totals
observed frequencies	outside	7425	7669	8931	12161	10798	7719	54703
	inside[2]	5335	5091	3829	599	1962	5041	21857
	totals	12760	12760	12760	12760	12760	12760	76560
Expected Values	outside	9117.167	9117.167	9117.167	9117.167	9117.167	9117.167	54703
	inside	3642.833	3642.833	3642.833	3642.833	3642.833	3642.833	21857
	totals	12760	12760	12760	12760	12760	12760	76560
Column Totals (%)	outside	58.190	60.102	69.992	95.306	84.624	60.494	71.451
	inside[3]	41.810	39.898	30.008	4.694[1]	15.376	39.506	28.549
	totals	100	100	100	100	100	100	100

1) Can be distinctly improved by post hoc optimization of the so-called 'arbitrary' constant (see text above).

2) The maximum value for this row is the number of tappings inside the indifference interval NTPII of the normalized tapping positions NTP which is 5666.

3) The maximum value for this row is the %-inside value for the indifference interval NTPII of the actual data which is 44.4 %.

quency distribution confirms the mentioned differences in magnitude of the estimations, as the two models (HWL, MPC) with estimations further away from the tapping locations show low to very low values for tappings inside the indifference intervals around their estimates with only about 15.4 % of the tappings inside the indifference interval HWLII for the HWL model and hardly any, that is about 4.7 % inside MPCII for the original p-center model of Marcus. For KLMII about 30 % are reported, which is slightly more than the expected value of 28.5 % under the assumption that there is no difference between the models and the overall distribution between values inside and outside the indifference interval is as given by the entire models. Assuming the indifference interval at the mean of a normal distribution with the standard deviation of the actual tapping data, the expected percentage of values within the indifference interval would be 40 %. This is about the value the analysis offers for the models of Schütte and Pompino-Marschall. Remembering the findings for the correlation analysis, of the computational models the psycho-acoustic model of Pompino-Marschall clearly is the model of choice to determine an alignment position, giving an estimate in mag-

nitude close to the actual tapping locations for words as well as for interjections and showing a very good correlation with the variation of the alignment position introduced by the stimuli.
The only estimate that gives a noticeable higher amount of tappings inside the indifference interval is the syllable-nucleus-onset at about 41.8 %. That is just about 2% below the experiment specific maximal value of 44.4 % and nearly 2% more than could have been expected from a normal distribution with this standard deviation, which indicates that the data is slightly squeezed into the middle of the distribution (positive kurtosis) and that the syllable-nucleus-onset is the alignment marker which accounts best for that fact.
Thus for the words presented here, the estimations derived from the measured syllable-nucleus-onset give the overall best prediction of the alignment position.

CONCLUSION

The results clearly show that the measured syllable-nucleus-onset is at least an equally good estimate for the location of the alignment position (here tapping) for equally timed rhythm perception as any of the other mentioned models of estimation, and therefore supports strongly the notion of a p-center syllable-nucleus-onset correspondence hypothesis or better, a congruence of the p-center and the mental representation of the syllable-nucleus-onset.

ACKNOWLEDGEMENTS

This work was supported by a grant of the "Forschungsschwerpunkt Allgemeine Sprachwissenschaft, Typologie und Universalienforschung der Förderungsgesellschaft Wissenschaftliche Neuvorhaben mbH Berlin".

REFERENCES

[1] TERHARDT, E.; SCHÜTTE, H. (1976), Akustische Rhythmuswahrnehmung: Subjektive Gleichmäßigkeit. *Acustica* 35, pp. 122-126.

[2] MORTON, J., MARCUS, S.M., & FRANKISH. C. R. (1976), Perceptual centres (P-centers), *Psychological Review,* vol. 83, pp. 405-408.

[3] MARCUS, S. M. (1976), *Perceptual Centres.* unpublished PhD thesis, Cambridge University Cambridge, England.

[4] HOWELL, P. (1988), Prediction of p-center location from the distribution of energy in the amplitude envelope: Part I & II, *Perception & Psychophysics,* vol. 43, pp. 90-93 & 99.

[5] JANKER, P. M. (1989), Der Einfluß von Segmentdauer- und Amplitudenmanipulation auf die P-center-Position einfacher CV-Silben, *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München,* vol. 27, pp. 71-141.

[6] JANKER, P. M. (1995), Sprechrhythmus, Silbe, Ereignis. Eine experimentalphonetische Untersuchung zu den psychoakustisch relevanten Parametern zur rhythmischen Gliederung sprechsprachlicher Äußerungen, *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München*, vol. 33, pp. 1-260.

[7] JANKER, P. M. (1995), On the influence of the internal structure of a syllable on the p-center perception, *Proc. 13th ICPhS Stockholm*, vol. 2, pp. 510-513.

[8] JANKER, P. M., POMPINO-MARSCHALL, B. (1991), Is the p-center influenced by 'tone'?, *Proc. 12th ICPhS Aix-en-Provence*, vol. 3, pp. 290-293.

[9] JANKER, P. M., The range of subjective simultaneousness in tapping experiments with speech stimuli, *Proc. of the ESCA Workshop on the Auditory Basis of Speech Perception*, Keele (UK) 15.-19. Juli 1996, pp. 204-207.

[10] KÖHLMANN, M. (1984), Rhythmische Segmentierung von Sprach- und Musiksignalen und ihre Nachbildung mit einem Funktionsschema, *Acoustica*, vol. 56, pp. 192-204.

[11] MARCUS, S. M.(1981), Acoustic determinants of perceptual center (P-center) location, *Perception & Psychophysics*, vol. 30(3), pp. 247-256.

[12] POMPINO-MARSCHALL, B. (1989), On the psychoacoustic nature of the P-center phenomenon, *Journal of Phonetics*, vol. 17, pp. 175-192.

[13] POMPINO-MARSCHALL, B. (1990), *Die Silbenprosodie. Ein elementarer Aspekt der Wahrnehmung von Sprechrhythmus und Sprechtempo*, Tübingen: Niemeyer.

[14] SCHÜTTE, H. (1978), Ein Funktionsschema für die Wahrnehmung eines gleichmäßigen Rhythmus in Schallimpulsfolgen, *Biological Cybernetics* 29, pp. 49-55.

[15] ALLEN, G. D. (1972), The location of rhythmic stress beats in English: An experimental study I & II. *Language and Speech* 15, pp. 72-100 & 179-195.

[16] RAPP, K. (1971) A study of syllabic timing, *Speech Transmission Lab., Royal Inst. Technology Quarterly Status and Progress Report 1*, pp. 14-19 Stockholm.

[17] HOLLISTER, R. D. T. (1973), Relation between hand and voice impulse movements, *Speech Monogr.* 4 (1), pp. 75-100.

[18] MEYER, E. A. (1898), Beiträge zur deutschen Metrik. Über den Takt, *Die neueren Sprachen*, vol. 6, pp. 1-37 & 122-140.

[19] JANKER, P. M. (1994), Nature and properties influencing the p-center perception. RPP94, Sheffield 8.-11. September 1994.

[20] JANKER, P. M. (1996), More on the nature and properties influencing the p-center perception. RPP96, Ohlstadt 8.-11. September 1996.

[21] JANKER, P. M. (1996, to appear), Some evidence for the p-center syllable-nucleus-onset correspondence hypothesis.

[22] DUNLAP, K. (1910), Reactions on rhythmic stimuli, with attempt to synchronize, *Psychological Review*, vol. 17, pp. 399-416.

[23] RADIL, T., MATES, J., ILLMBERGER, J., PÖPPEL, E. (1990), Stimulus anticipation in following rhythmic acoustical patterns by tapping, *Experimentia*, vol. 46, pp. 762-763.

[24] ASCHERSLEBEN, G., PRINZ, W. (1992), What gets synchronized with what in sensorimotor synchronisation, *Paper 13/1992, MPI für psychologische Forschung*, München.

[25] GEHRKE, J. (1996), Presentation at the RPP96, Ohlstadt 8.-11. Sep. 1996.

[26] ASCHERSLEBEN, G. (1994), *Afferente Information und die Synchronisation von Ereignissen*, Frankfurt/M: Lang.

[27] PÖPPEL, E. (1978), Time perception. Held, R., Leibowitz, H., Teuber, H.-L.: *Handbook of sensory physiology, vol. VIII: Perception*, pp. 713-729, Berlin: Springer.

[28] STEINBÜCHEL, N.v., PÖPPEL, E. (1991), Temporal order threshold and language perception. Bhathkar, V., Rege, K.: *Frontiers in knowledge-based computing*. New Delhi: Narosa.

[29]STEFFEN, A. (1995), *Zur Bestimmung der Ordnungsschwelle: Ein experimenteller Vergleich,* Magisterarbeit LMU München.

[30]MATES, J. (1992) Timing and corrective mechanisms in synchronization of motor acts with a sequence of stimuli, *Proc. Fourth Rhythm Workshop: Rhythm Perception and Production,* pp. 43-48 Bourges.

SUPPLEMENT

The 10 stimuli of the experiment not presented in the stimulus section.

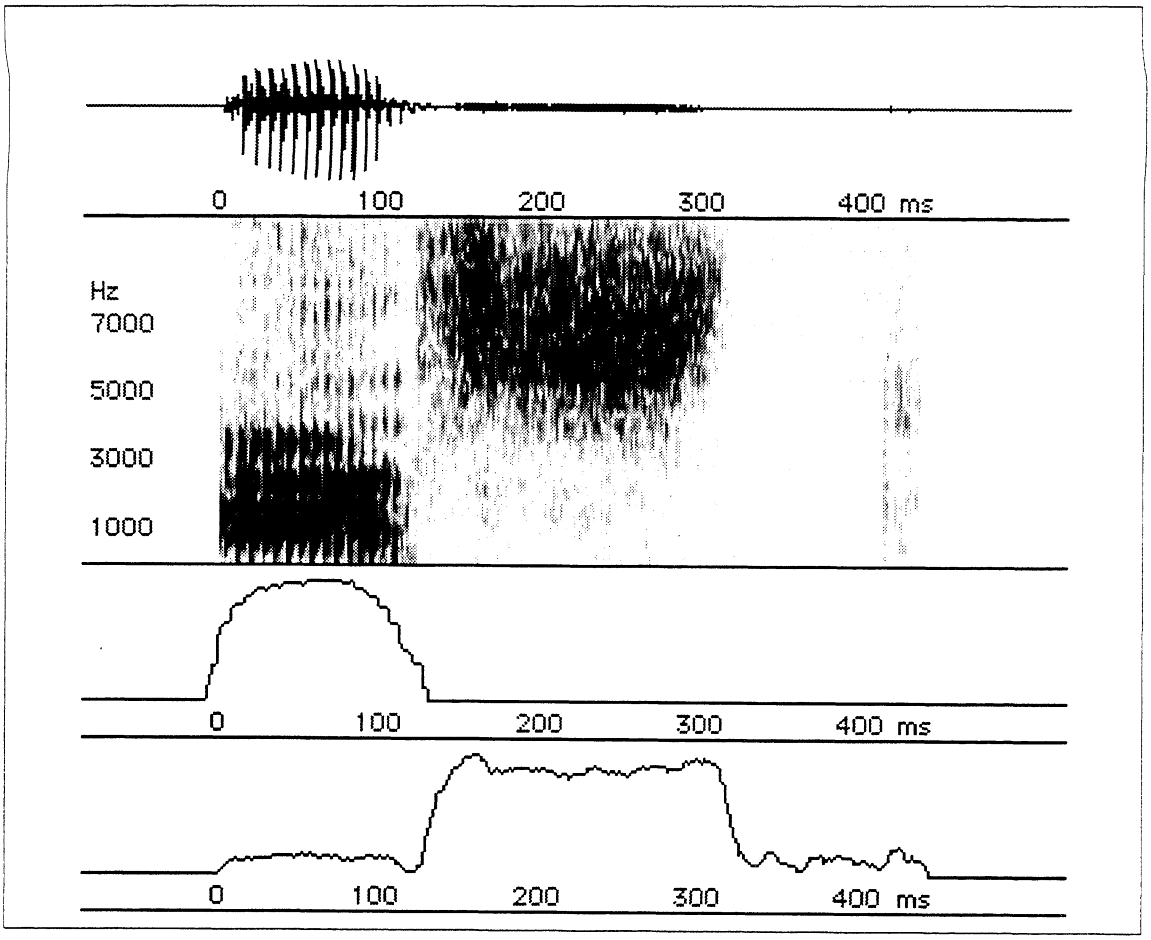


Figure 16: *<ast> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <ast> with an overall duration of 439 ms.*

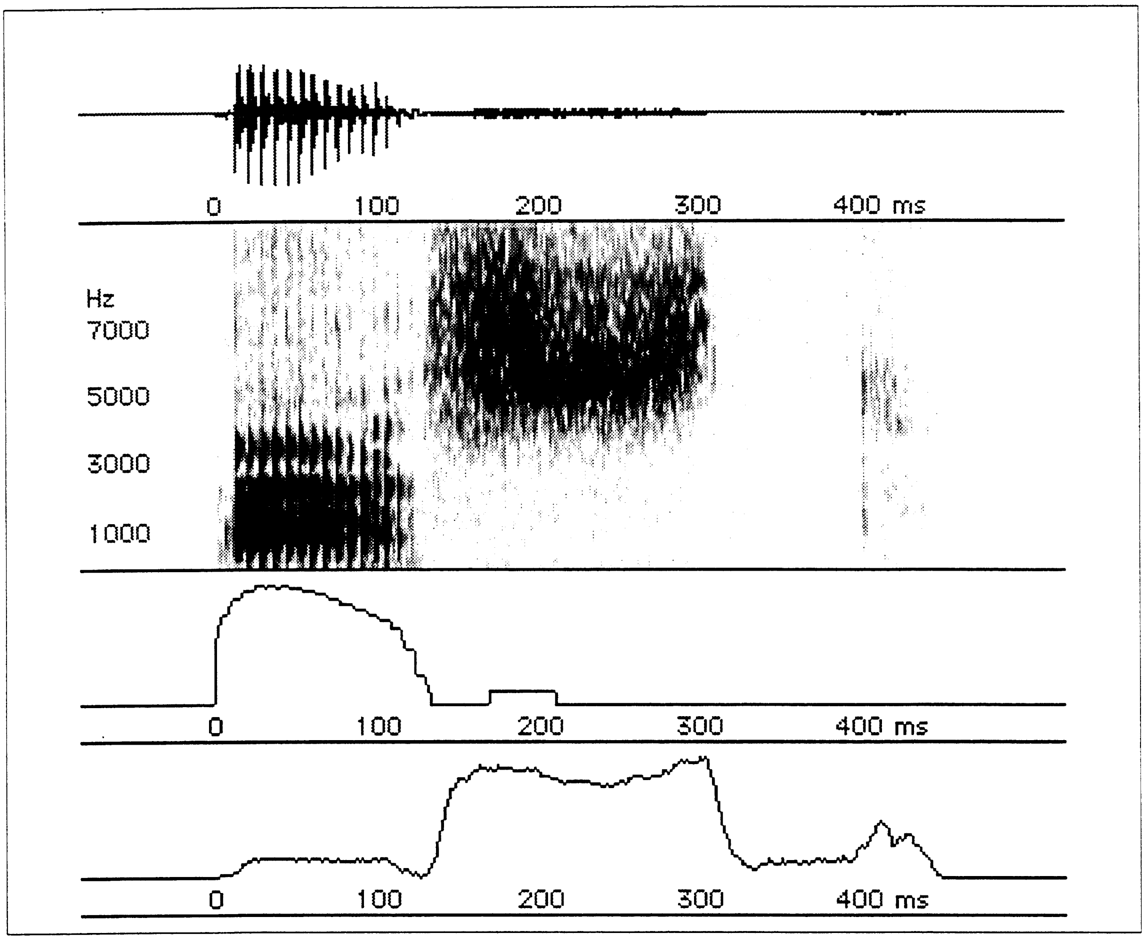


Figure 17: *<bast> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <bast> with an overall duration of 445 ms.*

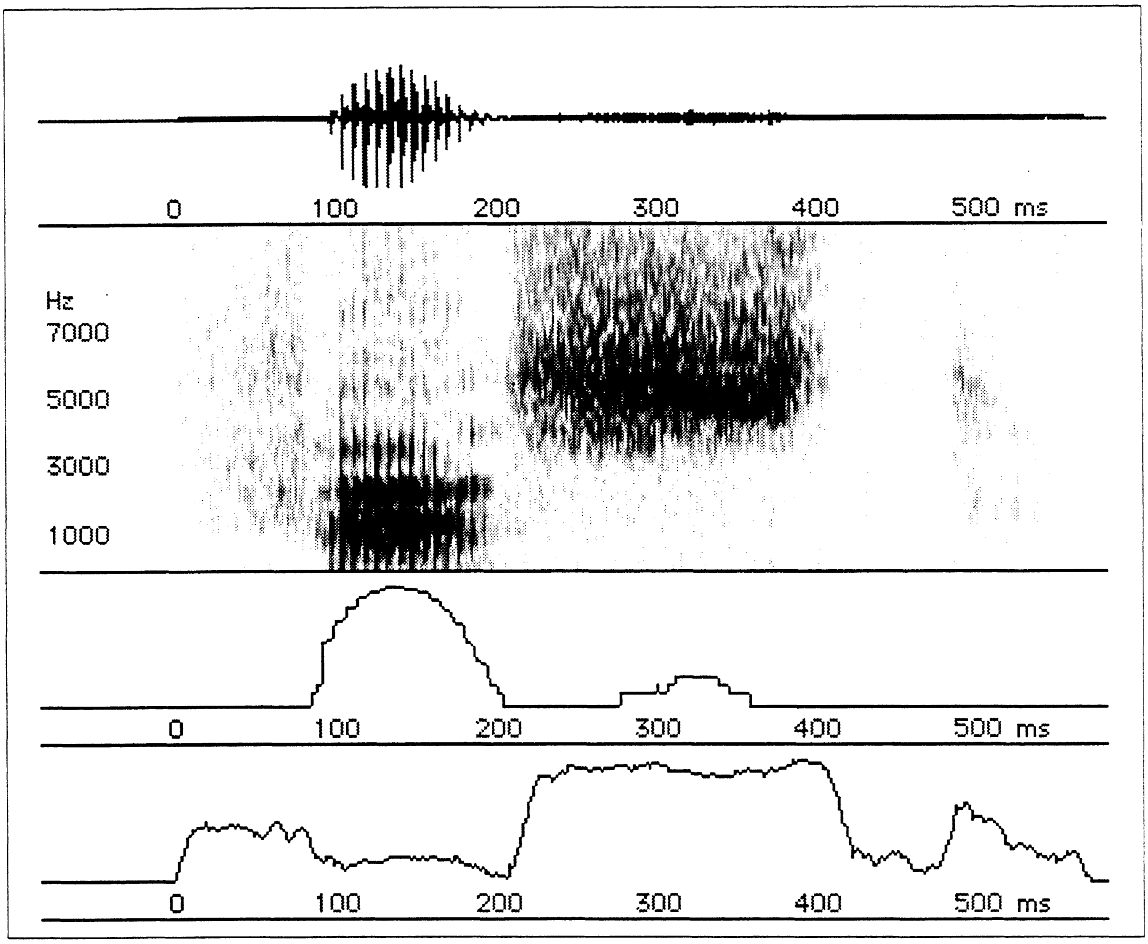


Figure 18: *<fast> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <fast> with an overall duration of 527 ms.*

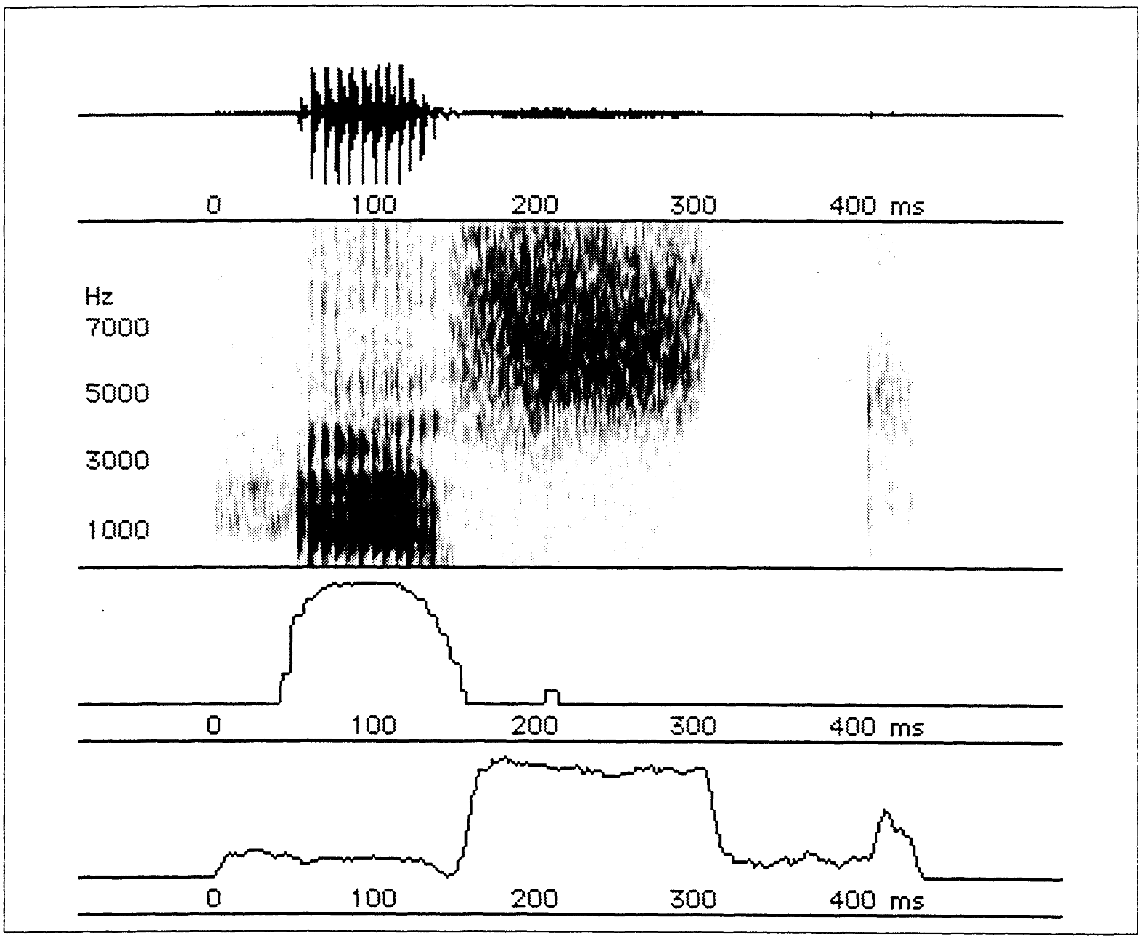


Figure 19: *<hast> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <hast> with an overall duration of 437 ms.*

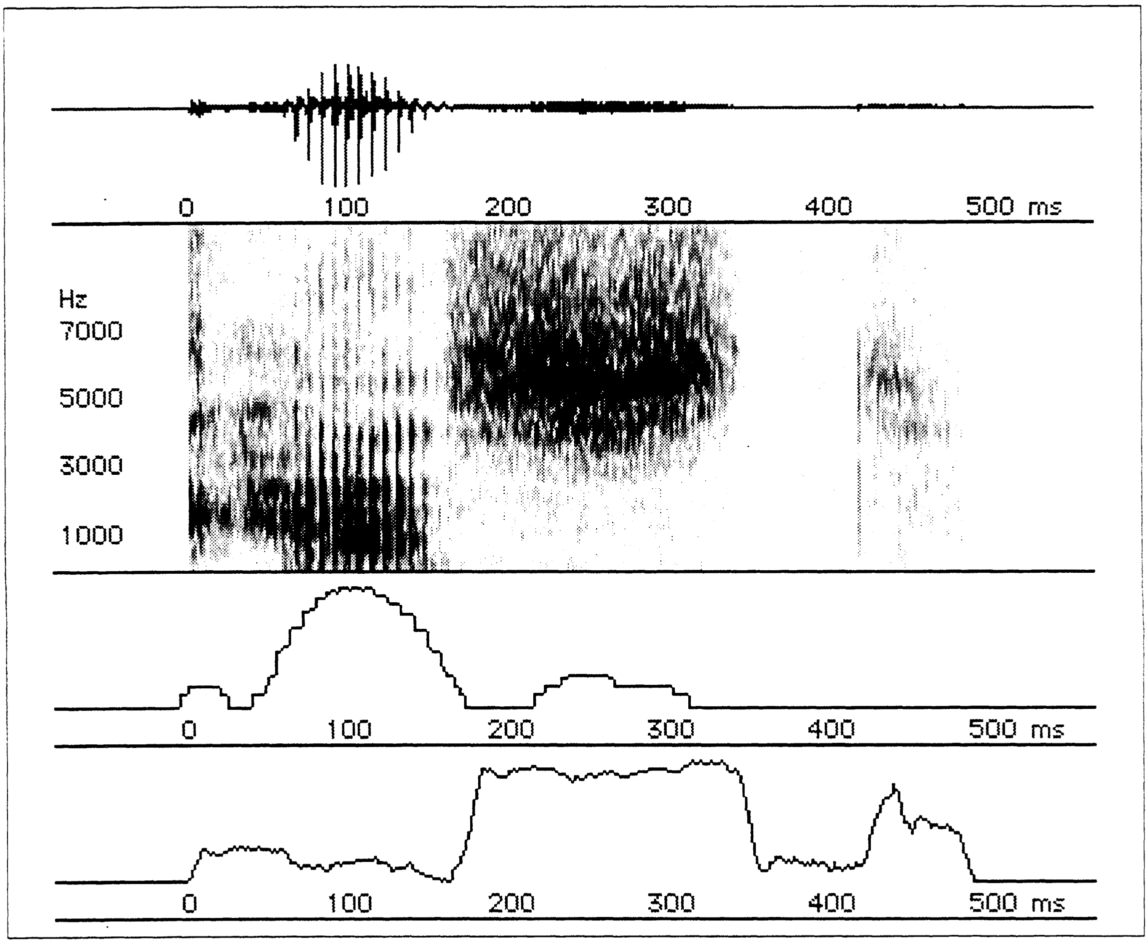


Figure 20: *<kast> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <kast> with an overall duration of 477 ms.*

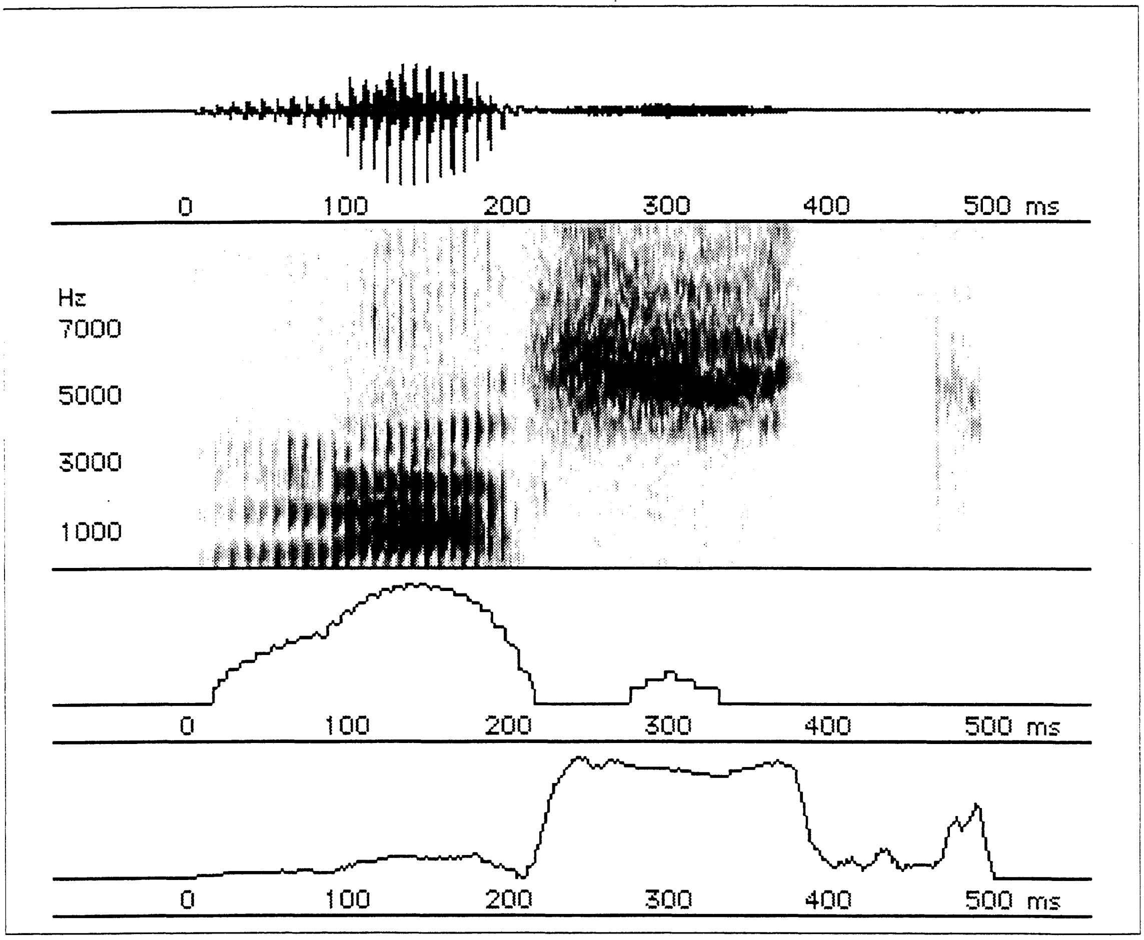


Figure 21: *<last> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <last> with an overall duration of 497 ms.*

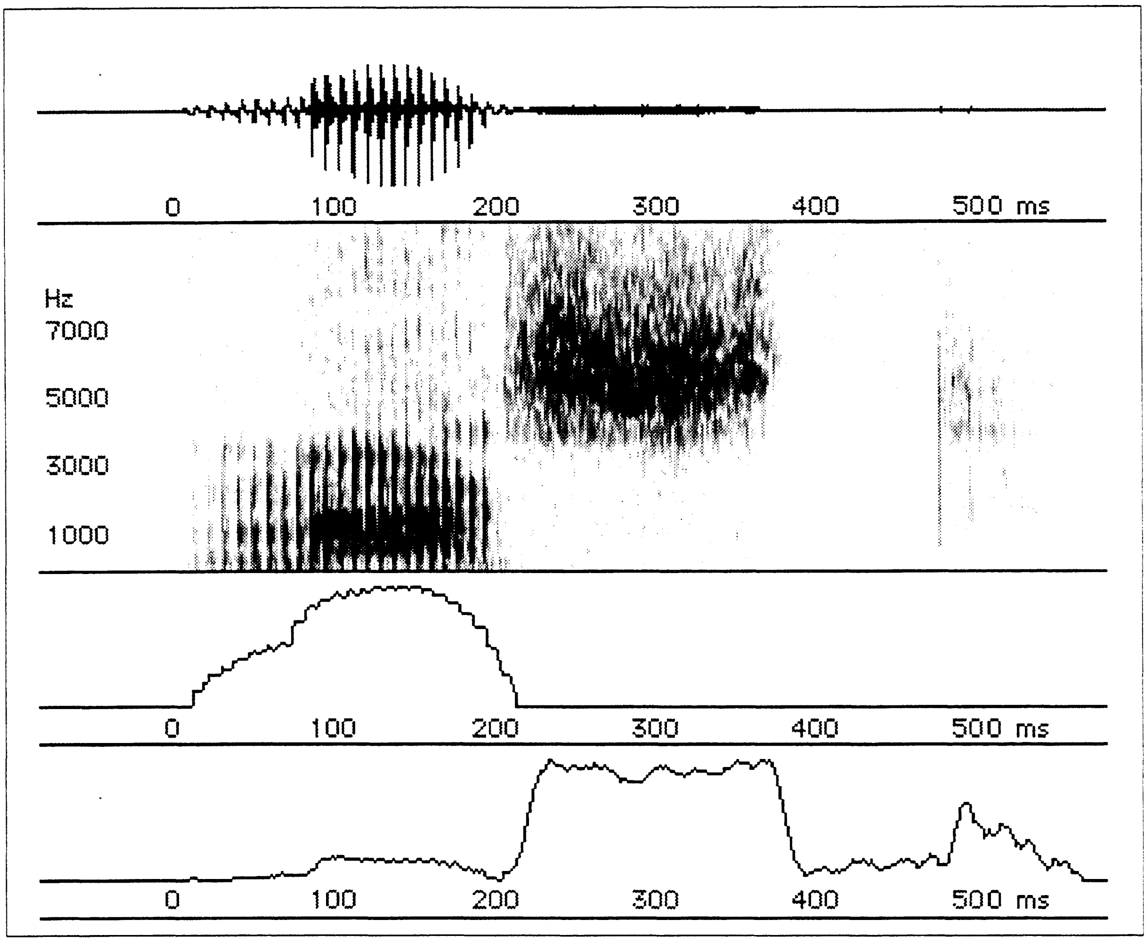


Figure 22: *<mast> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <mast> with an overall duration of 529 ms.*

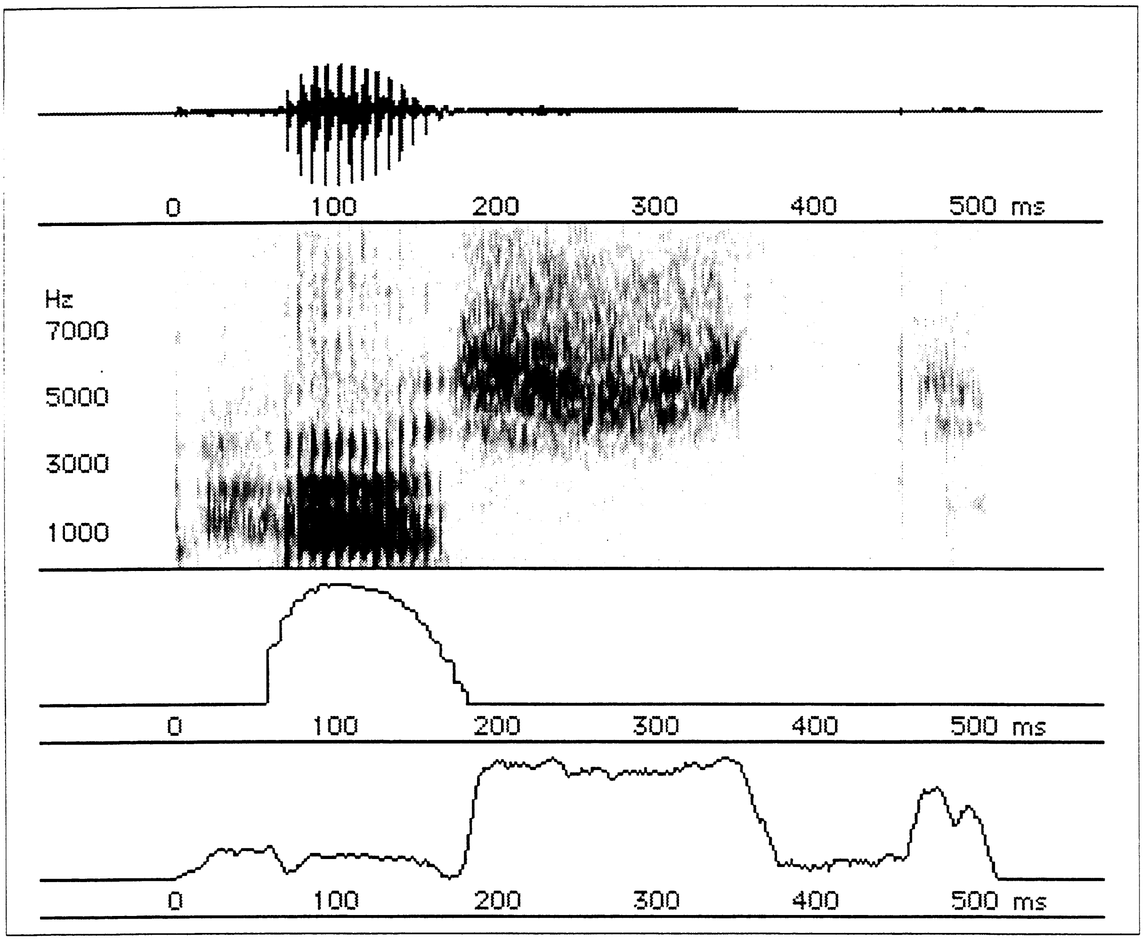


Figure 23: *<past> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <past> with an overall duration of 509 ms.*

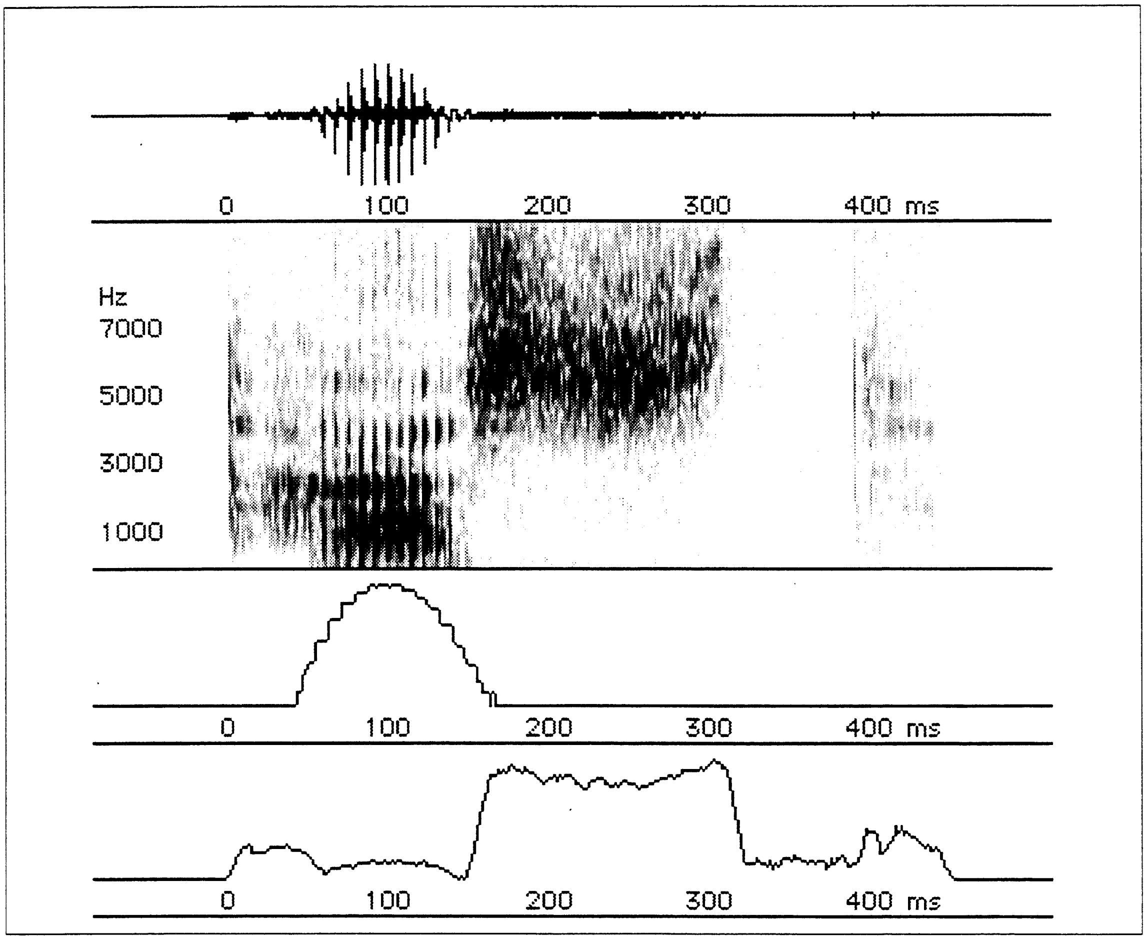


Figure 24: *<tast> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <tast> with an overall duration of 442 ms.*

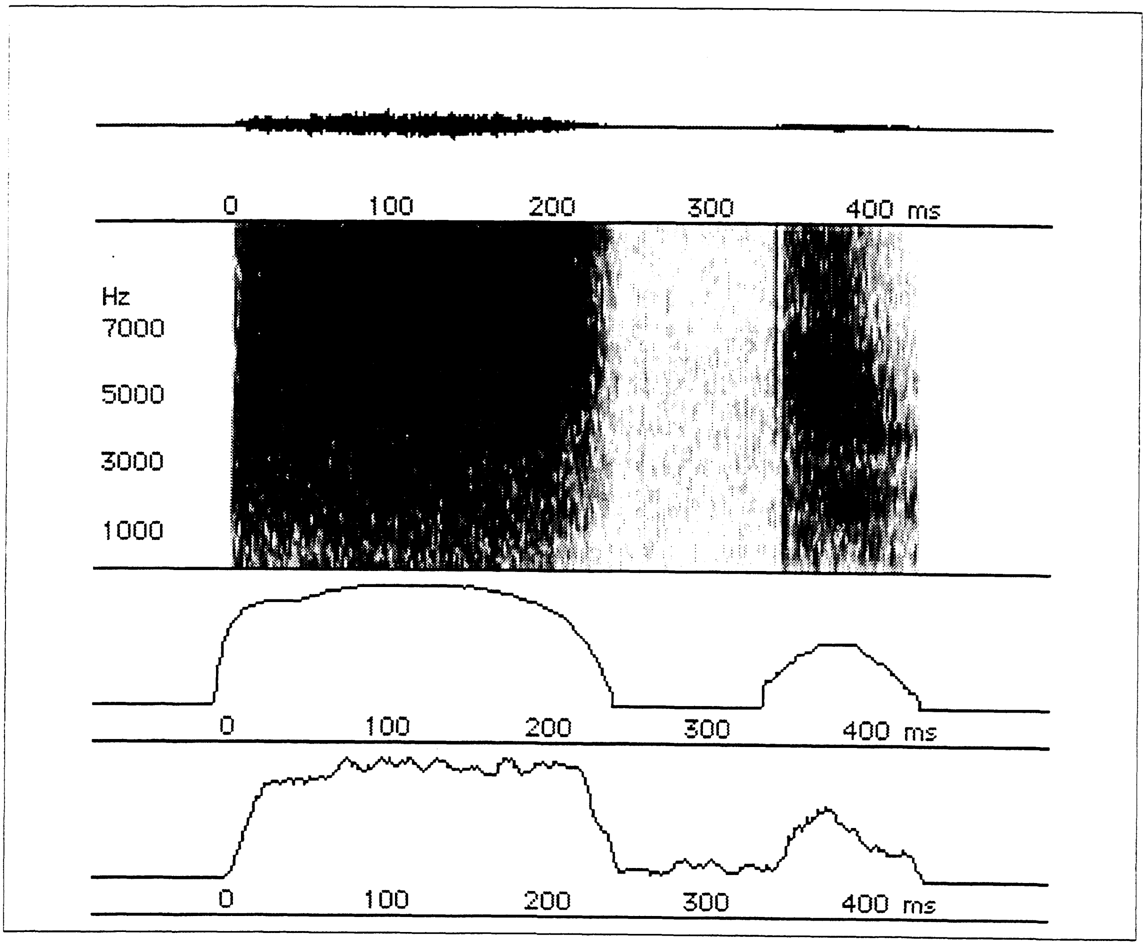


Figure 25: *<pst> Amplitude, sonagram, amplitude envelope (RMS, 30 ms) and zero crossing (10 ms) for stimulus <pst> with an overall duration of 428 ms.*

Manuela Noske

"József Attila" University, Szeged

Alignment and the distribution of German [ç] and [x]*

1 Introduction

This paper examines the distribution of the dorsal fricatives [ç] and [x] in Modern Standard German (MSG) within Optimality Theory (OT) (Prince & Smolensky 1993, McCarthy & Prince 1993a). The velar fricative [x] results from assimilation to a preceding back vowel, while the palatal fricative [ç] occurs elsewhere. Dorsal fricative assimilation does not apply in compounds or to the initial dorsal fricative of the diminutive suffix *-chen* which, though morphologically a suffix, is prosodically a separate phonological word of German (Iverson & Salmons 1992, Wiese 1996). This paper argues that dorsal fricative assimilation is constrained by CRISPEDGE (PrWd) which requires the prosodic word to have sharply-defined boundaries (Itô & Mester in press). I argue that spreading from a word-final back vowel to the initial dorsal of the following word results in a non-crisp word edge and so is ruled out, because CRISPEDGE (PrWd) ranks higher than the constraint governing spreading.

The paper is structured as follows: Section 2 introduces the data and reviews their treatment within the model of lexical phonology (Kiparsky 1982, 1985). Special emphasis is placed on the interaction of the rule of fricative assimilation and umlaut. Section 3.1 provides an account of the distribution of [ç] and [x] within the pre-correspondence version of Optimality Theory (McCarthy & Prince 1995). In that section I show that [x] is optimal after a back vowel and [ç] elsewhere, regardless of what assumptions we make about the underlying representation of the dorsal fricative, thus supporting Itô, Mester & Padgett's (1995) contention that underspecification is not a requirement on input forms in OT. Section 3.2 addresses the non-application of dorsal fricative assimilation in compounds and in diminutive constructions. In that section I argue that dorsal fricative assimilation is constrained by CRISPEDGE (PrWd). Independent support for the relevance of this constraint is provided from syllabification processes of German. Section 4 concludes.

2 Lexical phonology and the distribution of [ç] and [x] in MSG

Examples (1) demonstrate the complementary distribution of the palatal and velar fricative in MSG. The velar fricative [x] occurs only after a back vowel, while the palatal fricative [ç] is found after either a front vowel or a sonorant consonant.[1] These examples illustrate further that the distribution of these segments is independent of their syllabic position. The dorsal fricative can be either in the same syllable as the preceding vowel or form the onset of the next following syllable.

(1) [x] after a non-low back vowel / [ç] elsewhere

Buche	[bu:.xə]	'beech tree'	riechen	[ʀi:.çən]	'to smell'
suchte	[zu:x.tə]	'(s/he) searched'	Nichte	[nɪç.tə]	'niece'
Knochen	[knɔ.xən]	'bone'	sprechen	[ʃpʀɛ.çən]	'to speak'
lochte	[lɔx.tə]	'(s/he) made a hole'	Rechte	[ʀɛç.tə]	'rights'
Sprache	[ʃpʀa:.xə]	'language'	durch	[dʊʀç]	'through'
machte	[max.tə]	'(s/he) made'	Dolche	[dɔl.çə]	'swords'
			manch	[manç]]	'some'
			München	[mʏn.çən]	'Munich'

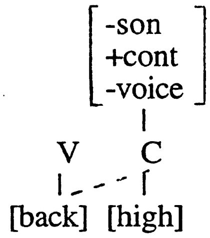
Word-initially, the palatal fricative [ç] is in free variation with either [ʃ] or [k] before front vowels, and with [k] before back vowels (Hall 1992, Wiese 1996):

(2)

Chirurg	[çi:.ʀʊʀk]	~	[ki:.ʀʊʀk]	~	[ʃi:.ʀʊʀk]	'surgeon'
Chemie	[çe.mi:]	~	[ke.mi:]	~	[ʃe.mi:]	'chemistry'
China	[çi:.na]	~	[ki:.na]	~	[ʃi:.na]	'China'
Cholesterin	[ço:.lɛs.te.ʀi:n]	~	[ko:.lɛs.te.ʀi:n]			'cholesterol'
Charisma	[ça.ʀɪs.ma]	~	[ka.ʀɪs.ma]			'charisma'

For several decades the prevailing assumption was that the velar fricative [x] is the less marked and hence basic segment type; more recently [ç] has been considered basic, because of its wider range of distribution. Hall's (1989) approach to this problem is innovative in that he regards neither segment as underlying; instead he proposes that the dorsal fricative is unspecified for backness underlyingly (/X/) and receives a specification for this feature either by spreading in (3) or by the default rule in (4).

(3) Fricative Assimilation (Hall 1989: 3)[2]



(4) [-back] default specification: [-son, +high, +cont] -> [-back]

Some well-known exceptions to the rule of fricative assimilation occur, however. FA does not apply in compounds (5a) or between words at the phrase level (5b).

(5)	a.	Biochemiker	[bi:oçe:mikʌ]	'bio-chemist'
		Indo-China	[ɪndoçi:na]	'Indo-china'
	b.	weil [du:çe:mi] studierst		'because you study chemistry'
		[zo: çi:ne:zɪʃə] Augen		'such chinese eyes'

There is yet another exception to FA which initially does not fall under this last generalization: The initial fricative of the diminutive suffix *-chen* surfaces as palatal and is thus invariant, even if preceded by a back vowel. The diminutive examples in (6a) contrast in this respect with the forms in (6b) in which the dorsal fricative and the preceding vowel are tautomorphemic.

(6)	a.	Kuh+chen	[ku:çən]	'cow, dim.'
		Pfau+chen	[pfao:çən]	'peacock, dim.'
		Tau+chen	[tao:çən]	'rope, dim.'
	b.	Kuchen	[ku:xən]	'cake'
		pfauch+en	[fao:xən]	'to hiss'
		tauch+en	[tao:xən]	'to dive'

Under the model of lexical phonology the exceptionality of this suffix suggests that FA is a lexical rather than a post-lexical rule, since only lexical rules admit of exceptions. If FA were post-lexical, it would be expected to apply "across-the-board", i.e., regardless of the morphological structure of the word. Hall (1989) hence considers FA a lexical rule and assigns it to level II of the lexical phonology of MSG.[3] The assumption that FA is a level II rule is based on its interaction with umlaut. MSG has several suffixes which cause fronting (and in case of a low vowel also raising) of a preceding stem vowel, as shown in (7).[4]

(7)	a.	Bach	[bax]	'brook'
		Bäch+lein	[bɛçlain]	'little brook'
	b.	back+en	[bakən]	'to bake'
		Bäck+er	[bɛkʌ]	'baker'
	c.	schwach	[ʃvax]	'weak'
		schwäch+er	[ʃvɛçʌ]	'weaker'

Hall (1989) argues that umlaut takes place before FA, so that all suffixes that induce umlaut (including those of level II) must be in place before FA has its first chance to apply. If FA were to apply before umlaut (for example at level I), a dorsal fricative would agree with a preceding vowel in backness before umlaut were able to front that vowel.

(8)		[b a X]
	level I	
	FA	[b a x]
	level II	
	<u>morphology:</u>	
	add *-lein*	[[b a x] l a i n]
	<u>phonology:</u>	
	umlaut	[[b ɛ x] l a i n]
		* [bɛxlain] 'brook, dim.'

If FA applied after umlaut (for example late at level II, as suggested by Hall) this problem would not arise.

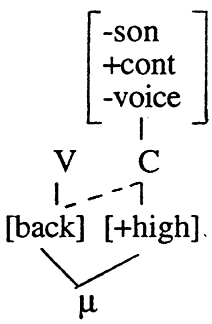
Hall recognizes that there is a problem with his proposal, however: If FA were a late level II rule, it would apply to the dorsal fricative of the diminutive suffix *-chen*. As a stress-neutral derivational affix, *-chen* belongs to level II of the lexical phonology of MSG and so is in principle subject to FA, as shown below.

(9)

level I	[k u:]
level II	
<u>morphology:</u>	
add *-chen*	[[k u:] X ə n]
<u>phonology:</u>	
FA	[[k u:] x ə n]
	* [ku:xən] 'cow, dim.'

To prevent the dorsal fricative from assimilating to a preceding back vowel, Hall stipulates that FA applies only to tautomorphemic sequences of segments, revising his original formulation as in (10).

(10) Fricative Assimilation



The restriction of FA to morpheme-internal environments is, however, at odds with the model of lexical phonology. Lexical phonology incorporates morphological restrictions on the application of phonological rules through level ordering. A lexical rule can be limited in its effect by introducing it with a given level of morphemes and by assuming that it ceases to apply before the next layer of formatives is added to the word. The application of FA to the diminutive suffix should hence be blocked by assigning it to an earlier level than the diminutive suffix *-chen*, and not by placing special restrictions on the rule of FA.

The conditioning of FA to tautomorphemic contexts is indeed unnecessary, since there is no need for umlaut to apply before FA. Umlaut is a typical example of a qualitative rule which applies on the melodic tier, regardless of the syllabic position of the segment it affects, and is thus not subject to Inalterability (Hayes 1986b).[5] That is, umlaut changes a [+back] feature to [-back], regardless of whether that feature is associated with a single vowel or a vowel plus a following dorsal fricative. There is thus no need for umlaut to apply before FA and FA can be treated as a level I rule.

```
(11)                      [ b a X ]
                              |
                           [+back]
     level I
        FA                [ b a x ]
                              |,-'
                           [+back]

     level II
       morphology:
        add -lein       [ [ b a x ] l a i n ]
                              | /
                           [+back]
       phonology:
                        [ [ b ɛ ç ] l a i n ]
                              | /
        umlaut             [-back]

                          [bɛçlain]   'brook, dim.'
```

With FA applying at level I, the dorsal fricative of the diminutive suffix -*chen* would be exempt from assimilating simply by being attached to the stem after FA has ceased to apply. It could thus surface with the default value [-back], even if preceded by a back vowel.

```
(12)                      [ k u: ]
     level I
        FA                    -

     level II
       morphology:
        add -chen       [ [ k u: ] X ə n ]
       phonology:
        default         [ [ k u: ] ç ə n ]

                          [ku:çən]   'cow, dim.'
```

Assigning FA to level I, we can furthermore account for its non-application in compounds. Assuming with Wiese (1986) that compounds are formed at level II, they would fail to undergo FA simply because FA is no longer effective at this level.

While the present analysis is internally coherent and consistent with the level ordering hypothesis of the model of lexical phonology, it is not consistent with other predictions that this model makes. FA introduces the feature [back] onto a segment for which this feature is not contrastive in MSG and so should be considered a post-lexical rule.[6] If FA applied post-lexically, however, we are at a loss as to how to explain the exceptional behavior of the diminutive suffix. Since all morpheme boundaries are erased by the time a representation enters the post-lexical component of grammar, the non-application of FA cannot be accounted for on morphological grounds.

There is, however, an alternative interpretation of these data which brings their phonological and not their morphological structure to bear on the analysis. Bloomfield pointed out in 1930 that forms like *Omachen* [omaçən] 'grandma, dim.' resemble compounds in not undergoing fricative assimilation and so could be analyzed as a sequence of two phonological words: $_{\omega'}$[$_{\omega}$[oma] $_{\omega}$[çən]]. A similar claim has been made in more recent times by Iverson & Salmons (1992), and is also at the heart of Borowsky's (1993) and Wiese's (1996) analysis of this phenomenon.[7] Integrating this insight into the model of lexical phonology, Iverson & Salmons (1992) suggest that FA operates in the post-cyclic lexical component at which structure preservation ceases to be in effect (Booij & Rubach 1984, 1987). Since post-cyclic rules apply only word-internally, the dorsal fricative of the diminutive suffix *-chen* is exempt from assimilation, if it is a "non-cohering" suffix and hence a prosodic word of German.

One fact that is problematic for the assumption that *-chen* is a prosodic word is that it can cause umlaut of a preceding back stem vowel. The noun roots in (13) which display umlaut in the diminutive form contrast in this respect with the examples in (14) in which umlaut fails to apply.

(13)	Schuh	[ʃu:]	'shoe'	Schüh+chen	[ʃy:çən]	'shoe, dim.'
	Haus	[haos]	'house'	Häus+chen	[hɔøsçən]	'house, dim.'
	Sohn	[zo:n]	'son'	Söhn+chen	[zø:nçən]	'son, dim.'

(14)	Mama	[mama]	'mom'	Mama+chen	[mamaçən]	'mom, dim.'
	Auto	[aoto]	'car'	Auto+chen	[aotoçən]	'car, dim.'
	Oma	[oma]	'grandma'	Oma+chen	[omaçən]	'grandma, dim.'

If *-chen* were a "non-cohering" suffix, we would expect it to leave the base of suffixation alone, phonologically speaking. That it can cause umlaut suggests instead that *-chen* forms a single phonological word with the base.

To solve the umlaut problem, Iverson & Salmons (1992) claim that there are two formatives of the shape *-chen* in MSG: An umlaut-inducing one which can potentially undergo FA and which indicates the diminutive only, and a variant with an affective connotation that does not cause umlaut and that fails to undergo FA. The first suffix is "relatively deep in the lexicon [...]," while the second instance of *-chen* is "[...] closer to the surface" (Iverson & Salmons 1992:42). Despite the fact that they assign these two suffixes to different strata of the lexicon of MSG, which invites an analysis in terms of the (cyclic) lexical phonology of this language, Iverson & Salmons ultimately seek an explanation for their different behavior in their phonological properties. They claim that only the first type of suffix is fused with the base to form a single phonological word, while the second type of affix forms a phonological word of its own.

However, the semantic distinction that justifies this classification is not supported by the German data at large. According to Fleischer (1975) diminutive forms like *Städtchen* [ʃtɛtçən] 'town, dim.' *Sümmchen* [zʏmçən] 'sum, dim.', *Mütterchen* [mʏtʌçən] 'mother, dim.' and *Küsschen* [kʏsçən] kiss, dim.' do not only express the diminutive, but often have an affective connotation. The problem that these and similar forms pose for an analysis along the lines of Iverson & Salmons is that they have an affective connotation and show umlaut. There is hence no clear semantic distinction between forms with umlaut which have a pure diminutive meaning and forms without umlaut with an affective connotation. Instead the affective connotation is found with both umlauted and non-umlauted stems, making a distinction between a diminutive, umlauting version of *-chen* and an affective, non-umlauting version of *-chen* untenable.

Assuming, then, that there is only a single suffix of the shape *-chen* in MSG, an analysis in terms of the model of prosodic lexical phonology is faced with a dilemma, the umlaut dilemma. The diminutive suffix forms a single phonological word with respect to umlaut, but a separate phonological word inasfar as dorsal fricative assimilation is concerned. We will see in section 3.2 that the contradictory behavior of the diminutive suffix finds a straightforward explanation if analyzed in terms of Generalized Alignment (McCarthy & Prince 1993b).

3 Optimality theory and the distribution of [ç] and [x] in MSG

3.1 Underspecification and Faithfulness

Optimality Theory (OT) (Prince & Smolensky 1993, McCarthy & Prince 1993a) assumes that a phonological grammar consists of a set of universal constraints which evaluate a representation with respect to its well-formedness. Constraints are ranked on a language-particular basis. They are violable, but violation is minimal; i.e., that candidate that best fits the well-formedness conditions expressed by the constraints will emerge as optimal, even though it might violate one or more of the existing constraints. A key element of this model is the assumption of a generator function Gen; for any given input form Gen produces a set of output candidates by making any number of changes to the input. These candidates are evaluated "in parallel"; i.e., unlike the model of lexical phonology which allows word formation rules and phonological rules to interact cyclically, OT assumes that all morphemes are part of the input.

It is the function of faithfulness constraints to insure that every element of the input is contained in every output candidate. Depending on the ranking of these constraints in the overall constraint hierarchy, either a minimal or maximal deviation of output from input is possible. Faithfulness constraints figure prominently in the following discussion of the underlying representation of dorsal fricatives in MSG and thus merit a brief review. Prince & Smolensky (1993) introduce two faithfulness constraints, PARSE and FILL, which militate against the deletion and insertion of input elements from the output. Itô, Mester & Padgett (1995) extend the use of these constraints to cover both features and association relations between elements in the input form, grouping them under the common name FAITH:

(15) FAITH (Feature Faithfulness) (Itô, Mester & Padgett 1995: 586)

PARSEFEAT

All input features are parsed.

FILLFEAT

All features are part of the input.

PARSELINK

All input association relations are kept.

FILLLINK

All association relations are part of the input.

In their view, not only the insertion of a feature, but also the insertion of an association line into a candidate form constitutes a violation of FAITH.

Let us begin the analysis of dorsal fricative assimilation in MSG by considering what the underlying representation of dorsal fricatives is. The discussion in the previous section assumed that dorsal fricatives are unspecified for backness underlyingly and that they receive a specification either by spreading or by default. Itô, Mester & Padgett (1995) maintain that OT does not require underspecification of input forms, since output constraints and their ranking alone provide an account of the pattern of segmental distribution. Using as an example the feature [voice] in sonorants and sonorant-obstruent clusters in Japanese, they show that, given a set of input forms which differ minimally in the specification of one of their segments and a fixed ranking of the constraints, the same output candidate emerges as optimal. Underspecification is thus not a necessary requirement of input forms.

The same conclusion could be reached regarding the underlying representation of dorsal fricatives in MSG. The following discussion will show that we can account for the distribution of [ç] and [x] without assuming that dorsal fricatives are unspecified for backness (/X/) in the input. As a matter of fact, we can even account for the occurrence of these segments under the assumption that dorsal fricatives are specified as [+back] underlyingly (/x/). For the purposes of this discussion let us distinguish between two kinds of environments, namely those in which the realization of the dorsal fricative is context-dependent and those in which it is not. As the earlier examples in (1) and (2) show, a dorsal fricative surfaces as [-back] in any environment other than after a back vowel. If we want to entertain the hypothesis that dorsal fricatives can be specified as [+back] in the input and account for their realization only in terms of the constraint hierarchy of MSG, we need to assume the existence of a constraint which bans the co-occurrence of the feature values [+back] and [+continuant] in dorsal consonants: *[x]. Assuming that this constraint is ranked higher than the constraint FAITH [αback] which preserves the underlying backness specification of a dorsal fricative, a [+back] dorsal fricative is excluded in a position such as word-initial, for example.

Consider first the form /xemi:/ 'chemistry' in table (16) below, whose initial fricative is specified as [+back] underlyingly. If *[x] is ranked above FAITH [αback], then the faithful candidate (a) is excluded due to its violating the co-occurrence constraint. Of the non-faithful candidates, (c) wins over (b), because it violates FAITH [αback] in only two instances. I will follow Itô, Mester & Padgett's use of rendering inserted features in italics, but use bold type-face to mark inserted association lines. The insertion or deletion of a feature, as well as any addition or deletion of an association line registers for one FAITH [αback] violation.

(16) Input: /x e m i:/ 'chemistry'
|
+bk

candidates	*[x]	FAITH [αback]
a. x e m i: | +bk	*!	
b. ç e m i: | *-bk* <+bk>		****!
c. ☞ X e m i: <+bk>		**

The winning candidate (c) demands some further explanation. Specifically, we need to consider how output forms with an unspecified dorsal are to be interpreted. Itô, Mester & Padgett (1995) argue that output underspecification is possible within the framework of OT. Infact, a key element of their analysis of sonorant voicing in Japanese is that sonorants are specified for [voice] in some contexts, namely if followed by an obstruent, while they are unspecified for this feature in other environments. Since sonorants are phonetically voiced, there must be either a statement or a rule which relates the incompletely specified output to the features of the phonetic signal. We could similarly assume that a dorsal fricative that is unspecified for backness in the output is realized at a palatal place of articulation. As regards their phonetic realization, there are thus no substantive differences between the forms (16b) and (16c) and I will treat both candidates as well-formed output representations corresponding to the phonetic form [çemi:] 'chemistry'.

Table (17) presents the other possible input form /çemi:/ with a [-back] word-initial dorsal. Of the candidates that are associated with this input, (a) is excluded because it violates the highest ranking constraint *[x]. The choice between the remaining two candidates therefore falls onto FAITH [αback]. As expected, the faithful candidate (b) wins over the non-faithful candidate (c).

(17) Input: /ç e m i:/ 'chemistry'
|
-bk

candidates	*[x]	FAITH [αback]
a. x e m i: | *+bk* <-bk>	*!	****
b.☞ ç e m i: | -bk		
c. X e m i: <-bk>		**!

The same result obtains when the dorsal fricative is unspecified for backness in the input, as shown in table (18). While candidate (a) is eliminated by the co-occurrence constraint, the faithful candidate (c) emerges as the winner over the non-faithful candidate (b).

(18) Input: /X e m i:/ 'chemistry'

candidates	*[x]	FAITH [αback]
a. x e m i: | *+bk*	*!	**
b. ç e m i: | *-bk*		**!
c. ☞ X e m i:		

Hence, we can account for the occurrence of [ç] in word-initial position without having to assume that either /ç/ or /X/ underlies this segment. This "freedom of analysis" is not without consequences, however: In order to maintain the assumption that surface [ç] can correspond to a [+back] dorsal fricative /x/ underlyingly, we need to appeal to the constraint *[x] which bans the occurrence of the feature value [+back] in dorsal fricatives. By contrast, if we assumed that either /ç/ or /X/, but not /x/ were basic, there would be no need to invoke this constraint. Instead of limiting the range of possible surface forms by assuming a close featural relationship between input and output, OT allows input forms that can drastically vary from the corresponding output forms, placing the burden of choosing the correct output on the constraint hierarchy. In the absence of an accompanying theory that assesses the cost-effectiveness of these analytical choices, it remains to be seen what the advantage of an OT approach is over an approach that favors minimization in the choice of features at the underlying level.

We can now turn to those examples in which the realization of the dorsal fricative is predictable through context. As already outlined in section 2, a dorsal fricative surfaces as [+back] after a back vowel. Invoking the important insight of autosegmental phonology that assimilation is feature spreading (Hayes 1986a), I posit the constraint in (19) which requires a back vowel and a following dorsal consonant to share a single specification [+back] (Itô & Mester 1995).[8] Since CVLINKAGE requires double linking of the feature value [+back], a VC sequence whose individual segments are specified as [+back] violates this constraint.

(19) CVLINKAGE:

```
              V     C
      root    •     •
              |     |
      dorsal  •     •
              |    /
           [+back]
```

If CVLINKAGE is ranked higher than FAITH [αback], only a candidate in which the feature [+back] is doubly linked emerges as optimal, regardless of what assumptions we make about the underlying representation of the dorsal fricative. Also, since [+back] is a possible feature in dorsal consonants after a back vowel, the constraint CVLINKAGE must rank higher than *[x].

Suppose that the dorsal fricative is specified as [-back] underlyingly, as shown in table (20). Gen submits, as members of the candidate set associated with this input, the output (a) in which the dorsal fricative agrees with the preceding vowel in backness. Despite the fact that it violates *[x] and FAITH [αback], it is the only form that satisfies the higher ranking constraint CVLINKAGE, and so emerges as the winner.

(20) Input: /k u: ç ə n/ 'cake'
| |
+bk -bk

candidates	CVLINKAGE	*[x]	FAITH [αback]
a. ☞ k u: x ə n | / +bk <-bk>		*	***
b. k u: x ə n | | +bk *+bk* <-bk>	*!	*	****
c. k u: ç ə n | | +bk *-bk* <-bk>	*!		****
d. k u: ç ə n | | +bk -bk	*!		
e. k u: X ə n | +bk <-bk>	*!		**

Consider now the table in (21) which evaluates candidates based on the input /ku:xən/ with a [+back] dorsal fricative. All candidates except (a) fall victim to CVLINKAGE and so are non-optimal.

(21) Input: /k u: x ə n/ 'cake'
        | |
      +bk +bk

<table>
<tr><th>candidates</th><th>CVLINKAGE</th><th>*[x]</th><th>FAITH [αback]</th></tr>
<tr><td>a. ☞ k u: x ə n
| /
+bk <+bk></td><td></td><td>*</td><td>***</td></tr>
<tr><td>b. k u: x ə n
| |
+bk +bk <+bk></td><td>*!</td><td>*</td><td>****</td></tr>
<tr><td>c. k u: ç ə n
| |
+bk -bk <+bk></td><td>*!</td><td></td><td>****</td></tr>
<tr><td>d. k u: x ə n
| |
+bk +bk</td><td>*!</td><td>*</td><td></td></tr>
<tr><td>c. k u: X ə n
|
+bk <+bk></td><td>*!</td><td></td><td>**</td></tr>
</table>

The last input to consider, (22), has an unspecified dorsal fricative. Even in this case, candidate (a) emerges as the winner due to CVLINKAGE outranking the co-occurrence and faithfulness constraints.

(22) Input: /k u: X ə n/ 'cake'
        |
      +bk

<table>
<tr><th>candidates</th><th>CVLINKAGE</th><th>*[x]</th><th>FAITH [αback]</th></tr>
<tr><td>a. ☞ k u: x ə n
| /
+bk</td><td></td><td>*</td><td>*</td></tr>
<tr><td>b. k u: x ə n
| |
+bk +bk</td><td>*!</td><td>*</td><td>**</td></tr>
<tr><td>c. k u: ç ə n
| |
+bk -bk</td><td>*!</td><td></td><td>**</td></tr>
</table>

d.	k u: X ə n | +bk	*!		

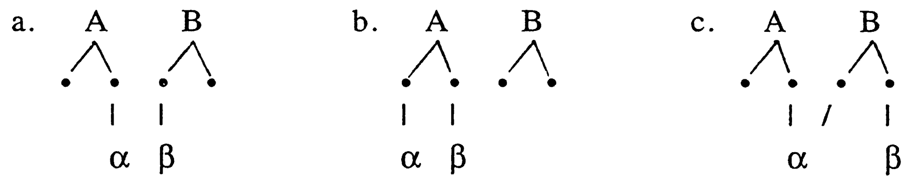
A comparison of tables (20) through (22) shows that we can account for the occurrence of [x] in a position after a back vowel without having to assume that dorsal fricatives are unspecified in the input. Furthermore, we have seen that we can account for the occurrence of [ç] in all other positions on the basis of the ranking *[x] >> FAITH [αback]. In the remainder of this paper I assume that dorsal fricatives are specified as [-back] underlyingly and so will omit the constraint *[x] from the following tables.

3.2 Alignment and Crisp Edges

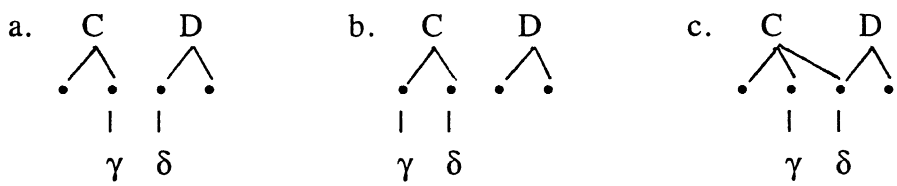
We now turn to the focus of this analysis, namely the non-application of dorsal fricative assimilation in compounds and the diminutive construction.

The notion of alignment developed by McCarthy & Prince (1993b) requires that a specified edge of a morphological constituent (a root, stem or affix) coincide with a specified edge of a prosodic constituent (a syllable, foot or prosodic word). The relationship between these categories is exclusive; i.e., no multiple linking of a feature across a boundary may occur. Itô & Mester (in press) propose a revision of alignment theory which distinguishes alignment proper from the requirement of prosodic constituents to have "crisp edges". In an effort to account for cases in which a consonant, for example, syllabifies into both the coda and the onset across a juncture, they claim that instances of multiple linking satisfy alignment. Consider the structures in (23) below, where A and B represent two morphological constituents, and α and β two prosodic constituents. Assuming that the right edge of A has to coincide with the right edge of α, then on Itô & Mester's view neither the structure in (23a) nor the one in (23c) violates ALIGN-R.

(23)



Alignment differs from the requirement of prosodic constituents to have "crisp edges"; i.e., the condition that every prosodic element is incorporated into a single higher prosodic unit which is captured by a family of CRISPEDGE constraints. CRISPEDGE rules out any linking across the boundary of a prosodic constituent. Assuming that C and D are prosodic categories, then the constraint CRISPEDGE (C) is violated by the structure in (24c), because element δ is linked to both prosodic categories C and D.

(24)


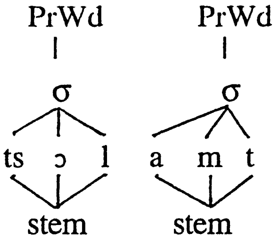
CRISPEDGE (PrWd) plays a crucial role in accounting for the non-application of dorsal fricative assimilation in German compound and diminutive constructions. This constraint is independently motivated in the analysis of syllabification in German.

McCarthy & Prince (1993b) observe that the left stem boundary in German is opaque to syllabification. The examples in (25) demonstrate that the final consonant of a prefix or the first constituent of a compound syllabifies into the coda, rather than into the onset of the following syllable, satisfying ALIGN (Stem, L, PrWd, L). The insertion of a glottal stop before the second vowel-initial morpheme is a consequence of the undominated position of ONSET in German.[9]

(25)	a.	auf+cssen	[ʔauf.ʔɛsn]	'to cat up'
		ver+irren	[fʌ.ʔɪʀn]	'to lose one's way'
	b.	Zoll+amt	[tsɔl.ʔamt]	'customs-house'
		berg+ab	[bɛʀk.ʔap]	'downhill'

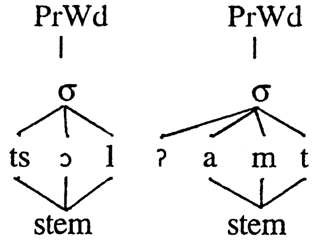
Consider the candidates in (26), all of which are built on the same input /tsɔl+amt/ 'customs-house'.

(26) a.



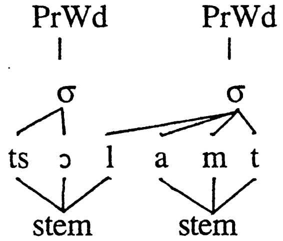
ONSET: *
ALIGN-L: √
FAITH: √

b.



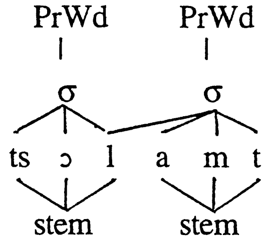
ONSET: √
ALIGN-L: *
FAITH: *

c.



ONSET: √
ALIGN-L: *
FAITH: √

d.



ONSET: √
ALIGN-L: √
FAITH: √

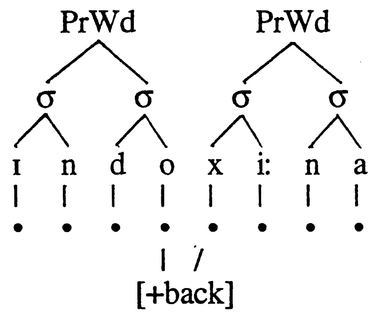
While the ranking of ONSET above ALIGN-L rules out the properly aligned but onsetless candidate (26a), under Itô & Mester's interpretation of alignment the ambisyllabic candidate (26d) also satisfies ALIGN-L and so would be given preference over the correct output form (26b), if no constraints other than ONSET and ALIGN-L are taken into consideration.[10] The difference between the optimal candidate (26b) and the suboptimal candidate (26d) is that the former has crisp word edges, justifying the appeal to the constraint CRISPEDGE (PrWd). If CRISPEDGE (PrWd) is ranked above ALIGN-L, candidate (26b) will emerge as the winner.

That it is indeed the prosodic word and not the syllable that is required to have crisp edges follows from the fact that multiple linking across a syllable boundary is not generally excluded in German. There is evidence from the distribution of long and short vowels that medial consonants in forms like *Hölle* [hœlə] 'hell' or the name *Otto* [ɔto] are ambisyllabic (Vennemann 1972, 1982;

Benware 1986; Wiese 1986; Ramers 1992, Féry 1995). Short lax vowels cannot occur in word-final position which suggests that German syllables are at least bimoraic. Word-medial consonants that are preceded by a short lax vowel must therefore be in the coda to satisfy the bimoraic requirement. Since they also form the onset of the following syllable they are ambisyllabic. That gemination is inhibited across a compound boundary is therefore not motivated by a restriction on German syllable structure, but must be rooted in the more specific requirement of prosodic words to have crisp edges.

Let us now consider the process of dorsal fricative assimilation in greater detail. As already mentioned above, dorsal fricative assimilation is strictly limited to word-internal contexts; i.e., it fails to apply in compounds or between words at the phrase level. Since assimilation applies freely across a syllable boundary, as demonstrated earlier by the form *Kuchen* [ku:.xən] 'cake', we cannot call upon a syllable structure requirement of German to explain why dorsal fricative assimilation fails to apply in compounds. We suggest instead that the non-application of fricative assimilation follows from the constraint CRISPEDGE (PrWd). Assuming that assimilation is feature spreading, the propagation of [+back] from the stem vowel in (27) to the following dorsal results in a multiply linked structure and consequently a non-crisp word edge.

(27)



*[ɪndoxi:na] 'Indo-china'

Provided CRISPEDGE (PrWd) ranks higher than CVLINKAGE, the initial dorsal fricative of the second stem fails to undergo assimilation and so surfaces as front.

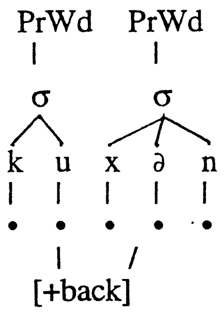
(28) Input: /ɪ n d o + ç i: n a/ 'Indo-china'

+bk -bk

candidates	CRISPEDGE	CVLINKAGE	FAITH [αback]
a. ɪ n d o x [i: n a | / +bk <-bk>	*!		***
b. ☞ ɪ n d o [ç i: n a | | +bk -bk		*	
c. ɪ n d o [x i: n a | | +bk *+bk* <-bk>		*	****
d. ɪ n d o [X i: n a | +bk <-bk>		*	**

If we assume that *-chen* is a prosodic word of German, we can furthermore explain why its initial dorsal fricative escapes assimilation. Assimilation is blocked in this environment, because it results in a non-crisp word edge.

(29)



*[ku:xən] 'cow, dim.'

(30) Input: /k u: + ç ə n/ 'cow, dim.'
| |
+bk -bk

candidates	CRISPEDGE	CVLINKAGE	FAITH [αback]
a. k u: x [ə n | / +bk <-bk>	*!		***
b. ☞ k u: [ç ə n | | +bk -bk		*	
c. k u: [x ə n | | +bk *+bk* <-bk>		*	****

d. k u: [X ə n | +bk <-bk>		*	**

Finally, let us reconsider the umlaut problem. As already pointed out in section 2, the fact that the diminutive suffix *-chen* causes umlaut of a preceding stem vowel is difficult to reconcile with the assumption that it is a "non-cohering" suffix or prosodic word of German. However, under the present approach, the claim that the diminutive suffix is a separate prosodic word of German can be maintained despite the fact that it causes umlaut of the base of affixation. Since umlaut consists of the anchoring of an underlyingly floating feature value [-back] with a back root vowel in the output (Klein 1995), it violates neither ALIGN-L nor CRISPEDGE (PrWd). As a matter of fact, the advantage of the alignment approach over an approach within prosodic lexical phonology emerges most clearly in the analysis of the interaction of dorsal fricative assimilation and umlaut in MSG. By focusing on the edge of the prosodic word, alignment theory is able to make use of a qualitative difference between these two processes to explain why assimilation fails to apply between two prosodic words, while umlaut can take place in this context: Fricative assimilation is a spreading operation, which results in blurred word edges; umlaut is a feature insertion process. Note that I am not claiming that umlaut can apply between words; I am claiming that umlaut simply does not bear on the question of what a phonological word is.

4 Conclusion

In this paper I have presented novel solutions to the invariance of the German diminutive suffix *-chen* within the model of lexical phonology and within optimality theory. In section 2 I have argued that the invariance of the diminutive suffix can be accounted for by assuming that dorsal fricative assimilation applies at level I, before the diminutive morpheme is suffixed. The solution within optimality theory, on the other hand, depends on the recognition of *-chen* as a separate prosodic word of German. If *-chen* is a prosodic word, the constraint CRISPEDGE (PrWd) accounts for the non-application of fricative assimilation to its initial dorsal fricative.

NOTES

* I would like to thank Lise Dobrin, Caroline Féry, Tracy Alan Hall, Ursula Kleinhenz and Sylvia Löhken for comments on an earlier draft of this paper.

[1]According to Kohler (1977), the velar variant [x] is found after the back tense vowels [u:] and [o:], the uvular fricative [χ] after the low vowel [a], while either [x] or [χ] can occur after the non-low back lax vowels [ʊ] and [ɔ]. This distinction will be ignored here.

[2]Hall (1989) includes the specification [-voice] in order to distinguish the voiceless dorsal fricative /X/ from /j/. Unlike Hall, I assume that /j/ is a glide and not a fricative, making any reference to the feature value [-voice] unnecessary. In the remainder of this paper I will refer to /X/ as simply the dorsal fricative. Also, note that Hall uses [back] to refer to both values [+back] and [-back]; hence, on Hall's account vowels are specified as either [+back] or [-back] when FA applies.

[3]Wiese (1986) assumes that there are three lexical strata in MSG: Stratum I consists of the stress-attracting derivational affixes, stratum II of the stress-neutral derivational affixes and compounding, and stratum III of the inflectional affixes.

[4]I assume that umlaut is induced by suffixes which contain a floating feature [-back] underlyingly (Lieber 1987, 1992).

[5]This fact was brought to my attention by John Goldsmith.

[6]Alternatively, we could assume that FA is a lexical rule that operates in violation of the principle of structure preservation. Much has been said in the literature about the implications this assumption has for the model of lexical phonology, cf. Borowsky (1986), Macfarland & Pierrehumbert (1991) among others.

[7]Wiese (1996) provides interesting evidence from a class of 'gapping' phenomena in favor of the assumption that *-chen* is an independent phonological word of MSG. He observes that the suffix *-chen* can be deleted from conjoined expressions such as ***Väter- und Mütterchen*** 'father-dim. and mother-dim.' or ***Brüder- und Schwesterchen*** 'brother-dim. and sister-dim.'. A comparison with ill-formed examples like ***Komponist- und Lehrerin*** 'composer-fem. and teacher-fem.' or **Versicher-*

und Verwaltung 'insurance and administration' shows that this is by no means a property of all derivational suffixes in German. One condition on deletion is that the suffix must be co-extensive with a syllable. Not just any syllable can be deleted, however, as evidenced by **male- oder wählerisch* 'picturesque or choosy'. Wiese (1996) concludes that only suffixes that form independent words can be deleted and *-chen* is one of those.

[8]Alternatively, we could formulate a constraint ALIGN-R ([+back], dorsal C) which requires the feature value [+back] to align with a dorsal consonant on its right (Kirchner 1993). However, since the domain of alignment is not a prosodic constituent, but simply the linear string VC, an analysis in terms of a linked VC domain appears more appropriate.

[9]To account for the contrastive pattern of epenthesis in the monomorphemic forms below, Féry (1995) - following Hall (1992) and Yu (1992) - suggests that the <u>foot</u> and not the prosodic word boundary is the location for glottal stop insertion. While the examples in (a) consist of two feet, the examples in (b) consist of a single foot.

a.	Ruin	[(ʀu:.)$_{Ft}$(ʔí:n)$_{Ft}$]	'ruin'
	Theater	[(te:.)$_{Ft}$(ʔá:tʌ)$_{Ft}$]	'theater'
b.	Fluor	[(flú:.ɔʀ)$_{Ft}$]	'fluorine'
	Boa	[(bó:.a)$_{Ft}$]	'boa'

On Féry's account it is the left edge of the foot that needs to align with the left edge of the syllable, ALIGN (Foot, L, σ, L). It could be argued that the constraint ALIGN (Stem, L, PrWd, L) is needed independently to account for glottal stop insertion in compounds like *Zollamt* [tsɔl.ʔamt] 'customs-house' or *Bauamt* [bau.ʔamt] 'building office'. According to Féry (1995) these forms consist of two feet, motivating the insertion of a glottal stop before the vowel of the second morpheme. The argument that the second stem in *Zollamt* [tsɔl.ʔamt] 'customs-house' forms a metrical foot of its own is in part based on the assumption that it carries secondary stress. It contrasts in this respect with the second syllable in words like *Hering* [hɛ.ʀɪŋ] 'herring' or *König* [køniç] 'king' which are unstressed. As Wiese (1996: 275) points out, however, speakers of German have no clear intuition about the stress patterns of bisyllabic words with a strong-weak pattern, making a distinction between syllables with secondary stress and unstressed syllables doubtful. This in turn removes some of the motivation for assuming that the second syllable in *Zollamt* [tsɔl.ʔamt] 'customs-house' forms an independent foot in German, leaving this issue unsettled.

[10]Candidate (26b) is ultimately chosen over (26c), because ALIGN-L is evaluated gradiently. The insertion of an empty position, which is filled out by a glottal stop in German, constitutes a less severe violation of ALIGN -L than syllabifying the liquid into the onset of the second syllable.

REFERENCES

Benware, Wilbur A. (1986). *Phonetics and phonology of modern German: An introduction.* Washington: Georgetown University Press.

Bloomfield, Leonard (1930). German ç and x. *Le maitre phonétique* 3:20. 27-28.

Booij, Geert & Jerzy Rubach (1984). Morphological and prosodic domains in lexical phonology. *Phonology yearbook* 1. 1-27.

- (1987). Postcyclic versus postlexical rules in lexical phonology. *LI* 18. 1-44.

Borowsky, Toni (1986). *Topics in the lexical phonology of English.* Ph.D. dissertation, University of Massachusetts, Amherst. Published by Garland Press, New York, 1990.

- (1993). On the word level. In Sharon Hargus and Ellen M. Kaisse (eds.) *Studies in lexical phonology.* San Diego: Academic Press. 199-234.

Féry, Caroline (1995). Alignment, syllable and metrical structure in German. Ms., Universität Tübingen.

Fleischer, Wolfgang (1975). *Wortbildung der deutschen Gegenwartssprache.* Tübingen: Niemeyer.

Hall, Tracy Alan (1989). Lexical phonology and the distribution of German [ç] and [x]. *Phonology* 6:1. 1-17.

- (1992). *Syllable structure and syllable-related processes in German.*

Tübingen: Niemeyer.

Hayes, Bruce (1986a). Assimilation as spreading in Toba Batak. *LI* 17. 467-99.

- (1986b). Inalterability in CV phonology. *Lg* 62:2. 321-52.

Itô, Junko & Armin Mester (1995). The core-periphery structure of the lexicon and constraints on reranking. In Jill N. Beckman et al. (eds.) *Papers in optimality theory*. Amherst, Massachusetts: GSLA. 181-210.

- (in press). Realignment. In René Kager et al. (eds.) *Proceedings of the Utrecht workshop on prosodic morphology*. Mouton: The Hague.

Itô, Junko, Armin Mester & Jaye Padgett (1995). Licensing and underspecification theory. *LI* 26:4. 571-613.

Iverson, Gregory K. & Joe Salmons (1992). The place of structure preservation in German diminutive formation. *Phonology* 9:1. 137- 43.

Kiparsky, Paul (1982). Lexical morphology and phonology. In I.-S. Yang (ed.) *Linguistics in the morning calm*. Seoul: Hanshin. 3-91.

- (1985). Some consequences of lexical phonology. *Phonology* 2. 85-138.

Kirchner, Robert (1993). Turkish vowel disharmony in optimality theory. Paper presented at the Rutgers Optimality Workshop #1, Rutgers University.

Klein, Thomas B. (1995). Umlaut in optimality theory. Ph.D. dissertation, University of Delaware.

Kohler, Klaus (1977). *Einführung in die Phonetik des Deutschen*. Berlin: E. Schmidt.

Lieber, Rochelle (1987). *An integrated theory of autosegmental processes*. New York: SUNY Press.

- (1992). *Deconstructing morphology*. Chicago: University of Chicago Press.

McCarthy, John & Alan Prince (1993a). Prosodic morphology I. Ms, University of Massachusetts, Amherst & Rutgers University.

- (1993b). Generalized alignment. In Geert Booij and Jaap van der Marle (eds.) *Morphology yearbook 1993*. Dordrecht: Kluwer.

- (1995). Faithfulness and reduplicative identity. In Jill N. Beckman et al. (eds.) *Papers in optimality theory*. Amherst, Massachusetts: GSLA. 249-384.

Macfarland, Talke & Janet Pierrehumbert (1991). On ich-laut, ach-laut and structure preservation. *Phonology* 8:1. 171-80.

Prince, Alan & Paul Smolensky (1993). Optimality theory: Constraint interaction in generative grammar. Ms, Rutgers University and University of Colorado, Boulder.

Ramers, Karl Heinz (1992). Ambisilbische Konsonanten im Deutschen. In Peter Eisenberg et al. (eds.) *Silbenphonologie des Deutschen*. 246-83.

Vennemann, Theo (1972). On the theory of syllabic phonology. *Linguistische Berichte* 18. 1-18.

- (1982). Zur Silbenstruktur der deutschen Standardsprache. In Theo Vennemann (ed.) *Silben, Segmente, Akzente*. Tübingen: Niemeyer. 261-305.

Wiese, Richard (1986). Schwa and the structure of words in German. *Linguistics* 24. 695-724.

- (1996). *The phonology of German*. Oxford: Clarendon Press.

Yu, Si-Taek (1992). *Unterspezifikation in der Phonologie des Deutschen*. Tübingen: Niemeyer.

Articulatory reduction in fluent speech
A pilot study on syllabic [n] in spoken Standard German

Bernd Pompino-Marschall

Zentrum für Allgemeine Sprachwissenschaft

Jägerstr. 10/11
D-10117 Berlin
Germany

SUMMARY

The paper reports of an acoustic and perceptual study on reduction of the German word final syllable *'_en'* [@n][1] in contrast to single word final [n]. Three utterances of *'ein'* and *'einen'* were cut from the same sentence utterances of four different speakers and analysed according to their segment durations. In a perception experiment these utterances had to be identified. Statistical analyses revealed that the perception of *'einen'* is mainly dependent on the duration of the final [n] segment itself. Detailed analysis of the results however showed that there may be quite complex articulatory reorganization resulting in the perception of the so called syllabic [n].

INTRODUCTION

Processes of reduction in fluent standard German have onlyoccasionally been studied (for more recent empirical studies see Wesenick 1994, Kohler 1996, for modelling Kohler 1990, 1992, Kröger 1993). The present paper describes an acoustic/auditory pilot study of schwa elision in nasal context resulting in the so called syllabic [n]. An identification test was run to determine the acoustic cues for the perception of syllabic [n].

STIMULI

As stimuli utterances of the words *'ein'* [QaIn] (German indefinite article, masculine, nominative case) and *'einen'* [QaIn@n] (indefinite article, accusative case masculine) from the PHONDAT 2 (cf. Pompino-Marschall 1992) Kiel corpus of read speech (IPDS 1995, Kohler 1994) were chosen. The utterances were selected to represent quite different forms of reductions of the *'einen'* items according to the original Kiel segmentation/transcription (IPDS 1995, Kohler 1994), i.e. items with clear as well as unclear segment boundaries between the single [n] instances and items with further [n] reduction. This selection resulted in the instances of the sentence utterances

Wann geht bitte <u>ein</u> Zug nach Hamburg mit möglichst wenig, ja, Bahnhöfen? (e068)

Was ist die billigste Möglichkeit für <u>ein</u> Bahnticket von hier nach Kiel? (e083)

Wann fährt bitte <u>ein</u> Zug von Regensburg nach Dortmund am Montag, am Montagmorgen? (e097)

Gibt es <u>einen</u> Zug, der in Köln länger hält und nach Dortmund weiterfährt? (e076)

---

[1] Throughout the paper transcriptions are given in SAMPA (Wells et al. 1992) notation.

Gibt es einen Nachtzug, der am dreiundzwanzigsten Zwölften morgens in Ulm ankommt? (s055)

Gibt es einen Zug, der am dreiundzwanzigsten Zwölften um zirka zwanzig Uhr in Oldenburg ist? (s063)[2]

produced by four speakers (3 male: dlm, hpt, uga; 1 female: rtd[2]). All speakers had a standard German pronunciation (northern provenance).

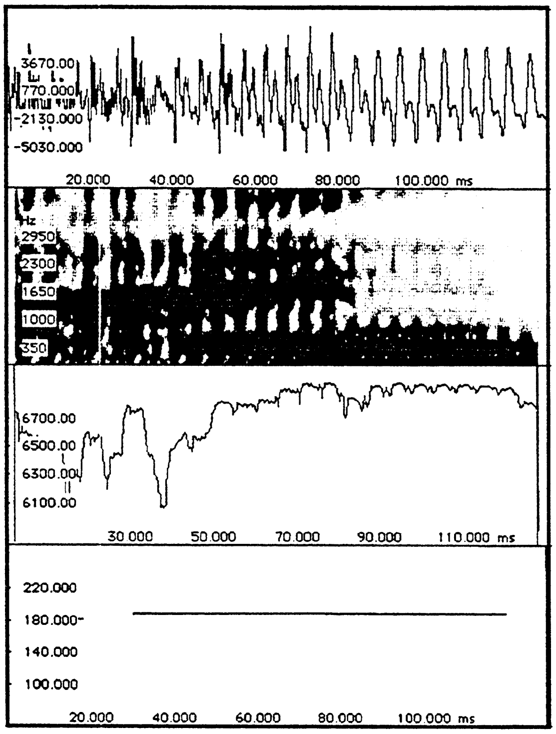
The stimuli for the perception experiment were cut from the sentence utterances according to the original Kiel segmentation (i.e. from the beginning of the (glottalized) diphthong [aI] till the end of the last segmented/labelled [n] of the indefinite article. In figures 1 to 4 the audio signals, sonagrams, RMS amplitude and fundamental frequency contours of representative items of *'ein'* and *'einen'* are depicted for the four speakers.[3] Table I lists the durations of the single segments according to the Kiel segmentation (reanalysed with respect to the duration of the glottalized segment).

Table I: Segment durations (in ms) of the selected utterances

				item				
			'ein'				'einen'	
speaker		e068	e083	e097		e076	s055	s063
dlm	q	-	-	46		-	-	48
	aI	68	40	37		95	77	64
	n	39	53	47		22	52	42
					n	17	-*	%22
hpt	q	67	-	62		86	68	39
	aI	69	43	40		79	62	150
	n	31	61	38		50	47	27
					n	-	%34*	26
rtd	q	13	42	63		88	87	57
	aI	67	42	46		71	75	83
	n	29	69	43		26	61	81
					n	%44	-*	-
uga	q	-	-	-		-	23	-
	aI	88	70	94		103	83	112
	n	33	69	50		35	55	40
					n	%42	55	41

% denotes a not clearly definable segment boundary,
items marked with an asterisk are followed by a not clearly segementable word initial [n].

[2] The canonical pronunciation of the sentences in SAMPA is given in the appendix.
[3] The analyses as well as the preparation of the test tape were conducted with the help of the Signalyze 3.11 software for Apple MacIntosh.



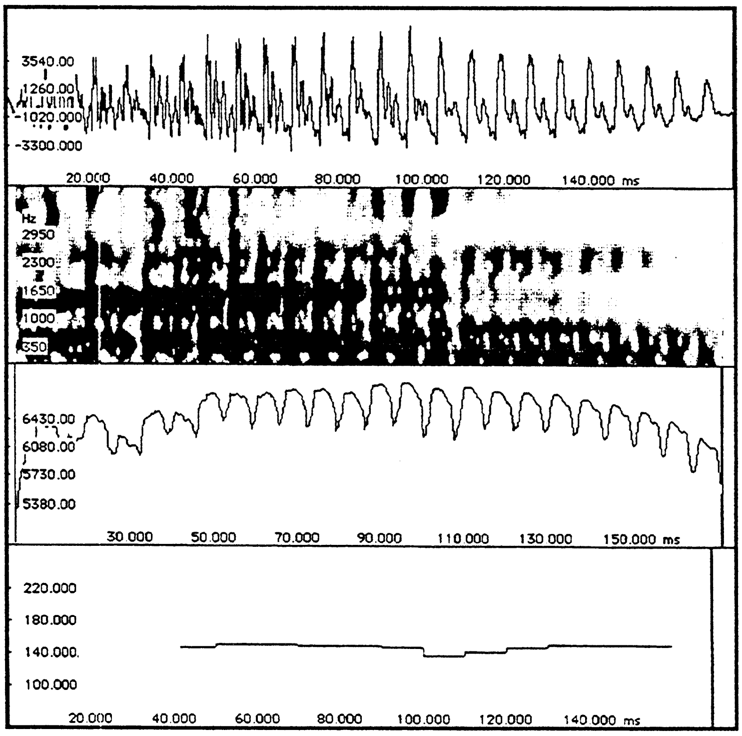


Fig. 1: Examples of utterances of *'ein'* (top; e097) and *'einen'* (bottom; s063) for speaker dlm; first row: audio signal; second row: wide band (200Hz) sonagram; third row: RMS amplitude (10 ms window); fourth row: fundamental frequency.

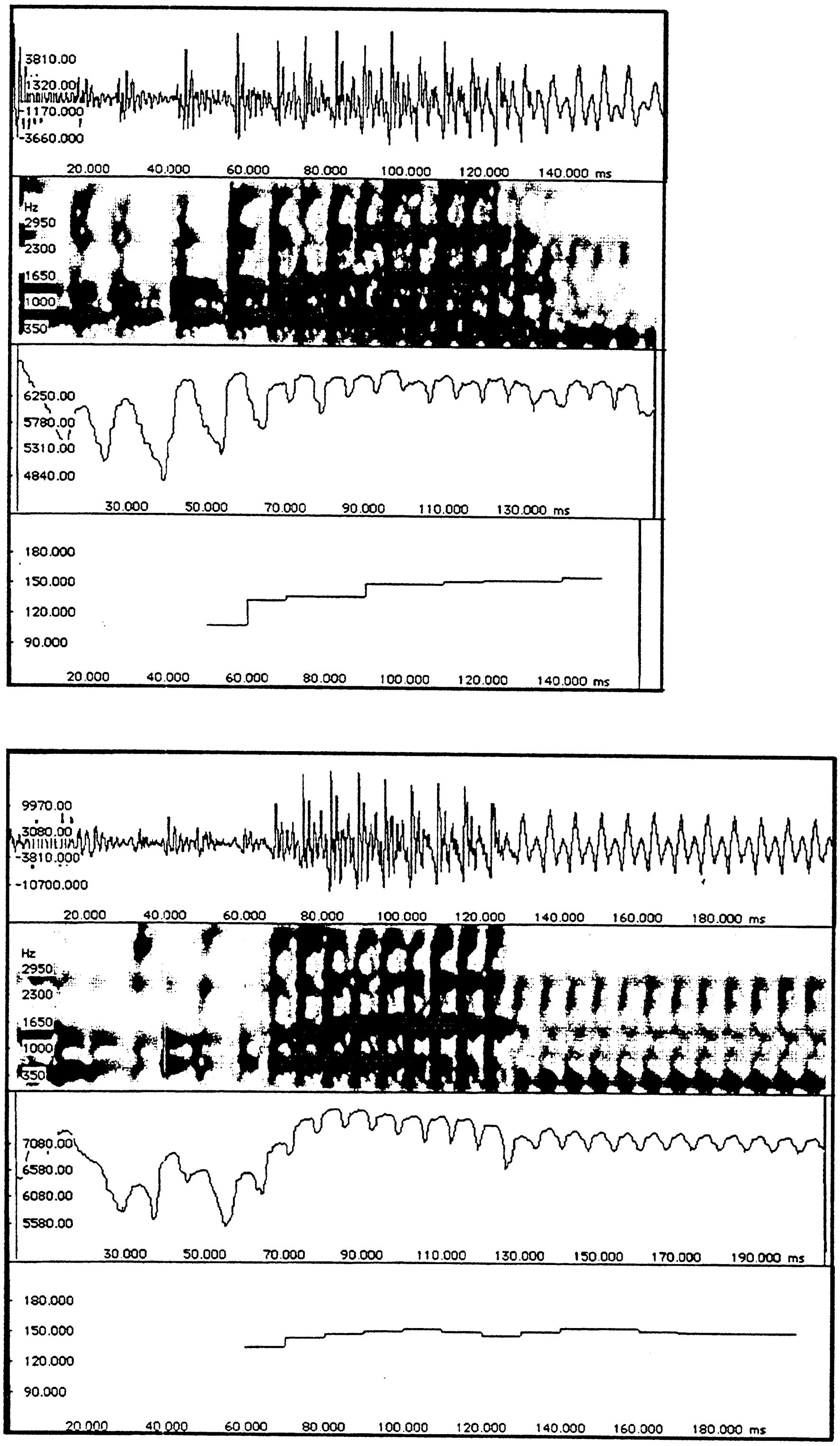
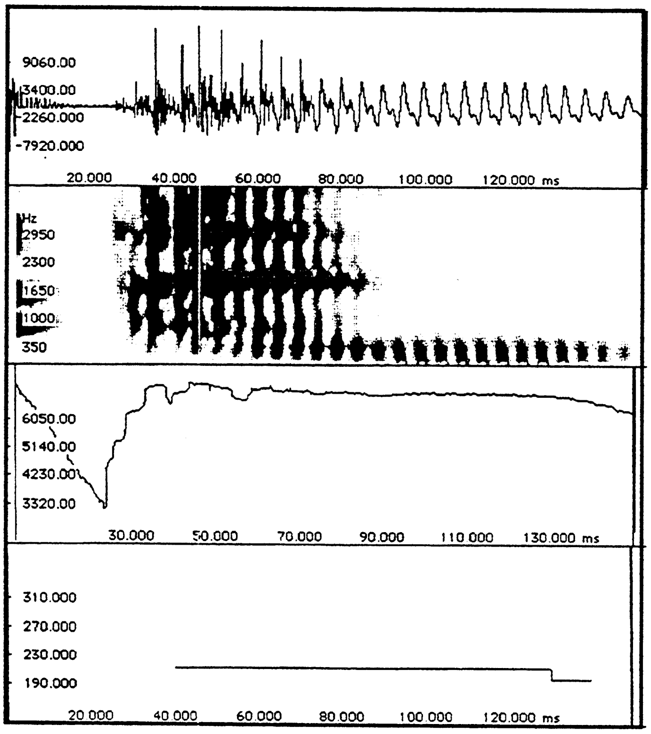


Fig. 2: Examples of utterances of *'ein'* (top; e068) and *'einen'* (bottom; s055) for speaker hpt (format as in figure 1).



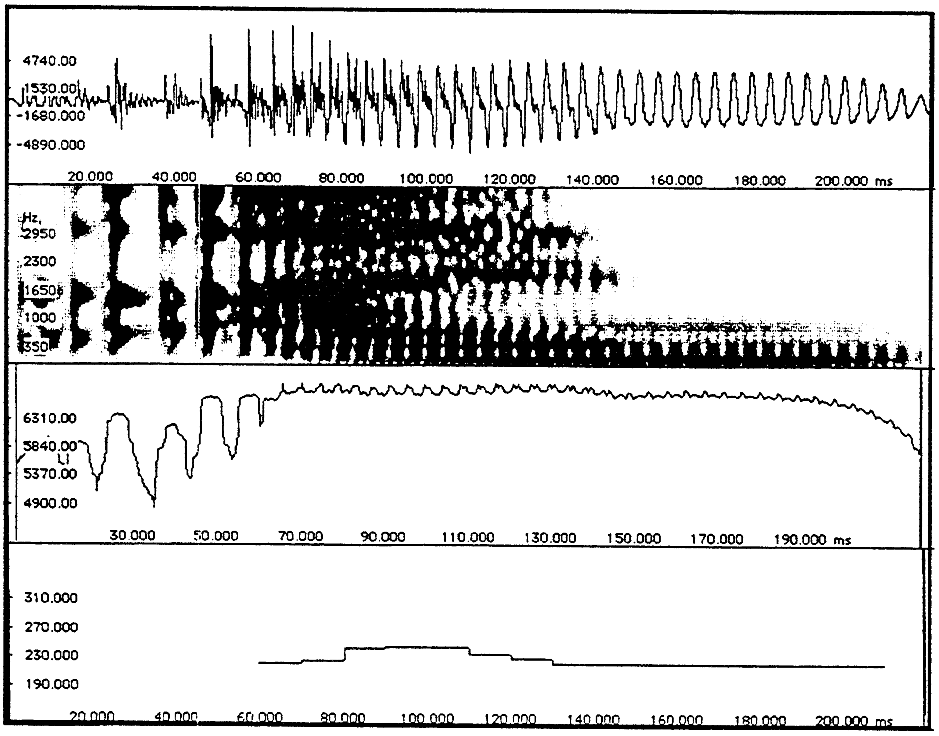
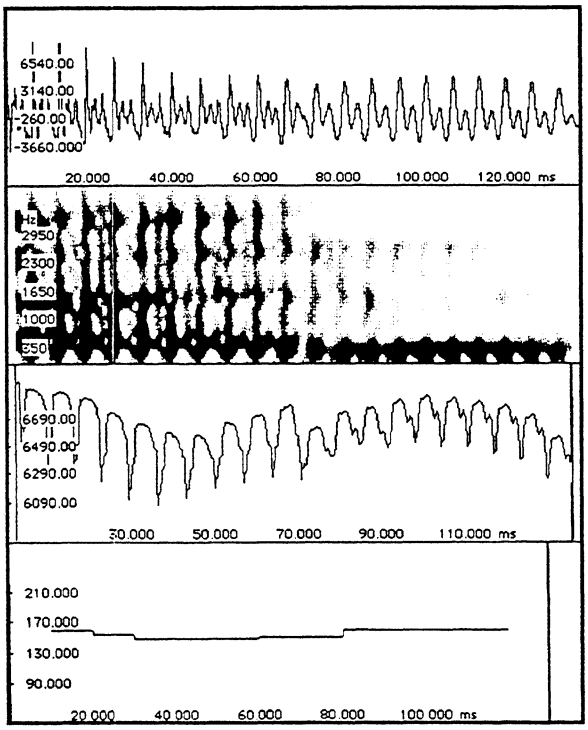


Fig. 3: Examples of utterances of *'ein'* (top; e083) and *'einen'* (bottom; s063) for speaker rtd (format as in figure 1).



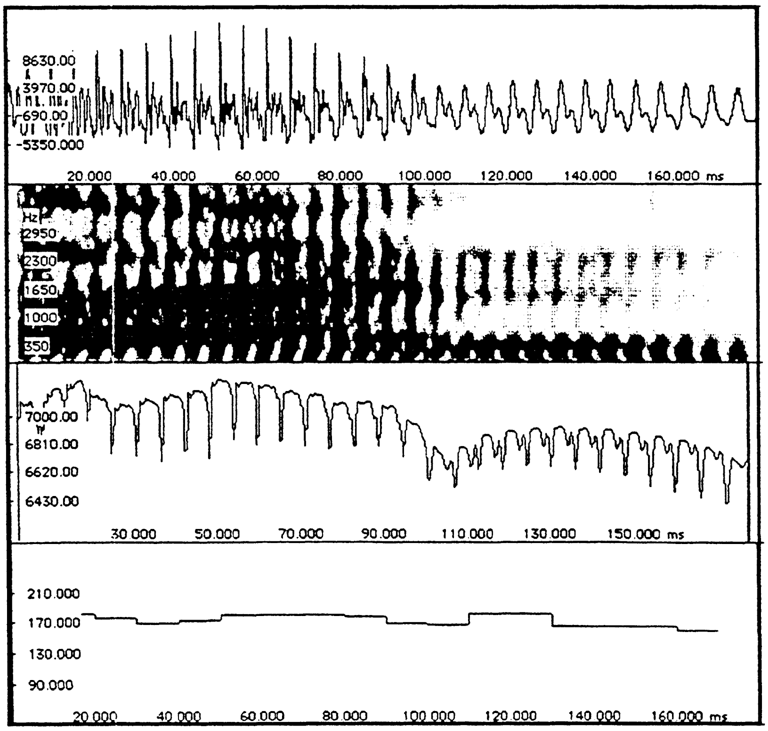


Fig.4: Examples of utterances of *'ein'* (top; e083) and *'einen'* (bottom; e076) for speaker uga (format as in figure 1).

EXPERIMENTAL PROCEDURE

The isolated *'ein'/'einen'* utterances were presented ten times in quasi randomized order to ten subjects (5 male, 5 female) who had to decide whether they heard *'ein'* or *'einen'*. The subjects listened to the stimuli via headphones at a comfortable loudness level at an interstimulus interval of about 3 s and a longer pause after blocks of 12 trials. They had to mark their choices on a prepared answer sheet.

RESULTS

The overall categorical responses to the single items is depicted in figure 5, the responses of the single subjects are listed in Table II.

Figure 5 clearly shows differences in the identification of the test items of different speakers and of single items as well. In Table II further differences in the responses of single listeners can be found (cf. below).

The *'ein'* utterances were identified correctly at 80.17 percent, the reduced *'einen'* utterances at only 66.08 percent. The latter is clearly due to the poor recognition of the speaker's dlm *'einen'* items; for the other speakers we get an identification score of 77% (again mainly resulting from difficulties with the single item e083 of speaker rtd; this item excluded: 81%) for *'ein'* and 80% for *'einen'*.

DISCUSSION

General observations

When cutting the test items from the utterances it was recognized that the glottalized portion of the diphthong [aI] often results in the perceptual illusion of a prefixed unstressed [d@] syllable. For the further analysis of segmental durations the utterances were therefore reanalysed according to the duration of glottalization as already shown in Table I.

Acoustic segment durations

Differences in segment durations between items of *'ein'* and *'einen'* are observable for the duration of the final [n] and the duration of the nonglottalized part of the diphthong [aI]. Overall, the difference in [n] duration - *'ein'*: 58.7 ms (19.89) vs. *'einen'*: 87.8 (24.60)[4] - is clearly significant ($t = 3.121$; $p < .01$): *'einen'* clearly has longer [n] durations. The different speakers show different lengthening of [n]: speaker dlm exhibits an only marginal lengthening of 12% whereas the speakers hpt, rtd, and uga show lengthenings of 42, 50, and 76% respectively. The diphthong [aI] - *'ein'*: 58.7 (19.89); *'einen'*: 87.8 (24.60) - is also significantly longer for *'einen'* ($t = 3.143$; $p < .01$). The different speakers again show different lengthening ratios not paralleling those for [n]. Clearly, the different context in sentence e083 - [fy:6 QaIn] - resulted in an extremely short diphthong in these items.[5] But even when these items were excluded from statistical analysis a significant difference in diphthong duration between *'ein'* and *'einen'* remained ($t = 2.362$; $p < .025$).

---

[4] Means and standard deviations (in brackets) in ms.

[5] In colloquial German there would even be the possibility of a total elision - [fy:6 n].

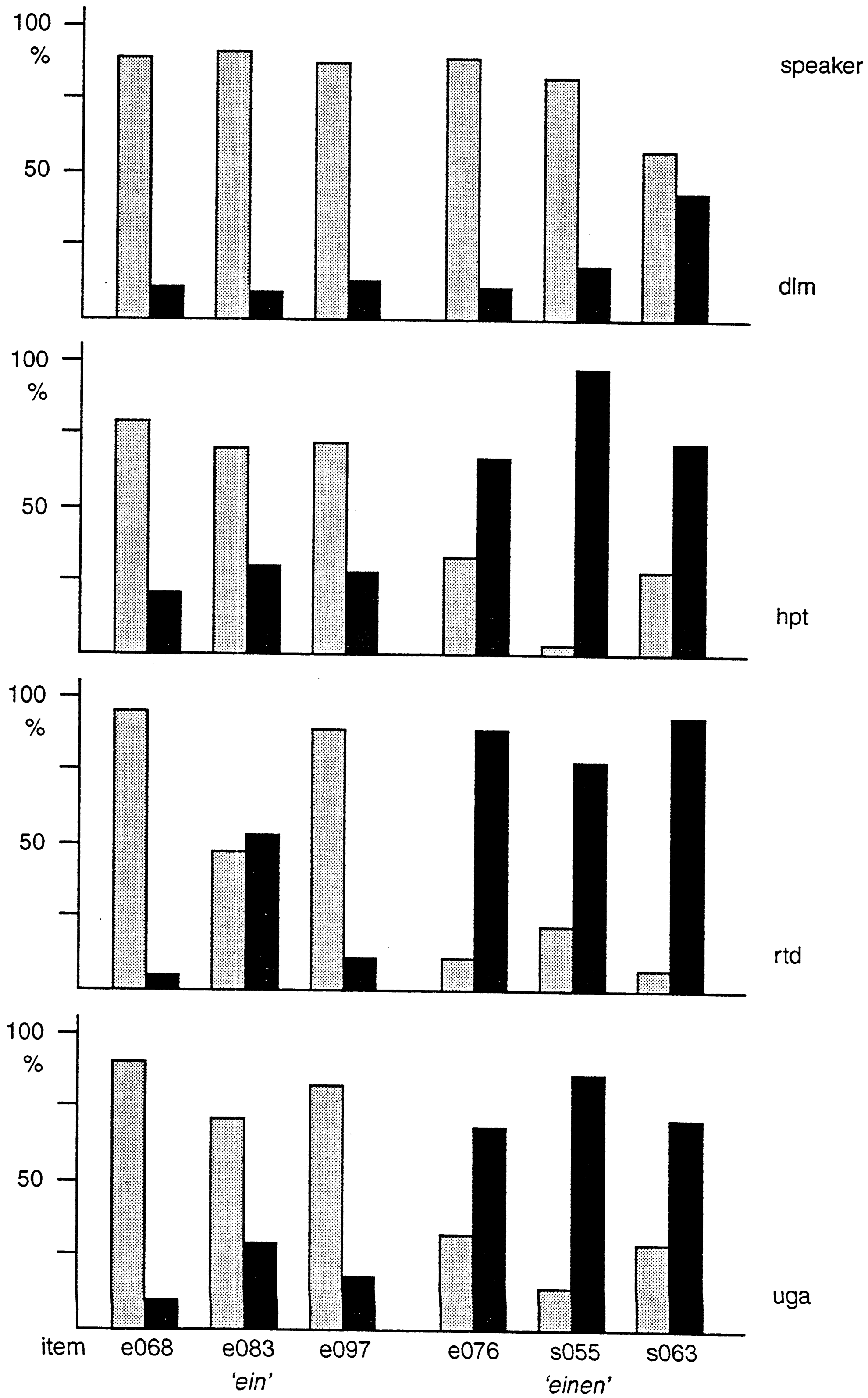


Fig. 5: Overall identification scores in percent; gray: *'ein'*, black: *'einen'*.

Table II: Response frequencies of *'ein'* (left columns) and *'einen'* (right columns) for the single subjects (single rows) and mean percentage (rows labelled sum).

speaker		10	-	9	1	10	-	10	-	10	-	6	4
		9	1	10	-	10	-	10	-	2	8	6	4
		10	-	10	-	10	-	10	-	10	-	7	3
		10	-	10	-	9	1	9	1	9	1	6	4
dlm		7	3	6	4	5	5	8	2	5	5	4	6
		10	-	10	-	8	2	8	2	10	-	2	8
		10	-	8	2	10	-	10	-	10	-	9	1
		8	2	9	1	7	3	8	2	9	1	2	8
		8	2	9	1	9	1	9	1	8	2	7	3
		7	3	10	-	9	1	7	3	9	1	8	2
	sum	89	11	91	9	87	13	89	11	82	18	57	43
		9	1	9	1	6	4	1	9	-	10	-	10
		10	-	10	-	10	-	9	1	-	10	10	-
		7	3	10	-	9	1	2	8	1	9	-	10
		5	5	8	2	5	5	-	10	-	10	1	9
hpt		4	6	1	9	3	7	3	7	-	10	1	9
		10	-	6	4	10	-	5	5	-	10	2	8
		8	2	5	5	8	2	10	-	1	9	9	1
		7	3	2	8	7	3	-	10	-	10	1	9
		10	-	9	1	6	4	3	7	-	10	3	7
		9	1	10	-	8	2	-	10	1	9	1	9
	sum	79	21	70	30	72	28	33	67	3	97	28	72
		10	-	9	1	8	2	-	10	-	10	-	10
		9	1	2	8	10	-	2	8	5	5	-	10
		10	-	7	3	9	1	-	10	4	6	-	10
		10	-	10	-	10	-	-	10	-	10	-	10
rtd		7	3	4	6	8	2	1	9	5	5	2	8
		10	-	3	7	10	-	1	9	-	10	1	9
		10	-	2	8	9	1	2	8	4	6	2	8
		10	-	-	10	10	-	1	9	2	8	-	10
		10	-	6	4	9	1	3	7	1	9	1	9
		9	1	4	6	6	4	1	9	1	9	1	9
	sum	95	5	47	53	89	11	11	89	22	78	7	93
		8	2	8	2	10	-	2	8	-	10	1	9
		10	-	2	8	7	3	2	8	2	8	3	7
		10	-	10	-	10	-	5	5	4	6	7	3
		10	-	9	1	9	1	6	4	-	10	3	7
uga		5	5	1	9	3	7	3	7	-	10	2	8
		10	-	7	3	8	2	-	10	-	10	1	9
		10	-	9	1	10	-	7	3	3	7	8	2
		8	2	8	2	7	3	-	10	-	10	1	9
		10	-	9	1	9	1	5	5	3	7	3	7
		9	1	8	2	9	1	2	8	2	8	-	10
	sum	90	10	71	29	82	18	32	68	14	86	29	71
	item	e068		e083		e097		e076		s055		s063	

Acoustic cues for syllabic [n]

Since no clear indications for a syllabic [n] - partly contrary to the expectations according to the original Kiel segmentation/transcription - could be found in the RMS amplitude contour as well as the fundamental frequency contour (cf. figures 1 to 4), the correlations between different segment durations and response frequencies were calculated.

Clearly, here the correlation between the duration of the (complex) [n] segment and the frequency of *'einen'* responses wins with $r = .7878$ ($p < .001$). Contrary to the primary expectation the correlation between the ratio of the duration of the nonglottalized vocalic portion and [n] failed significance ($r = -.4916$; n.s.).[6]

Specific observations

Some results concerning the different speakers, indiviual test items as well as single listeners require some further attention. The overall poor recognition (cf. figure 5 & table II) of the items of speaker dlm even when the original Kiel segmentation/transcription would predict otherwise (cf. table I, especially dlme076) may be interpreted as due to this speaker's weak variation in [n] duration (cf. above) between *'ein'* and *'einen'*. A clear effect of [n] duration also seems to be the cause for the poor identification of rtde083: here the [n] duration of *'ein'* is in the same range as in this speaker's *'einen'* items. For some listeners this [n] lengthening in e083 also results in categorical reversals for speaker uga.

The most interesting result is the categorical reversal in the identification of *'einen'* (s063) of speaker hpt by listeners 2 and 7. When inspecting this item (cf. figure 6) it can be seen that there is no obvious [n] lengthening but that the vocalic portion consists of two parts divided by a decrease in amplitude: an [a] part of about 49 ms and a diphthongal part. If the former is cut from the quite normal sounding original utterance as visualized in figure 6 a clearly ungrammatical utterance will result in that it is clearly perceived as *'Gibt es ein* (nom.) *Zug ...'* instead of *'Gibt es einen* (acc.) *Zug ...'*. What we get here is not a syllabic [n] but a bisyllabic indefinite article. Without the sentence context however, this bisyllabicity seemed to be no cue for these two listeners.

General discussion

The main cue for the perception of syllabic [n] therefore seems to be the duration of the [n] segment itself. If there is a further dependency on adjacent vowel duration which cannot finally be excluded from the present results this vowel duration is the duration of the nonglottalized vowel portion alone.

In the present data no indication of a really syllabic [n] could be found. In connected speech however, bisyllabicity on the other hand may not be cued by the [n] segment itself but also by an articulatory reorganization of adjacent segments. A prototypical example of such reorganization was recently reported by Tillmann (1995): The - by German hearers - clearly perceived utterance of the word *'und'* [QUnt] in the sentence *'Ich will nach Hamburg* [pause] *<u>und</u> möchte dort etwa gegen sieben Uhr abends ankommen'*[7] was only a glottalized pitch rise on the following [m] resulting in a syllabic [m] before the following word. A similar reorganization in articulatory timing resulted in the bisyllabicity of the vocalic nucleus of the *'einen'* item of hpts063 discussed above.

---

[6] No correlations at all resulted when the glottalized portion was included in this proportion.

[7] Also from the PHONDAT 2 corpus; Munich speaker.

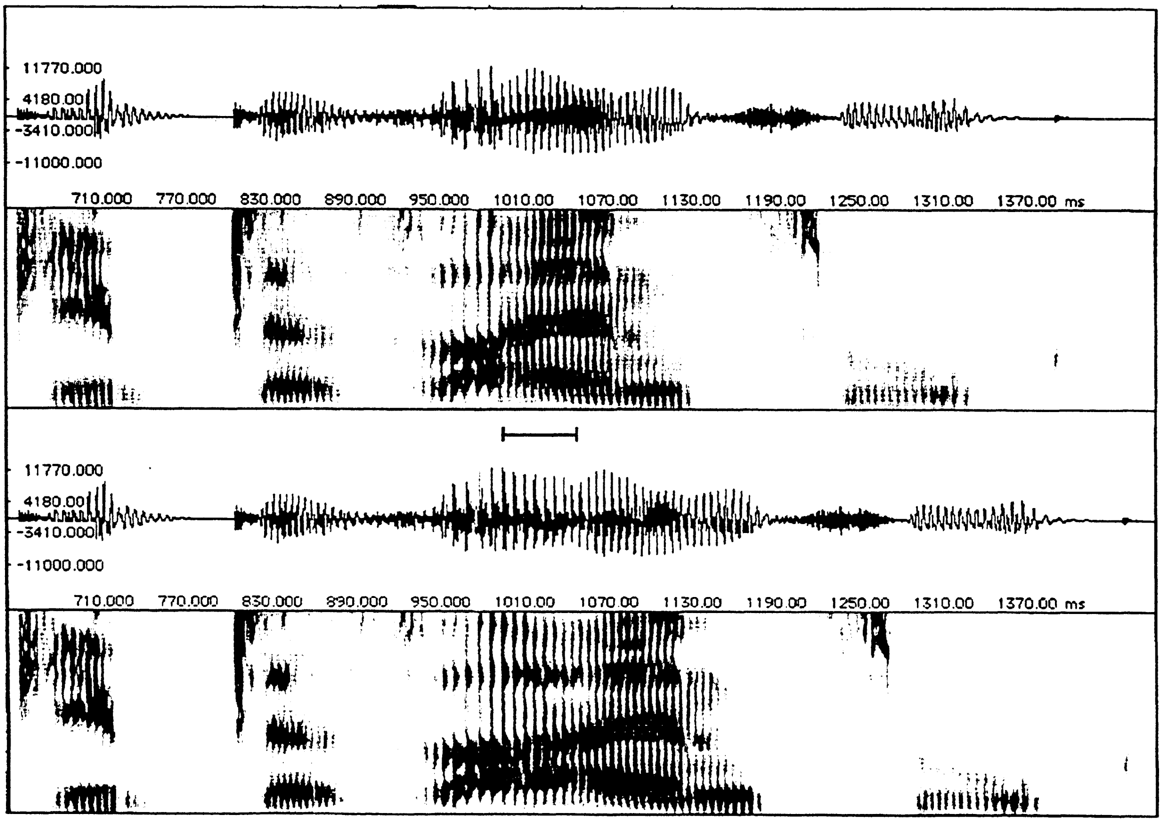


Fig. 6: Audio signals and sonagrams of the manipulated (top) and the original (bottom) sentence part "Gibt es einen Zug ...?" (s063 of speaker hpt): In the manipulated utterance the marked [a] segment of 49 ms was cut.

The results of the present pilot investigation clearly show that much can be learned about the articulatory organization in fluent speech by a thorough acoustic analysis (accompanied by perceptual evaluation) of larger data bases of connected/spontaneous speech as are only recently available for German through the PHONDAT project.

REFERENCES

IPDS (1995), CD-ROM#1: The Kiel Corpus of Read Speech, vol. I. Kiel.

Kohler, K.J. (1990), Segmental reduction in connected speech in German: phonological facts and phonetic explanations. In: Hardcastle, W.J. & Marchal, A. (eds.), Speech Production and Speech Modelling. Dordrecht, 69-92.

Kohler, K.J. (1992), Gestural reorganization in connected speech: a functional viewpoint on 'articulatory phonology'. Phonetica 49, 205-211.

Kohler, K.J. (1994), Lexica of the Kiel PHONDAT Corpus Read Speech, Kiel [= Arbeitsberichte Institut für Phonetik und digitale Sprachverarbeitung Universität Kiel (AIPUK) 27 & 28].

Kohler, K.J. (1996), Articulatory reduction in German spontaneous speech. In: Proceedings of the 1st ESCA Tutorial and Research Workshop on Speech Production Modelling & 4th Speech Production Seminar. Autrans, 1-4.

Kröger, B.J. (1993), A gestural production model and its application to reduction in German. Phonetica 50, 213-233.

Pompino-Marschall, B. (1992), PHONDAT. Verbundvorhaben zum Aufbau einer Sprachsignaldatenbank für gesprochenes Deutsch. In: Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM) 30, 99-128.

Tillmann, H.G. (1995), Kleine und Große Phonetik. Phonetica 52, 144-159.

Wells, J.; Barry, W.; Grice, M.; Fourcin, A. & Gibbon, D. (1992), Stage report SEn.3 Standard computer compatible transcription (SAM-UCL-037). In: ESPRIT project 2589 (SAM) Multilingual Speech Input/ Output Assessment, Methodology and Standardisation. Final Report.

Wesenick, M.-B. (1994), Entwurf eines überspezifizierenden Regelsystems der Aussprache des Deutschen als Basis für empirische Untersuchungen (M.A. thesis, München).

APPENDIX

Orthographic text and canonical pronunciation (in SAMPA[8]) of the selected sentences of the Kiel corpus

e068:
Wann geht bitte *ein* Zug nach Hamburg mit möglichst wenig, ja, Bahnhöfen?
v a n+ g 'e: t b 'I t @ Q aI n+ t s 'u: k n a: x+ h 'a m b U 6 k m I t+ m '2: k l I C s t v 'e: n I C , j a:+ , b 'a: n #h "2: f @ n ?

e083:
Was ist die billigste Möglichkeit für *ein* Bahnticket von hier nach Kiel?
v a s+ Q I s t+ d i:+ b 'I l I C s t @ m '2: k l I C k aI t f y: 6+ Q aI n+ b 'a: n #t "I k I t f O n+ h 'i: 6 n a: x+ k 'i: l ?

e097:
Wann fährt bitte *ein* Zug von Regensburg nach Dortmund am Montag, am Montagmorgen?
v a n+ f 'E: 6 t b 'I t @ Q aI n+ t s 'u: k f O n+ r 'e: g @ n s b U 6 k n a: x+ d 'O 6 t m U n t Q a m+ m 'o: n t a: k , Q a m+ m 'o: n t a: k m 'O 6 g @ n ?

e076:
Gibt es *einen* Zug, der in Köln länger hält und nach Dortmund weiterfährt?
g 'i: p t Q E s+ Q aI n @ n+ t s 'u: k , d e: 6+ Q I n+ k '9 l n l 'E N 6 h 'E l t Q U n t+ n a: x+ d 'O 6 t m U n t v 'aI t 6 #f "E: 6 t ?

s055:
Gibt es *einen* Nachtzug, der am dreiundzwanzigsten Zwölften morgens in Ulm ankommt?
g 'i: p t Q E s+ Q aI n @ n+ n 'a x t #t s "u: k , d e: 6+ Q a m+ d r 'aI U n t #t s v "a n t s I C s t @ n t s v '9 l f t @ n m 'O 6 g @ n s Q I n+ Q 'U l m Q 'a n #k "O m t ?

s063:
Gibt es *einen* Zug, der am dreiundzwanzigsten Zwölften um zirka zwanzig Uhr in Oldenburg ist?
g 'i: p t Q E s+ Q aI n @ n+ t s 'u: k , d e: 6+ Q a m+ d r 'aI U n t #t s v "a n t s I C s t @ n t s v '9 l f t @ n Q U m+ t s 'I 6 k a: t s v 'a n t s I C Q 'u: 6 Q I n+ Q 'O l d @ n b U 6 k Q I s t+ ?

[8] In the PHONDAT version of SAMPA the symbol + is added as a marker to the last segment of unstressed function words and the symbol # denotes compound boundaries.

Kinematic and dynamic analysis of syllable articulation
A pilot study on German syllables with tense and lax vowels

Bernd Pompino-Marschall, Peter M. Janker and Christine Mooshammer

Zentrum für Allgemeine Sprachwissenschaft

Jägerstr. 10/11
D-10117 Berlin
Germany

SUMMARY

The present study is aimed at (1) testing the possibility of extracting dynamic gestural parameters from measured kinematic movement data of lingual articulation, and (2) asking the question whether these dynamic prameters may help to resolve the problem of what is articulatorily underlying the so called 'syllable cut' ('Silbenschnitt') in German. The results of the analysis of articulatory data for a first native speaker of German and observations on the extractability of dynamic parameters are reported. The study will be continued by different re-analyses and data from additional subjects.

INTRODUCTION

German 'Silbenschnitt'

Whether the 'long' and 'short' vowels of German represent a quantity opposition (long vs. short) or a quality opposition (tense vs. lax) has baan an object of debate for years in the phonological literature. The phonetic fact remains that with the exception of [aː] vs. [a] and [ɛː] both phonetic features correlate in the German vowel system. Another tradition - usually attributed to Sievers (1901) but reaching back to mid 18th century (e.g. Klopstock) - regards this opposition to be a prosodic one, an opposition in 'syllable cut' (Sievers 1901) or 'close' vs. 'loose' contact (Jespersen 1904; Trubetzkoy 1939). 'Close contact'/'strong cut' was understood to mean that the following consonant cuts off the (lax) vowel before it reaches its sonority peak. In recent times arguments for the prosodic interpretation of this opposition were presented by Vennemann (cf. Vennemann 1991).

Being a basically perceptual phonetic category, the phenomenon of 'syllable cut' could not be reduced experimentally to any independently motivated parameter in the acoustic or articulatory domain than to vowel duration (cf. Fischer-Jørgensen & Jørgensen 1969). But in a recent study analysing lingual articulatory movement data Hoole et al. (1994) were able to show that articulatory differences between tense and lax vowels in German do not show up in the controlled movements (of vocalic opening/consonantal closing) itself but rather in the timing between these two movements. These results are further supported by the study of Kroos (1996). This speaks for a prosodic base of this opposition within the German vowel system, i.e. for a dependency of this phenomenon on the timing of successive articulatory gestures.

In the present study we want to (1) replicate some of the findings of Hoole et al. (1994) and (2) ask the question whether there are invariant relative articulatory timing constraints underlying this prosodic vowel contrast of German. The parameters of relative timing (cf. Fowler 1977) are not measurable in the kinematic movement data but are themselves relatively abstract parameters of a dynamic model supposed to underly the actually measured movement data. Therefore the next paragraph is devoted to the description of the dynamic gestural model used here to experimentally estimate the values of these parameters.

The force field approach to gestural modelling

As has been shown, the behaviour of articulatory movements can be described very effectively as the dynamics of their underlying gestures (cf. Saltzman & Munhall 1989, Browman & Goldstein 1986ff). The term gesture here is meant to denote "a member of a family of functionally equivalent articulatory movement patterns that are actively controlled with reference to a given speech-relevant goal" (Saltzman & Munhall 1989: 334).

In the literature gestures are normally described by second-order harmonic oscillators (mass-spring systems):

$$m\ddot{y} + b\dot{y} + k(y - y_{tg}) = 0$$

(where m denotes the mass, b the damping of the system, k the stiffness of the spring, y the momentaneous position, $\dot{y}$ the momentaneous velocity, $\ddot{y}$ the momentaneous acceleration, and y_{tg} the rest position of the mass). Thus, in Articulatory Phonology different articulatory trajectories can be modelled by the same underlying critically damped mass-spring system (with mass m set to 1) by only changing the stiffness parameter and the coordination between neighboring gestures need not to be expressed in absolute units of time[1] but only relatively in terms of gestural phase.[2]

Since gestures are not meant here to describe the movement of single muscles but functional goal directed articulatory behaviour, the mass-spring analogy (especially the term 'stiffness' that here has nothing to do with e.g. stiffness of a muscle) should better be avoided. For the case of critical damping the above formula can be expressed in terms of the eigenfrequency ($\omega/2\pi$ or the eigenperiod $T = 2\pi/\omega$) of the undamped system:

$$\ddot{y} + 2\omega\dot{y} + \omega^2(y - y_{tg}) = 0$$

(with ω^2 denoting the force field per unit mass acting on the articulator position in direction of the target position; cf. Kröger et al. 1995: 1880).

Figure 1 depicts the general structure of the gestural model: in panel (a) the tongue dorsum movement along the main trajectory direction for successive /i:/ and /a/ gestures - along with the abstract (model derived) target positions (thick horizontal bars) is shown; panel (b) shows the assumed activation intervals (between the extrema of the displacement of the articulators) and the corresponding rectangular force function as used in Articulatory Phonology; in panel (c) the resulting force field (depending on articulator-target distance) for these /i:/ and /a/ gestures can be seen.

When applying this model to fit actually measured trajectory data Kröger et al. (1995) noticed that it seems to be too inflexible to fit the naturally occuring displacement, velocity and acceleration time functions. The rectangular force functions that are unrealistic on independent grounds (e.g. electromyographic (EMG) data of neuronal muscular control) are replaced in the model of Kröger (1996) and Kröger et al. (1995) by force functions that are characterized by a sinoidally rising onset interval, a steady-state phase of full activation (as well as a here neglected offset interval with sinoidally declining force; cf. figure 3 below).

---

[1] Nontheless the activation interval for a single gesture is normally specified in units of time.

[2] Especially this feature will be used in our analysis on the syllable cut prosody below.

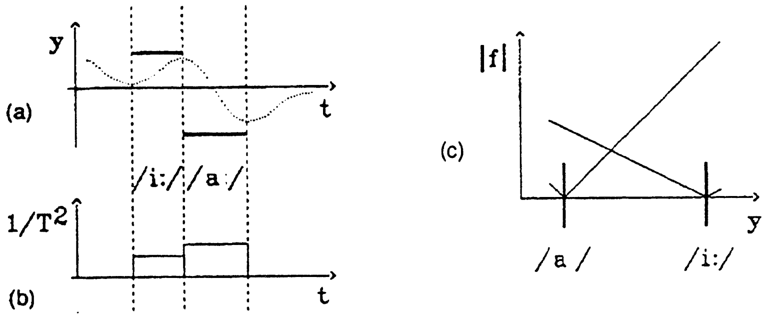


Fig. 1: Tongue body movement and gestural targets (bold horizontal bars) for successive /i:/ and /a/ gestures (a), rectangular force functions within the activation interval (b) and resulting force fields depending on distance from target position (c; thick vertical bars equalling horizontal ones in (a); after Kröger et al. 1995: 1880).

Below the formula is given for the force function in the onset interval (from t_{1on} to t_{2on}):

$$\omega(t) = \omega_0 \sin\left(\frac{2\pi(t - t_{1on})}{4(t_{2on} - t_{1on})}\right)$$

During the steady-state phase the force function remains constant

$$\omega(t) = \omega_0$$

The effect of this modification of the model is illustrated in figure 2 where the force field of a rectangular force function is compared (for clarity of exposure only) to a force field with a stepwise increasing force function. Thus, especially the unnaturally high acceleration peaks at gestural onset as found in the fits of the basic model could be avoided.

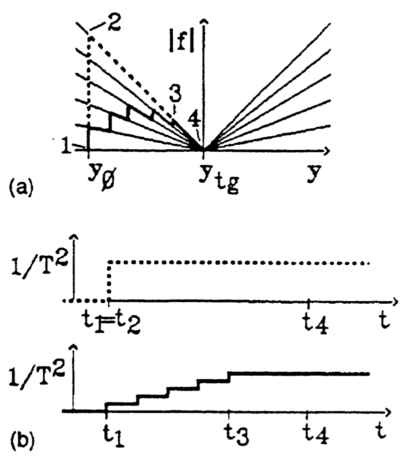


Fig. 2: Force fields (a) of different force functions (b): rectangular (dashed line), stepwise increasing (thick solid line; after Kröger et al. 1995: 1881).

This model is referred to as the six-parameter model in Kröger et al. (1995). The six gestural parameters are: (1) eigenperiod (T), (2) target position (y_{tg}), (3) beginning of the onset interval (t_{1on}, henceforth numbered [1]), end of the onset interval (t_{2on}; equal to the beginning of the steady-state phase, henceforth numbered [2]), beginning of the offset interval (and end of the steady-state phase, t_{1off}, henceforth [3]) and end of the offset interval (t_{2off}; not used here). Figure 3 shows an example fit for successive /i:/ and /a/ gestures.

A multidimensional minimization algorithm - the downhill simplex method (cf. Press et al. 1992: 408ff) - is used for adjusting the gesture parameters to fit the model's displacement function to the measured one within the fitting interval (from [1] to [3]) of N sample points:[3]

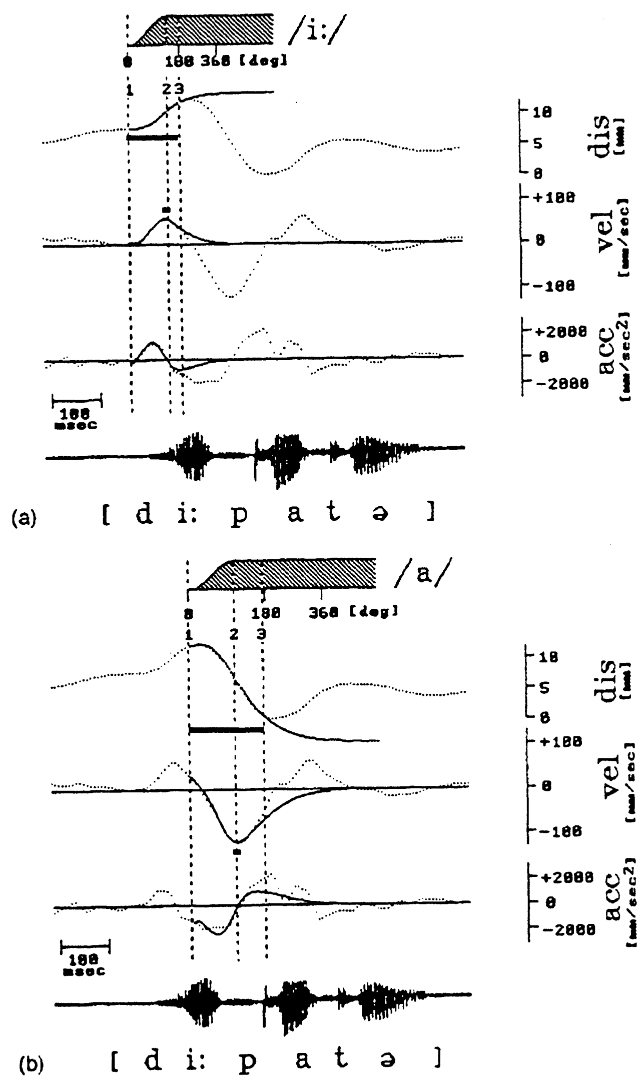


Fig. 3: Fit of successive /i:/ (a) and /a/ gestures (b) by using a continuous force function (1: beginning of the onset interval; 2: end of the onset interval; 3: end of the steady-state phase (= end of fit interval); thick horizontal bars at displacement and velocity functions show interval of minimization (see text for details); after Kröger et al. 1995: 1882).

[3] As well as the velocity function within a 10 ms window around the point of maximal velocity.

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i-y\,(t_i,a_1,\dots,a_M)}{s_i}\right)^2 \rightarrow min$$

(where M are the six model parameters, y_i the measured data points (from 1 to N), s_i the standard deviation of the measured data and $y(t_i,a_1 \dots a_M)$ the computed data).

The fitting proceedes in two steps from the starting values as depicted in figure 4. These starting values are defined as follows: (1) the end of the onset interval [2] is set to the point of maximal velocity (in figure 4 arrows labelled /iː/ and /a/), (2) the zero crossings of the velocity function surrounding this peak are used as a starting value for the beginning of the onset interval [1] and the beginning of the offset interval [3] (in figure 4 arrows labelled by roman numbers with index i and a respectively). Eigenperiod is initially set to twice the time interval between these points marking the extremal displacements since gestural offset roughly corresponds to a phase of 180° where 18% relative target distance is reached. The starting value for the target position (y_{tg}) is set the value of the maximal displacement.

The first fitting procedure then estimates three parameters simultaneously (the others held fixed at their initial values): target position (y_{tg}), eigenperiod (T) and beginning of the onset interval [1]. The second fitting procedure optimizes eigenperiod (T), the beginning of the onset interval [1], beginning of the steady-state phase [2] and the end of the steady-state phase/fitting interval [3] (the other parameters again remaining fixed).

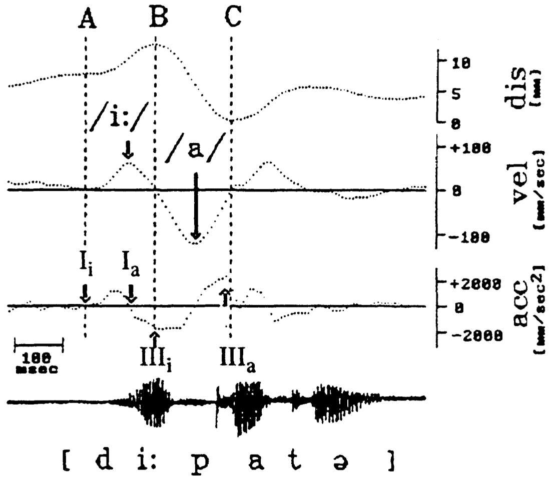


Fig. 4: Starting values of the fit algorithm for the successive /iː/ and /a/ gestures of fig. 3 (see text for details; after Kröger et al. 1995: 1883).

EXPERIMENTAL PROCEDURE

The general experimental setup is depicted in figure 5.

Material

To be able to control for a maximum of variables nonsense words were used as test utterances. Words containing CV or CVC syllables with tense and lax vowel nuclei in stressed position were constructed according to German phonotactics in the form [gə.C1V.C2ə] or [gə.C1VC2] (where C1 = C2: [t, z/s, l], V: [iː, ɪ, uː, ʊ, ɑː, ɑ], points marking syllable boundaries; e.g. [gəˈtɑːtə]/[gəˈtɑːt], [gəˈzʊsə]/[gəˈzʊs], [gəˈliːlə]/[gəˈliːl]). These words were embedded in the carrier sentences “Ich habe _ gesagt” (“I said _.”) or “Ich habe _ erwähnt” (“I mentioned _.”) respectively to hold the nearer context of the interesting articulations as stable as possible (i.e. [... ə _ ə ...] or [... ə _ ɛɐ ...]). The text of the test sentences were presented to the native German subject five times in randomized order via monitor.

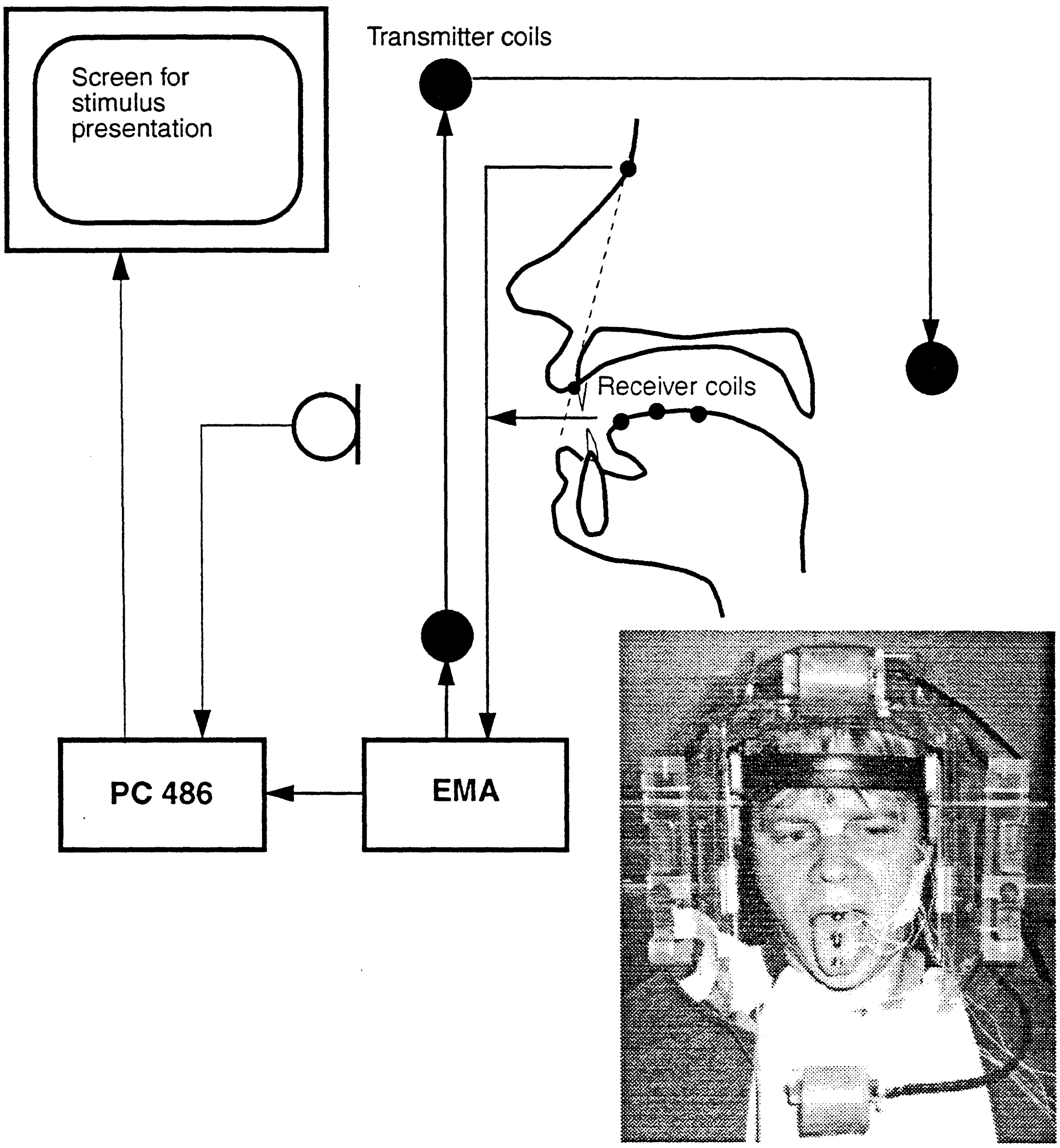


Fig. 5: Experimental setup and placement of receiver coils (bottom right: front view of the subject with tongue streched out to demonstrate the placement of the receiver coils).

Recording procedure

Tongue movements were monitored by means of electromagnetic articulography (AG100 Carstens Medizinelektronik, Göttingen, Germany). This method involves the use of three transmitter coils (mounted on a helmet) to generate an alternating magnetic field at three different frequencies. The field strength detected by sensor coils mounted on the articulators is roughly inversely proportional to the cube of the distance between sensor and transmitter (see Perkell et al. 1992, 1993; Schönle 1988 for background to electromagnetic transduction systems). The raw distance signals are then converted by software to x-y coordinates in the midsagittal plane. In order to guarantee the quality of the articulatory data, additional procedures were implemented allowing more accurate calibration and better detection of unreliable data (see Hoole 1993 for details).

Details of the sensor positions are as follows: Three transducers were mounted on the midline of the tongue from about 1-5 cm from the tongue tip. Two reference coils were attached to upper incisors and the bridge of the nose to correct for head movements.

The modified recording software (Hoole 1993) stored the movement data of the five receiver coils (recorded at 400 Hz) together with the information of the instantaneous tilt and the synchronous audio signal (16 bit, 16 kHz) in compressed form.

Besides the articulatory data at the end of the test session a tracing of the hard palate of the subject was made by using a sensor attached to the finger of one of the investigators.

The raw data were preprocessed to (1) correct for the remaining measurement error[4], (2) rotate to the vertical axis defined by the positions of the coils at the bridge of the nose and the upper incisors, (3) decompress the audio file, and (4) splitting the tilt data from the position data. The audio data were further (1) transformed to Signalyze format for acoustical measurements and, parallel, reduced to 8 bit 8 kHz for the analysis software ARTIC (cf. Kröger 1993, 1996; Kröger et al. 1995) for kinematic and dynamic analysis.

Analysis procedures

Durational measurements in the acoustical signal were conducted with the Signalyze software for Apple Macintosh. In the acoustic signal the following durations/time points were determined manually under auditory and visual (especially sonagraphic) feedback: the duration of the preceding sentence frame "Ich habe" as well as the remaining sentence frame "gesagt"/"erwähnt", the release for the initial [g] of the testword (as reference point for the kinematic/dynamic analysis parameter for [z/s] and [l] gestures, the duration of the first unaccented syllable [gə] of the testword (in plosive context ending at the offset of the higher formants), the duration of the first consonant (in the case of [t] as separate time points of release and of voicing onset), the duration of the target vowel (in plosive context again deliminated by the offset of higher formants), duration of the second consonant (again for stops differentiated between release and voicing onset) and the duration of the following vocalic segment.

Kinematic and dynamic analyses of the vocalic and consonantal gesture of the testsyllable(s) were conducted with a modified version of the ARTIC software (Kröger 1993, 1996; Kröger et al. 1995).[5]

In order for a trajectory description along the lines of a dynamical model as described in the introductory section first of all the two dimensional movement data has to be reduced to a single (main movement) dimension. In the case of the analysis with ARTIC this is done by rotating the y-axis into the main direction of the movement. In a first step this was done for all different combinations of the three tense/lax vowels with the three consonants and for all three tongue coils separately. For the later analysis the mean value of this rotation was used in all repetitions of the vocalic and consonantal gestures of a given segmental composition type (e.g. for all [...

[4] By using the computed error during calibration.

[5] We thank Dr. Bernd Kröger (Institut für Phonetik der Universität Köln) for allowing us to use his software and for the cooperation in implementing and modifying it according to our needs.

tiːt ... , ... tɪt ...]) when measuring a single coil (so in the consonantal context [t] the rotation for the tongue tip coil was 50° for the /a/ vowels, 60° for /i/ and 70° for /u/).

Figure 6 demonstrates the rotation procedure used to determine the main articulatory movement to be fit by the minimization procedure described in the introductory paragraph.

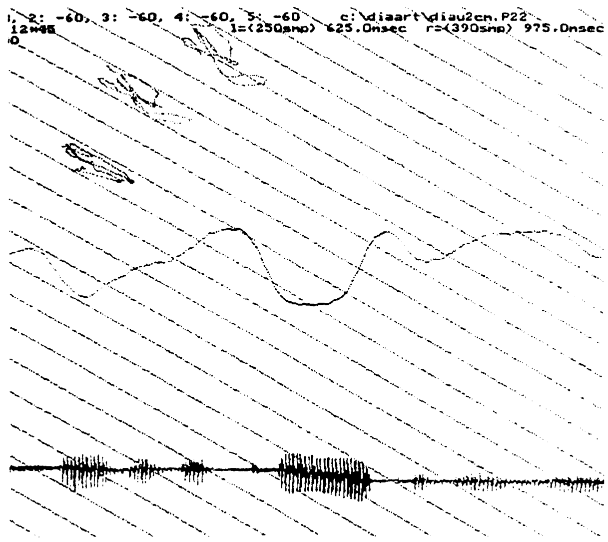


Fig. 6: Screen shot during the estimation of the angle of main articulatory movement direction: in the left upper corner the trajectories of the three tongue coils can be seen (leftmost: tongue tip; the highlighted part representing the CVC movement of [...taːt ...]), in the middle the movement of the tongue tip coil along the axis at an angle represented by the grid (60°) is shown, below the synchronous audio signal.

In a second step the algorithm determines the positions of the beginning and end of an articulatory movement around a chosen velocity maximum. For the kinematic measurements these automatically set cursor positions for the start and the end of the movement had to be manually corrected (especially in the case of tense vowels). Whereas the algorithm here strictly uses the zero crossings of the velocity function it seems reasonable to look for nearby local minima in order not to overestimate movement duration (cf. figure 7).[6]

Figure 7 shows a screen shot during the fitting procedure demonstrating this manual correction of the starting values of gestural on- and offset.[7] From these values the fitting algorithm of the model of time-varying force fields (Kröger 1996; cf. introductory section) starts its calculation.

Figure 8 shows the trajectories of the three tongue coils during the kinematically determined interval from the beginning of the vowel opening till the end of the consonantal closure of a single utterance of the test word [gəˈtaːtə]:

The tongue tip coil moves nearly parallel from the position of the [t] closure (N.B. the coil position not necessarily representing the actual contact point on the tongue surface) into a back position for the vowel and returning again into the alveolar closing position. The fit interval of the dynamic gestural model is marked by thicker trajectories in figure 8. As can easily be seen, during these controlled phases of tongue tip movement the other coils do not move synchronously in a controlled fashion. Whether these movements have to be regarded as simple co-ar-

[6] Hoole et al. (1994) here are using an arbitrarily set threshold value (20%) of displacement for delimiting movement duration. As has been shown by Kroos (1996) this results in more consistent data..

[7] The program ARTIC was modified as to record the values of the manually corrected movement start and end points that are to be compared to the model derived onset and offset points of the gesture.

ticulatory movements or as differently timed controlled movements of their own must remain an open question in this study.

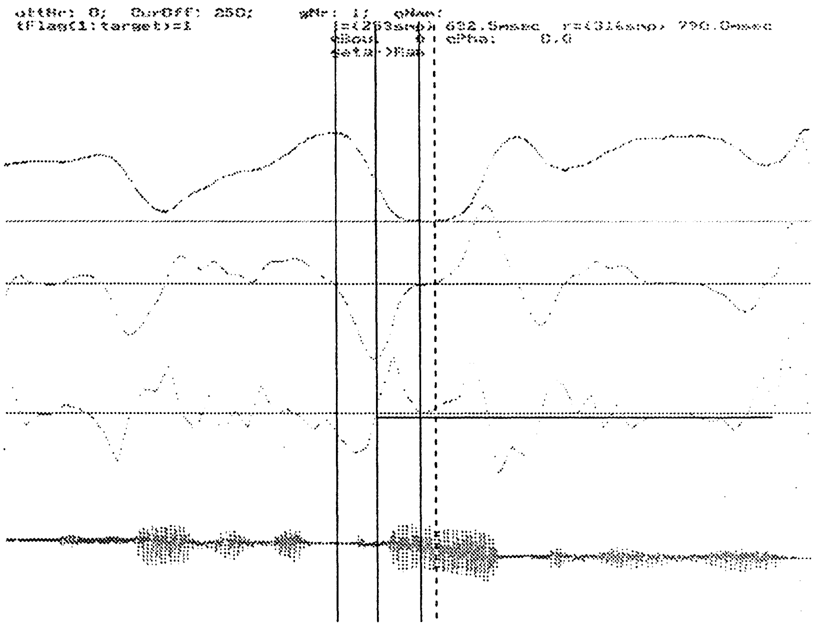


Fig. 7: Screen shot of the beginning of the fitting procedure (manual re-adjustment of the kinematic measurement points): the dashed line represents the automatically detected zero-crossing of the velocity function, the rightmost solid line the corrected gestural offset at the local minimum of the acceleration function (signal traces from top: displacement, velocity, acceleration, audio; markers from left to right: beginning of gesture, point of maximal velocity).

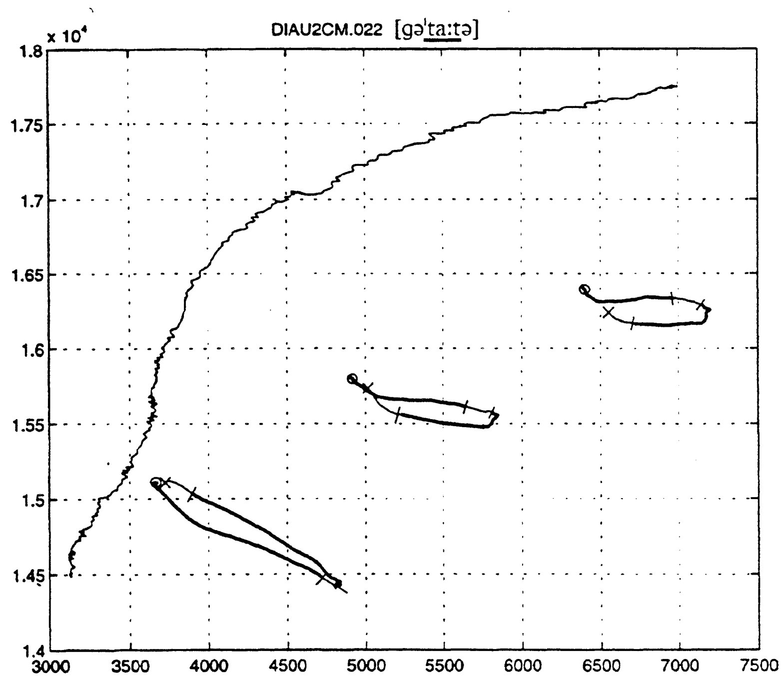


Fig. 8: CVC trajectories of the three tongue coils during the kinematically measured articulatory movement of the tongue tip coil; the thick parts (delimited by tick marks) show the fit interval of the gestural model; superimposed the contour of the hard palate is shown (vertical axis orientation between upper incisors and bridge of the nose; values in 10^{-2} mm distance from chin transmitter).

RESULTS

In the following paragraphs the results of the acoustic, kinematic and dynamic measurements are reported individually according to the following factors studied: tenseness of the vowel (2: tense, lax), vowel quality (3: /i/, /u/, /a/), syllable type (2: open, closed), consonant (3: [t], [z/s], [l]), and tongue point analysed (3: tip, front dorsum, back dorsum). The data reported here concentrate on the tongue tip behaviour as the main consonantal articulator in the [t] context since here the most stable results are to be expected.

Acoustic measurements

Here we want to report on only three different durational measurements of the acoustic signal. First of all vowel duration per se (measured from voicing onset to offset of the higher formants) is shown in the box plot of figure 9 (a) and given in table I for the different contexts (n = 5):

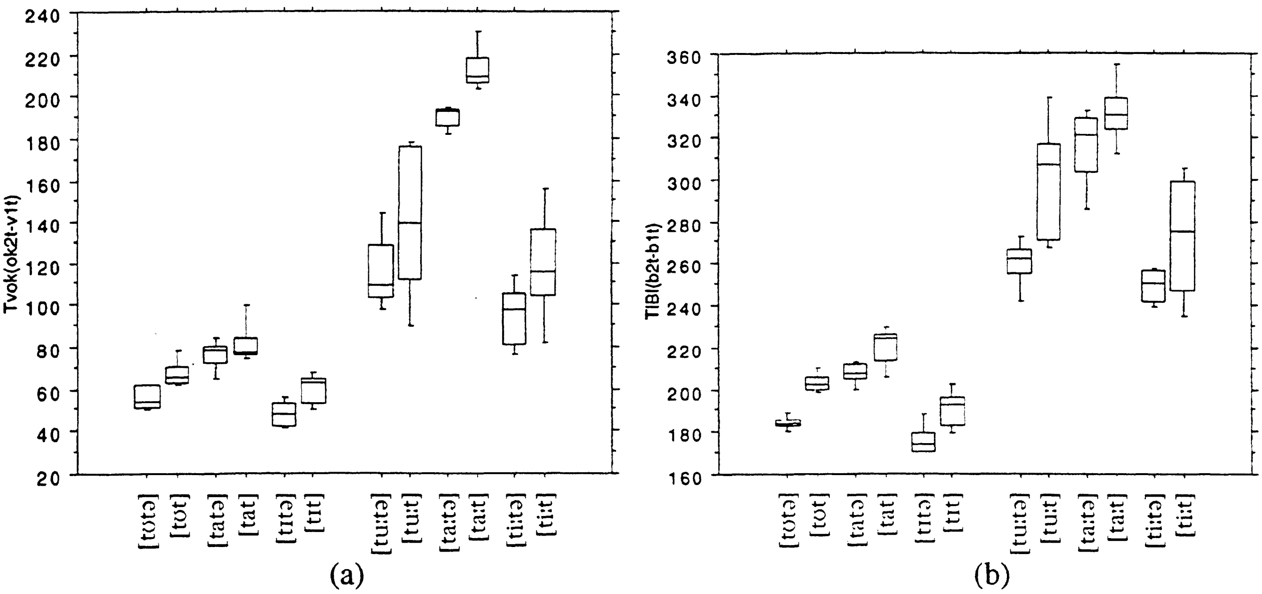


Fig. 9: Box plots of the voiced vowel duration (a) and duration of the burst-to-burst interval (b; in ms; boxes represent quartiles and median, bars the 10th and 90th percentile).

Table I: Acoustic vowel duration (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	55.663	5.462	[tu:tə]	115.862	18.350
[tʊt]	67.100	6.719	[tu:t]	140.300	37.167
[tatə]	76.275	7.111	[ta:tə]	189.787	5.191
[tat]	81.375	10.216	[ta:t]	212.375	10.904
[tɪtə]	47.900	6.303	[ti:tə]	94.450	15.329
[tɪt]	59.675	7.378	[ti:t]	119.050	26.840
mean	64.665	7.349		145.304	21.677

The data clearly show the expected effects of tenseness and vowel quality on duration: tense vowels being longer than lax ones and closed vowels intrinsically being shorter than open vowels. A parallel outcome is also seen in the durations of the burst-to-burst interval (including the voice onset time (VOT) of the first and the closure duration of the second [t]) as shown in figure 9 (b; cf. table II).

Table II: Duration of the burst-to-burst interval (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	184.100	2.990	[tu:tə]	260.025	11.370
[tʊt]	203.025	4.490	[tu:t]	299.037	29.389
[tatə]	207.700	5.222	[ta:tə]	315.625	18.759
[tat]	220.112	9.303	[ta:t]	331.725	15.007
[tɪtə]	176.150	7.285	[ti:tə]	249.275	8.180
[tɪt]	190.605	8.923	[ti:t]	272.662	29.747
mean	196.949	6.777		288.058	12.595

The different durational measurements taken (CVC duration, temporal distance between [ə_] offset and voiced [_ə] onset, between [t] burst and [_ə] onset, and the burst-to-burst interval correlate with one another as can be seen from figure 10:

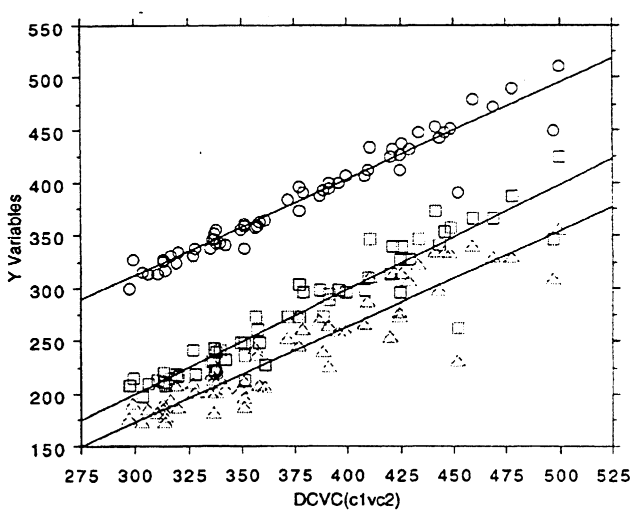


Fig. 10: The different durational measurements compared to the measured CVC duration: temporal distance between [ə_] offset and voiced [_ə] onset (circles; r = .970), between [t] burst and [_ə] onset (squares; r = .943), and burst-to-burst interval (triangles; r = .922).

Figure 11 and table III show thirdly the behaviour of voice onset time VOT for the prestressed [t] in the different contexts showing a lengthening in non-low tense vowels. This speaks in favour of the assumed missing quality difference in the case of the German /a/ vowels, all other vowels showing a lengthening of VOT with tenseness.

Table III: VOT duration (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	60.062	1.908	[tu:tə]	84.625	8.246
[tʊt]	69.650	2.365	[tu:t]	107.537	14.838
[tatə]	65.762	5.934	[ta:tə]	61.613	13.314
[tat]	66.875	9.263	[ta:t]	63.375	4.137
[tɪtə]	62.725	2.428	[ti:tə]	84.625	8.586
[tɪt]	63.942	4.994	[ti:t]	91.675	10.922
mean	64.836	5.181		82.242	10.611

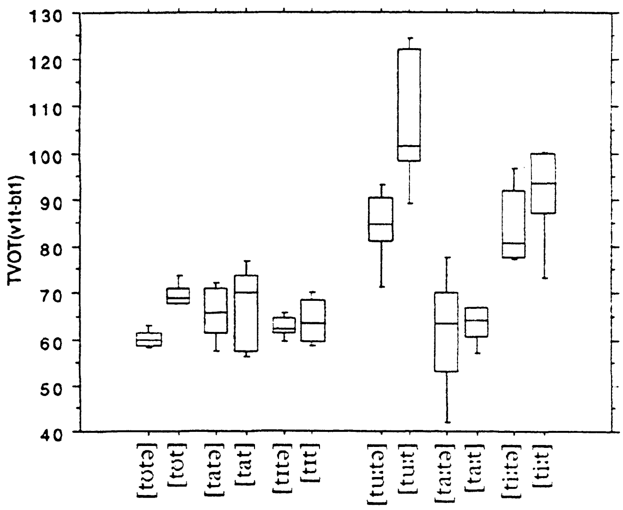


Fig. 11: Box plot of voice onset time durations (VOT in ms; boxes represent quartiles and median, bars the 10th and 90th percentile).

Kinematic measurements

The following figures and tables present the articulator displacement (im mm) and the movement duration (in ms) of the vocalic opening gesture and of the [t] closing gesture under the different context conditions.

Table IV: Articulator displacement during the vocalic opening gesture (in mm)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	-9.694	.263	[tu:tə]	-12.540	1.269
[tʊt]	-11.350	1.022	[tu:t]	-13.220	1.291
[tatə]	-9.710	.958	[ta:tə]	-12.510	.839
[tat]	-10.060	.570	[ta:t]	-12.962	.456
[tɪtə]	-4.780	.625	[ti:tə]	-3.636	.622
[tɪt]	-4.964	.282	[ti:t]	-3.738	.459
mean	-8.426	.686		-9.768	.893

Table V: Duration of the vocalic opening gesture (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	101.000	8.768	[tu:tə]	144.000	11.673
[tʊt]	112.500	5.863	[tu:t]	158.000	9.083
[tatə]	121.000	2.850	[ta:tə]	149.500	9.083
[tat]	131.000	13.532	[ta:t]	143.500	9.117
[tɪtə]	105.500	7.159	[ti:tə]	104.500	14.405
[tɪt]	111.000	2.236	[ti:t]	104.500	8.367
mean	113.667	7.732		134.000	10.503

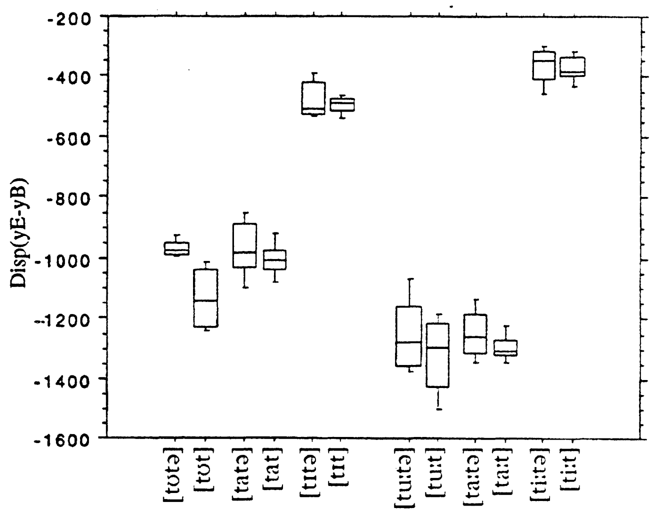


Fig. 12: Displacement (Disp(yE-yB)) of the tongue tip coil during the vocalic opening gesture (values are given in 10^{-2} mm).

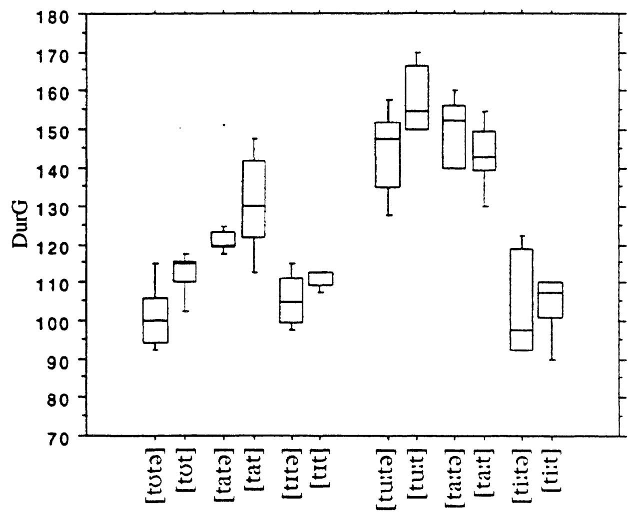


Fig. 13: Duration of the vocalic opening gesture (tongue tip coil; in ms).

Table VI: Articulator displacement during the [t] closing gesture (in mm)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	9.098	.425	[tu:tə]	12.634	1.547
[tʊt]	10.828	1.070	[tu:t]	13.056	1.397
[tatə]	9.536	.833	[ta:tə]	11.994	.447
[tat]	10.030	.553	[ta:t]	11.914	.994
[tɪtə]	4.270	.327	[ti:tə]	3.458	.434
[tɪt]	4.638	.202	[ti:t]	3.442	.336
mean	8.067	.642		9.416	.986

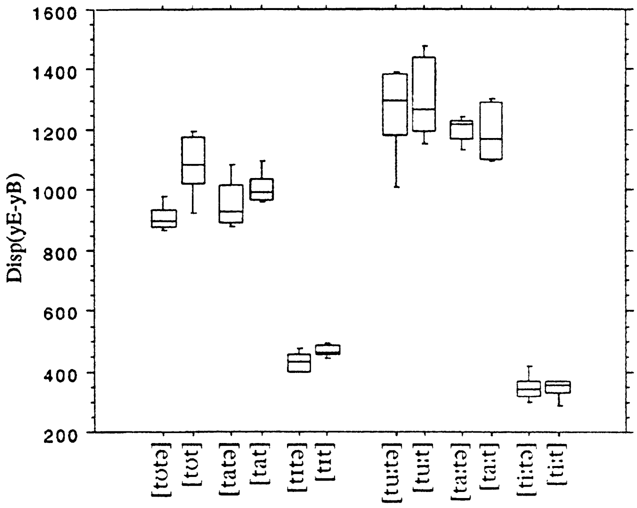


Fig. 14: Displacement (Disp(yE-yB)) of the tongue tip coil during the [t] closure gesture (values are given in 10^{-2} mm).

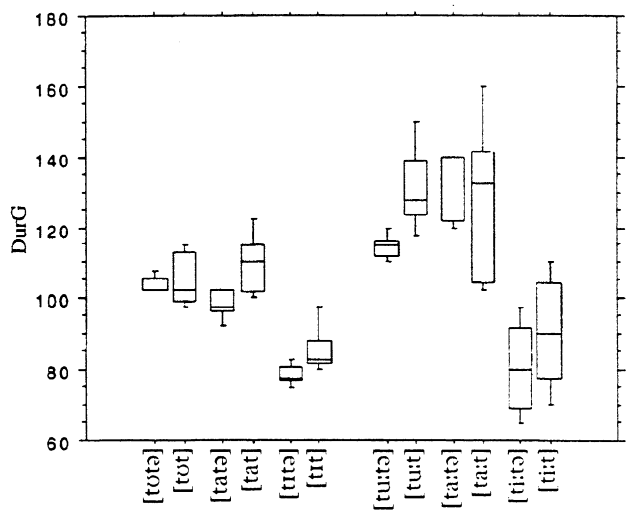


Fig. 15: Duration of the [t] closing gesture (tongue tip coil; in ms).

Table VII: Duration of the [t] closing gesture (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	141.000	2.236	[tu:tə]	114.500	3.708
[tʊt]	105.500	7.786	[tu:t]	131.100	12.265
[tatə]	98.500	4.183	[ta:tə]	132.500	10.308
[tat]	109.500	8.909	[ta:t]	127.000	23.809
[tɪtə]	78.500	2.850	[ti:tə]	80.500	13.509
[tɪt]	85.500	6.937	[ti:t]	90.500	16.240
mean	103.083	6.038		112.683	14.626

As can be seen the displacement of the tongue tip coil is of the same magnitude (but different in orientation) for the vocalic opening and the [t] closing gesture. For /i/ vowels there is less displacement in the case of the tense vowel reflecting its closed nature in the region measured by this coil. In contrast, for /a/ and /u/ the displacement is larger in the case of tense vowels reflecting the greater distance of these (in contrast to their lax counterparts) from the front position of the tongue tip coil at the alveolar closure.

The durations of the gestures on the other hand show a parallel behaviour to displacement: larger gestures are also longer. But this tendency is not so pronounced since it is compensated for by an often observed mechanism as shown in the following figures, i.e. that with larger gestures the maximal velocity of this gesture is risingas well. Expressed in terms of the mass-spring model of gestural dynamics this correlation represents the parameter of stiffness.

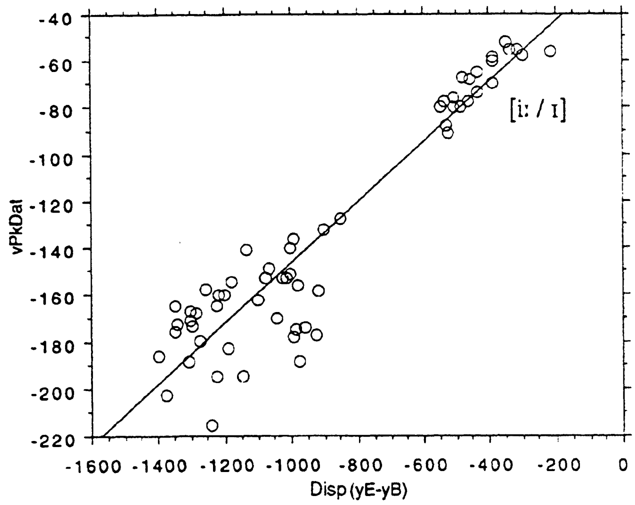


Fig. 16: Scatterplot of displacement (in 10^{-2} mm) vs. peak velocity (in 10^{-2} mm/s) demonstrating the concept of 'stiffness' (vPkDat = -17.784 + .128 * Disp(yE-yB); r = .944) for the vocalic gesture.

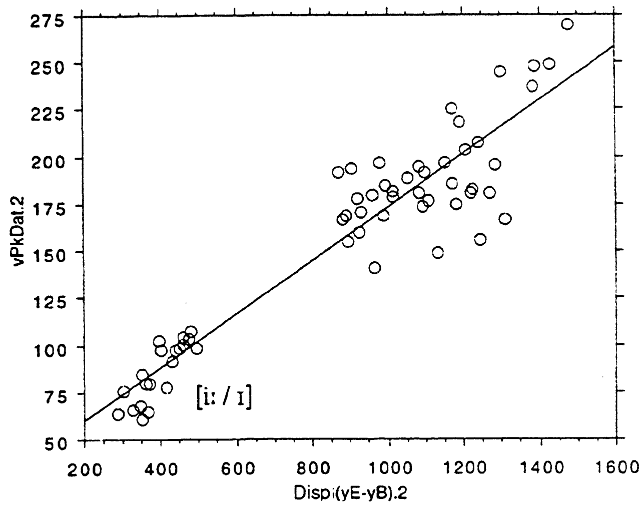


Fig. 17: Scatterplot of displacement (in 10^{-2} mm) vs. peak velocity (in 10^{-2} mm/s) demonstrating the concept of 'stiffness' (vPkDat.2 = 32.141 + .141 * Disp(yE-yB).2; r = .940) for the [t] gesture.

In these figures we can clearly see a dissociation between front and back vowels but both show a correlation between displacement and peak velocity that can be described by the same regression line. No differences for tense and lax vowels can be found.

To further test for possible differences between the gestures in the case of tense vs. lax vowels the next pair of figures and tables gives the parameter c for the velocity profile as proposed by Ostry & Munhall (1985):

(peak velocity / maximal displacement) * movement duration = c.

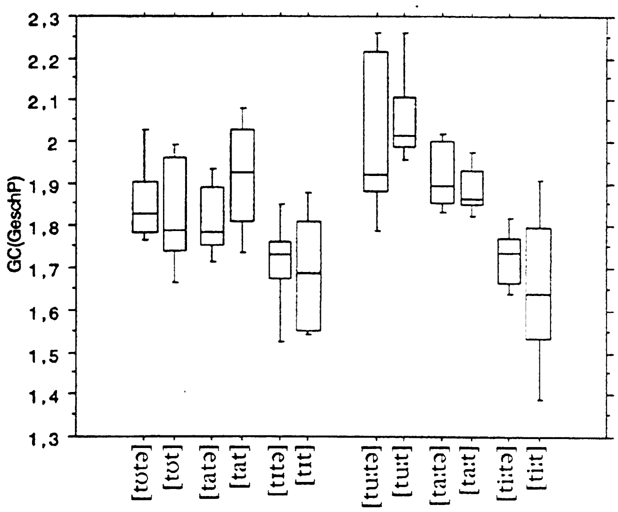


Fig. 18: Velocity profile parameter c for the vocalic opening gesture.

Table VIII: Velocity profile parameter c for the vocalic opening gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	1.855	.104	[tu:tə]	2.016	.203
[tʊt]	1.833	.134	[tu:t]	2.057	.119
[tatə]	1.816	.088	[ta:tə]	1.920	.083
[tat]	1.919	.136	[ta:t]	1.888	.059
[tɪtə]	1.713	.117	[ti:tə]	1.724	.070
[tɪt]	1.689	.145	[ti:t]	1.655	.195
mean	1.804	.122		1.877	.135

Table IX: Velocity profile parameter c for the [t] closing gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	2.125	.114	[tu:tə]	2.032	.091
[tʊt]	1.920	.163	[tu:t]	2.266	.319
[tatə]	1.777	.122	[ta:tə]	1.857	.129
[tat]	1.869	.074	[ta:t]	1.904	.232
[tɪtə]	1.829	.221	[ti:tə]	1.720	.111
[tɪt]	1.877	.070	[ti:t]	1.810	.189
mean	1.900	.138		1.932	.195

As can be seen there seem to be more intrinsic vowel differences in connection with this parameter than differences corresponding to the tense-lax opposition. This result suggests no differences with respect to this opposition in the articulatory movements themselves.

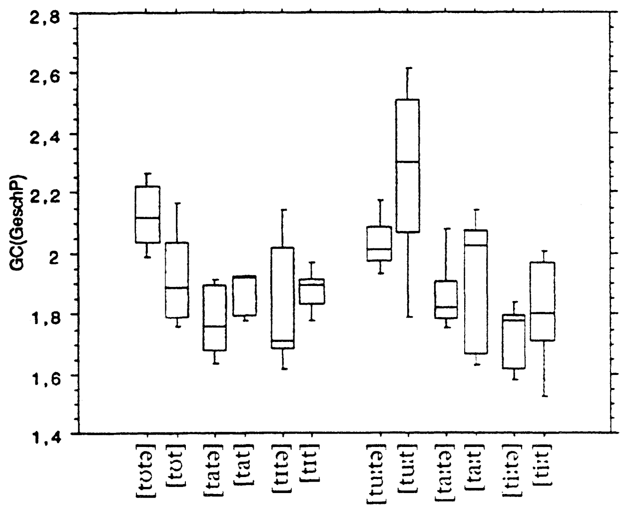


Fig. 19: Velocity profile parameter c for the [t] closing gesture.

As a last kinematic measurement we calculated the temporal distance between the end of the vocalic opening gesture and the beginning of the [t] closing gesture. In the test items with lax vowels these time points coincided with only one exception in the word [gəˈtɪtə] as can be seen in figure 20 and table X. In the test items with tense vowels on the other hand the onset of the closing gesture is delayed with respect to the offset of the opening gesture (cf. Hoole et al. 1994).

Table X: Temporal distance between the end of the vocalic opening movement and the beginning of the [t] closing movement (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	.000	.000	[tu:tə]	7.000	15.652
[tʊt]	.000	.000	[tu:t]	13.000	30.943
[tatə]	.000	.000	[ta:tə]	43.500	26.961
[tat]	.000	.000	[ta:t]	71.500	25.100
[tɪtə]	1.500	3.354	[ti:tə]	42.000	12.298
[tɪt]	.000	.000	[ti:t]	68.000	52.542
mean	.250	1.369		40.833	30.197

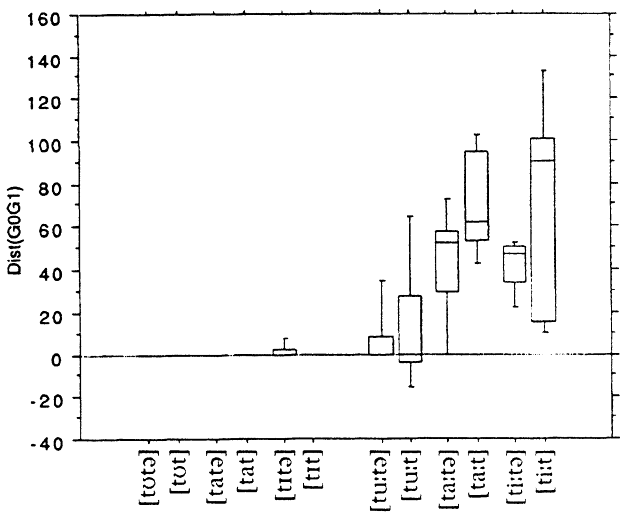


Fig. 20: Temporal distance between the end of the vocalic opening gesture and the beginning of the [t] closing gesture (in ms).

Dynamic measurements

Of special interest for this present study was the question of whether the dynamic parameters may help to fix differences in articulatory behaviour that are otherwise hardly definable.

The following figures and tables show the distances between the articulator position at the end of the fit interval and the abstract target position of the dynamic model (please note the different scaling in these figures: the variation shown in figure 21 represents one of the magnitude of the lax vowels in figure 22 left). The values of the target position per se are not reported here because they are not directly comparable since different measurement angles (cf. above and figures 25 - 28 below) lead to different values.

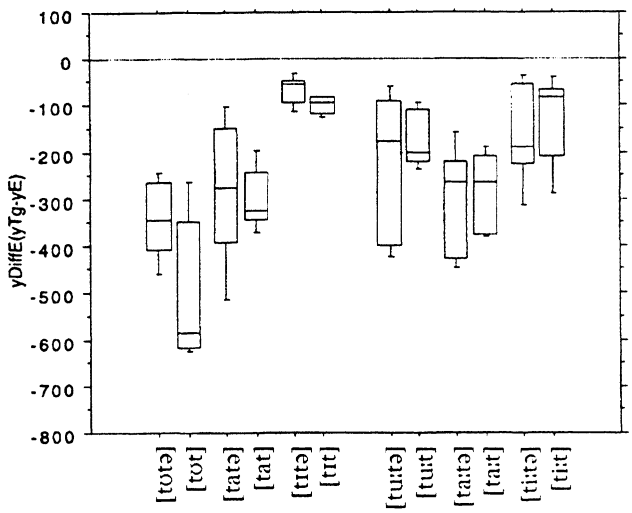


Fig. 21: Distance between articulator and target position at the end of the model's fit interval for the vocalic opening gesture (values are given in 10^{-2} mm).

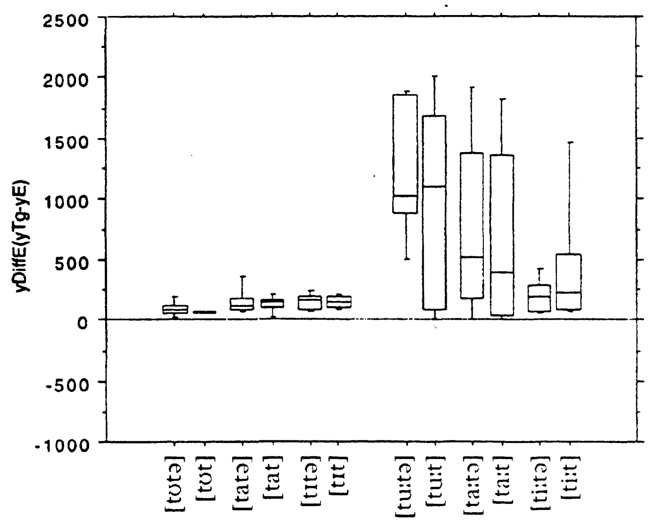


Fig. 22: Distance between articulator and target position at the end of the model's fit interval for the [t] closing gesture (values are given in 10^{-2} mm).

Table XI: Distance between articulator and target position (in mm) at the end of the model's fit interval for the vocalic opening gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	-3.426	.878	[tu:tə]	-2.310	1.666
[tʊt]	-4.938	1.637	[tu:t]	-1.732	.633
[tatə]	-2.814	1.610	[ta:tə]	-3.072	1.235
[tat]	-2.986	.684	[ta:t]	-2.848	.889
[tɪtə]	-.664	.331	[ti:tə]	-1.582	1.136
[tɪt]	-1.000	.198	[ti:t]	-1.342	1.005
mean	-2.638	1.332		-2.148	1.140

Table XII: Distance between articulator and target position (in mm) at the end of the model's fit interval for the [t] closing gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	.850	.626	[tu:tə]	12.518	5.986
[tʊt]	.576	.107	[tu:t]	9.574	8.878
[tatə]	1.458	1.211	[ta:tə]	7.702	7.781
[tat]	1.246	.694	[ta:t]	6.908	7.975
[tɪtə]	1.458	.684	[ti:tə]	1.908	1.496
[tɪt]	1.390	.533	[ti:t]	4.168	5.898
mean	1.163	.719		7.130	6.780

Clearly here the consonantal closing gesture after tense vowels shows higher values and a much greater variability in comparison to the values in the lax vowel context (cf. also figures 31 - 34 below).

As the second model parameter the following figures and tables show the eigenperiod values (please note again the different scaling in the individual figures).

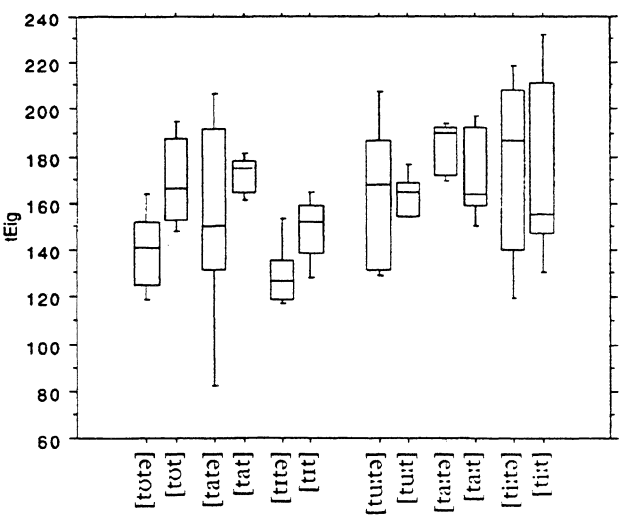


Fig. 23: Eigenperiod values (in ms) of the vowel opening gestures under the different contextual conditions.

Table XIII: Eigenperiod values (in ms) for the vocalic opening gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	139.620	17.682	[tu:tə]	163.180	33.258
[tʊt]	169.840	19.999	[tu:t]	163.380	9.159
[tatə]	154.780	47.577	[ta:tə]	183.700	11.755
[tat]	172.220	8.239	[ta:t]	172.880	19.870
[tɪtə]	129.220	14.330	[ti:tə]	175.380	40.995
[tɪt]	148.760	14.389	[ti:t]	175.020	41.574
mean	152.407	24.002		172.257	29.246

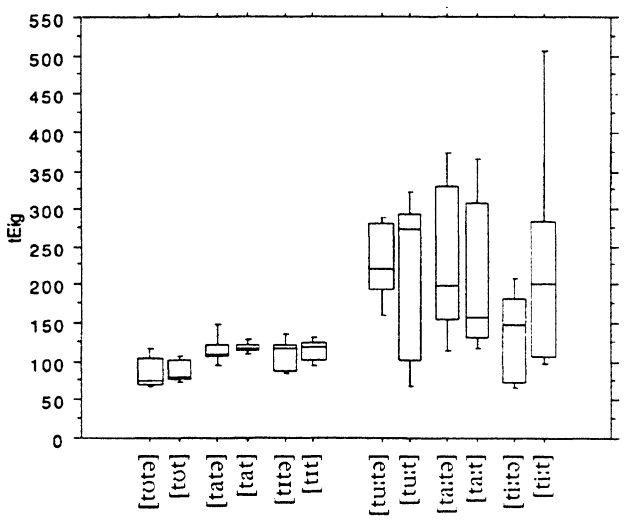


Fig. 24: Eigenperiod values (in ms) of the [t] closing gestures under the different contextual conditions.

Table XIV: Eigenperiod values (in ms) for the [t] closing gesture

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	86.340	20.975	[tu:tə]	230.760	53.097
[tʊt]	86.960	14.686	[tu:t]	212.200	113.228
[tatə]	115.000	19.736	[ta:tə]	234.160	107.304
[tat]	117.740	7.546	[ta:t]	212.880	108.656
[tɪtə]	108.180	20.621	[ti:tə]	133.760	61.717
[tɪt]	114.320	14.669	[ti:t]	224.740	165.476
mean	104.756	17.042		208.082	108.129

Again, we find higher values with much larger variation in the consonantal closing gesture in tense vowel context.

Since target position and eigenperiod duration are not independent from one another for the fitting algorithm of the model, in the following figures we looked for possibly occuring correlations between these two dynamic parameters.

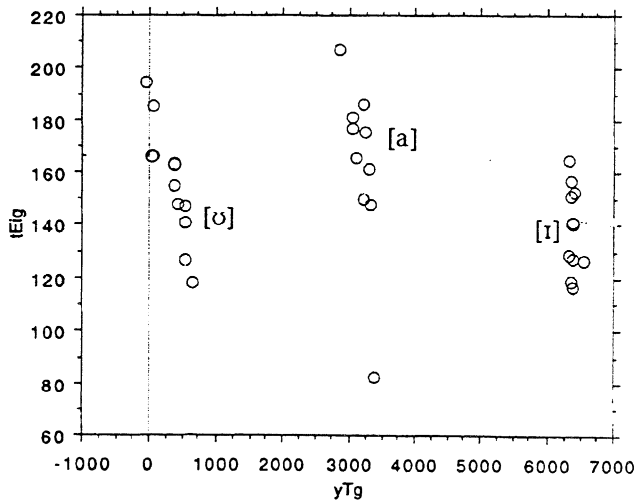


Fig. 25: Scatterplot of target position vs. eigenperiod value of the vocalic opening gesture for the lax vowels.

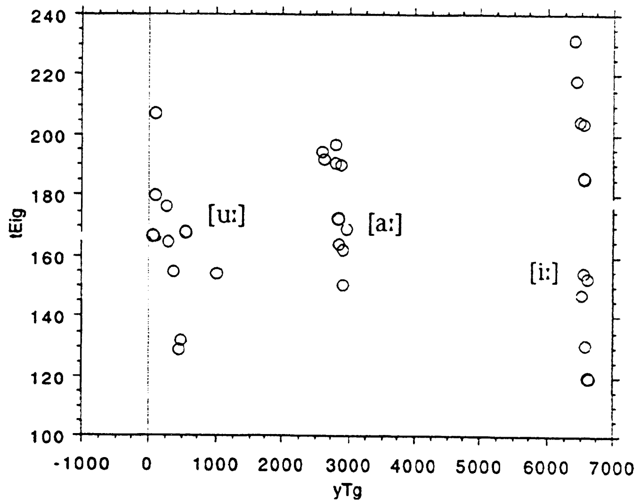


Fig. 26: Scatterplot of target position vs. eigenperiod value of the vocalic opening gesture for the tense vowels.

For the vocalic opening gesture no correlation between eigenperiod value and target position[8] can be detected (neither for tense nor for lax vowels).

Figure 27 and 28 show the same data for the consonantal closing gesture with the superimposed regression lines for the different vowel contexts.

[8] Please note that the differing values of the target position for the different vowels are mainly due to different angles of rotation (cf. above).

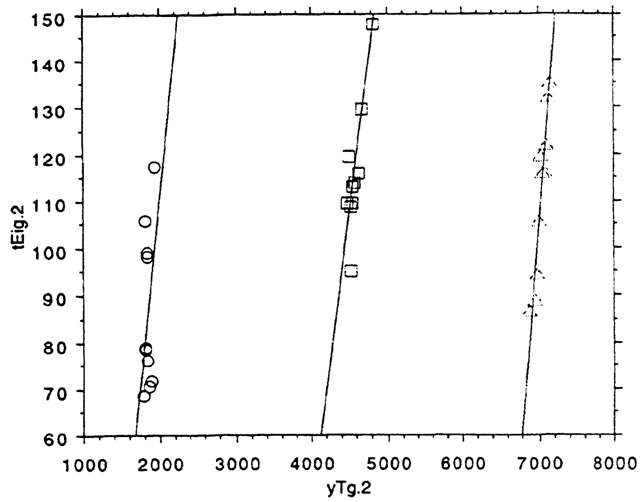


Fig. 27: Scatterplot of target position vs. eigenperiod value of the [t] closing gesture for the lax vowels with the calculated regression lines shown superimposed (circles: [ʊ], r = .397, n.s.; squares: [a], r = .849; triangles: [ɪ], r = .964).

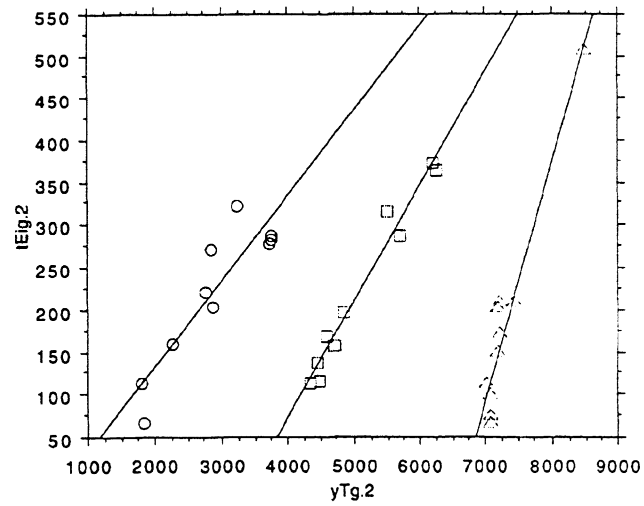


Fig. 28: Scatterplot of target position vs. eigenperiod value of the [t] closing gesture for the tense vowels with the calculated regression lines shown superimposed (circles: [u:], r = .901; squares: [a:], r =.984; triangles: [i:], r = .962).

Since there is a correlation between target position and eigenperiod value at least for the [t] closing gesture after tense vowels, the different results in the lax environment as well as in contrast to the vocalic opening gestures may be due to peculiarities of the fitting algorithm only. This question must be left unanswered here but will be studied further in more detail.

The following figures demonstrate the kinematic measurements as well as the model derived target positions. In these figures the trajectories during the measured gesture interval of the tongue tip coil are shown superimposed (starting points marked o, endpoints marked x) along with the contour of the palate and the target positions (crosses; and the mean target (large cross) for all five items). During the opening gesture for [ɪ] measured at the tongue tip coil in figure 29 one can see that the other coils do not show controlled movement in parallel but rather in the opposite direction.[9] As the normal case for the vocalic opening gesture the target positions do not scatter much and are quite close to the end point of the movement.

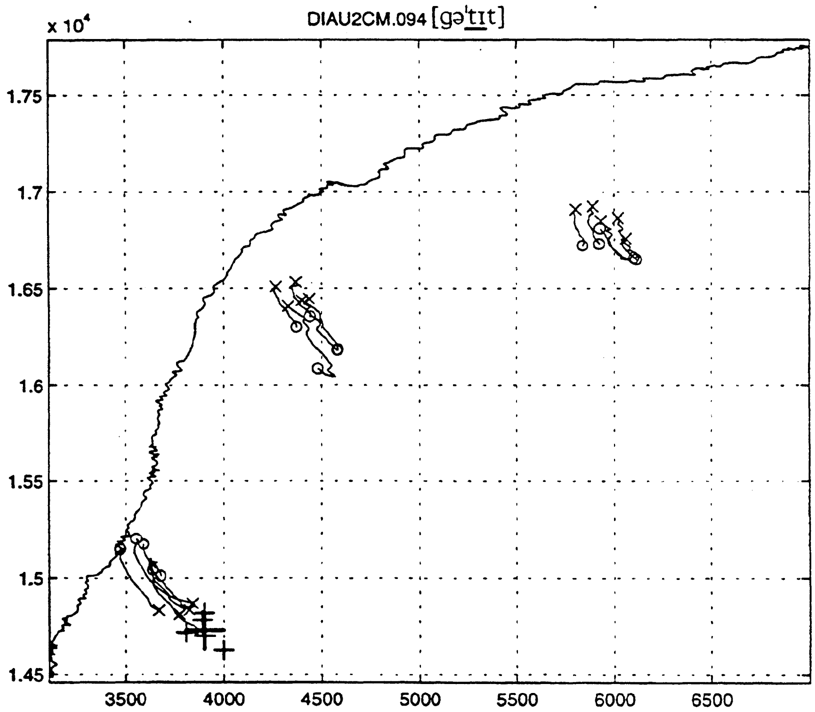


Fig. 29: Articulatory trajectories of the vocalic opening gestures in [gə'tɪt] as delimited in the kinematic analysis of the tongue tip coil (o marking measured movement onset, x movement offset) superimposed along with the palate tracking and the model derived gestural target positions (crosses) and the mean target location (large cross).

In figure 30, which shows the [t] closing gestures after [iː] in the same format as in figure 29 we can see the same direction reversals with respect to the different coils. Here the model derived targets scatter slightly more, the mean lying beyond the border of the hard palate since it is not to be confused with the target position of an articulator but as the origin of the force field acting on this articulator.

Figure 31 and 32 demonstrate the differences observed for the opening vs. closing gesture in tense vowel context (here for [aː]): an opening gesture modelled by a minimally varying target location near the extreme displacement of the articulator and a closing gesture modelled by a largely varying target position of the moving articulator.

[9] In general, the front tongue dorsum and back dorsum coil are difficult to measure according to the criteria applied in this study, e.g. for the front dorsum coil there is no specific angle of rotation that can be used for the different contexts.

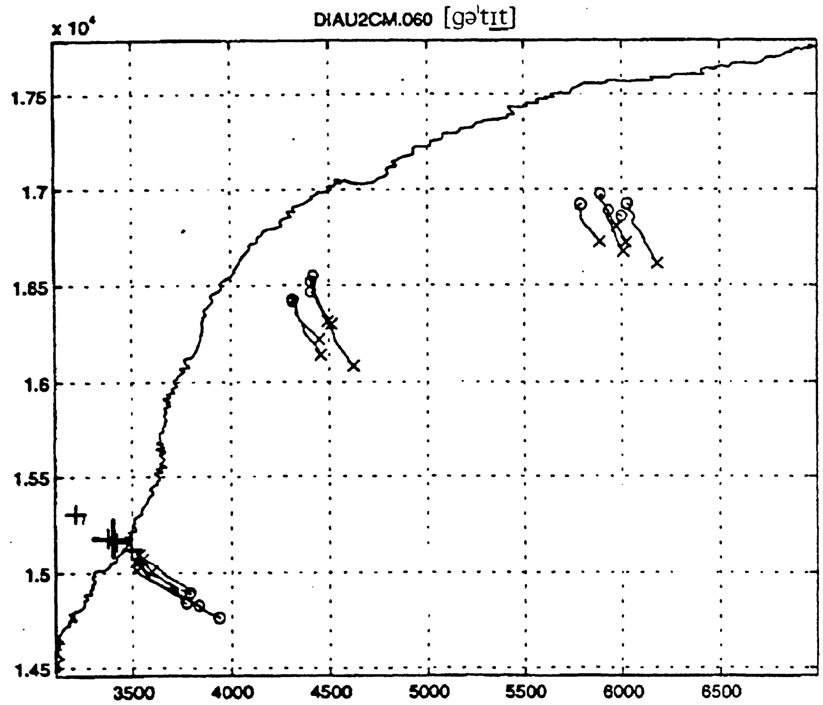


Fig. 30: Articulatory trajectories of the [t] closing gestures in [gə'tiːt] as delimited in the kinematic analysis of the tongue tip coil (o marking measured movement onset, x movement offset) superimposed along with the palate tracking and the model derived gestural target positions (crosses) and the mean target location (large cross).

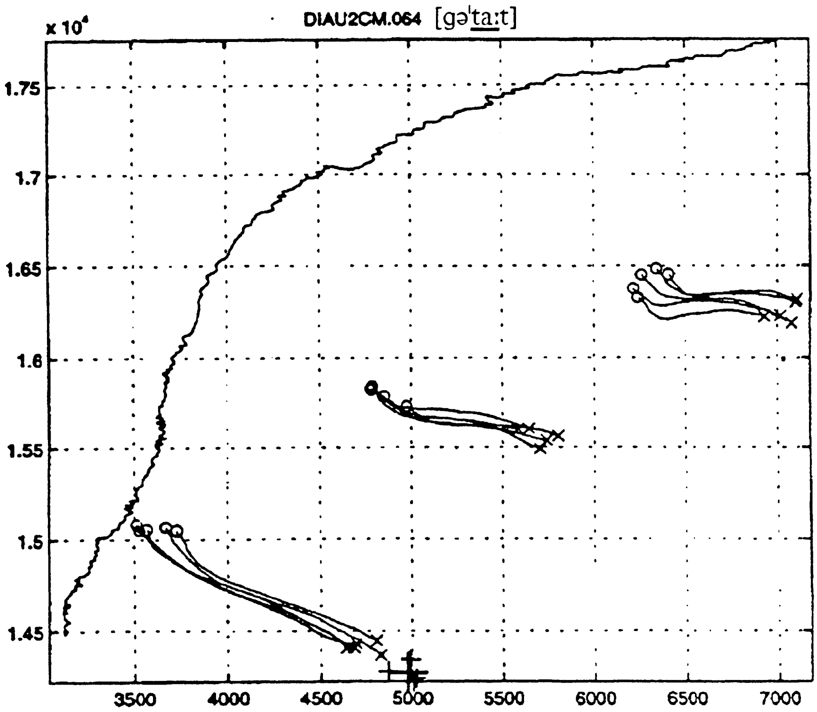


Fig. 31: Articulatory trajectories of the vocalic opening gestures in [gə'taːt] superimposed with the palate tracking and the model derived gestural target positions (crosses) and the mean target location (large cross).

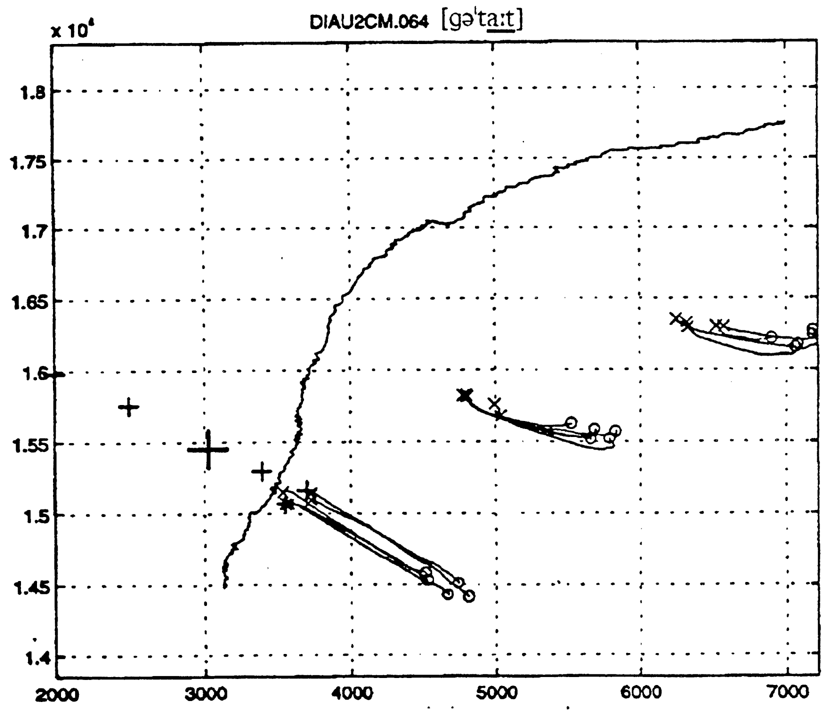
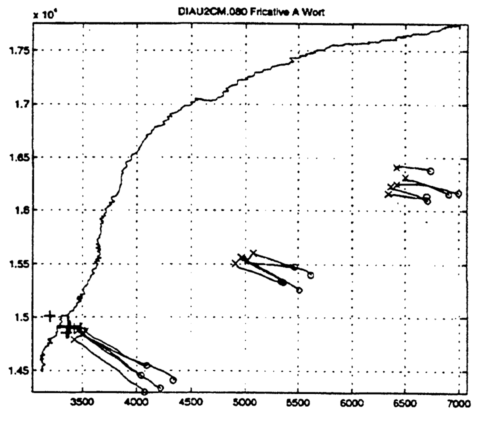


Fig. 32: Articulatory trajectories of the [t] closing gestures in [gə'ta:t] superimposed with the palate tracking and the model derived gestural target positions (crosses) and the mean target location (large cross).

In the figures 33 and 34 the parallel difference between the consonantal closing gestures when comparing preceding lax to tense vowels is seen for fricatives and laterals respectively.



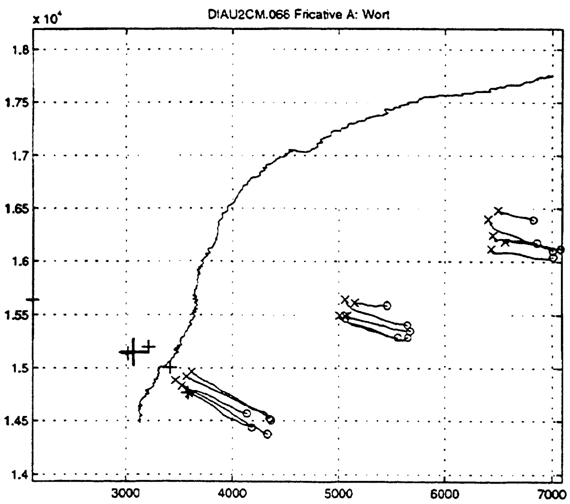


Fig. 33: Comparison of the trajectories and gestural target positions between lax [a] (left) and tense [a:] (right) vowel contexts for the fricative closing gesture.

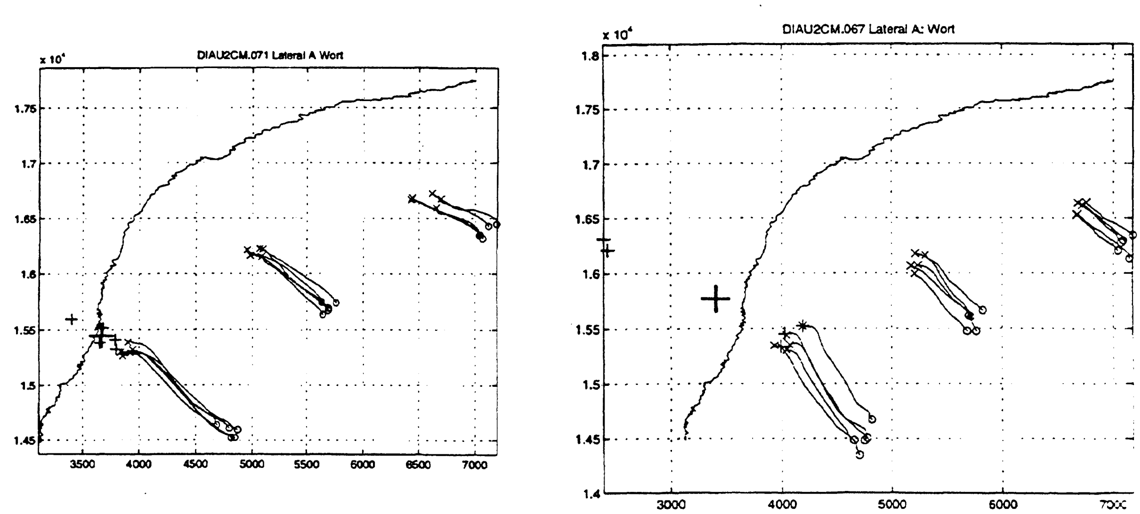


Fig. 34: Comparison of the trajectories and gestural target positions between lax [a] (left) and tense [a:] (right) vowel contexts for the lateral closing gesture.

Parallel to our kinematic analysis of the temporal distance between the vocalic opening and the consonantal closing movement we calculated the distance between the model derived gestures as shown for the plosive environment in figure 354 and table XV.

The same temporal distance, expressed in the relative timing measure of the preceding gesture's phase angle is shown in figure 36 and table XVI.

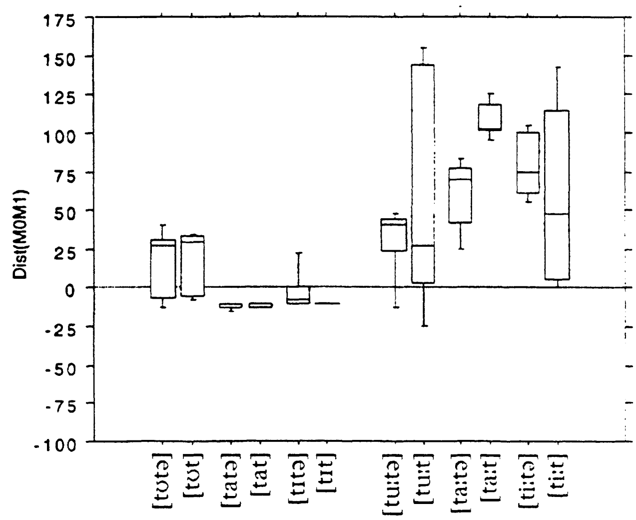


Fig. 35: Temporal distance between the end of the vocalic opening gesture and the beginning of the [t] closing gesture (in ms).

Table XV: Interval between the end of the vocalic opening gesture and the beginning of the [t] closing gesture (in ms)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	15.500	22.872	[tu:tə]	30.500	24.457
[tʊt]	17.000	21.316	[tu:t]	62.000	80.533
[tatə]	-11.500	2.236	[ta:tə]	60.000	23.519
[tat]	-11.500	1.369	[ta:t]	108.000	11.911
[tɪtə]	-2.500	14.031	[ti:tə]	79.000	21.694
[tɪt]	-10.000	.000	[ti:t]	60.500	61.932
mean	-.500	14.031		66.667	44.879

Table XVI: Interval between the end of the vocalic opening gesture and the beginning of the [t] closing gesture (relatively in degrees of the vocalic opening gesture)

Item	mean	std.dev.	Item	mean	std.dev.
[tʊtə]	239.420	77.063	[tu:tə]	259.788	52.827
[tʊt]	215.479	34.600	[tu:t]	351.979	126.115
[tatə]	178.251	23.749	[ta:tə]	307.386	31.433
[tat]	186.622	12.887	[ta:t]	396.111	64.613
[tɪtə]	237.212	30.436	[ti:tə]	276.924	35.052
[tɪt]	204.809	10.772	[ti:t]	269.524	56.941
mean	210.299	38.532		310.286	68.714

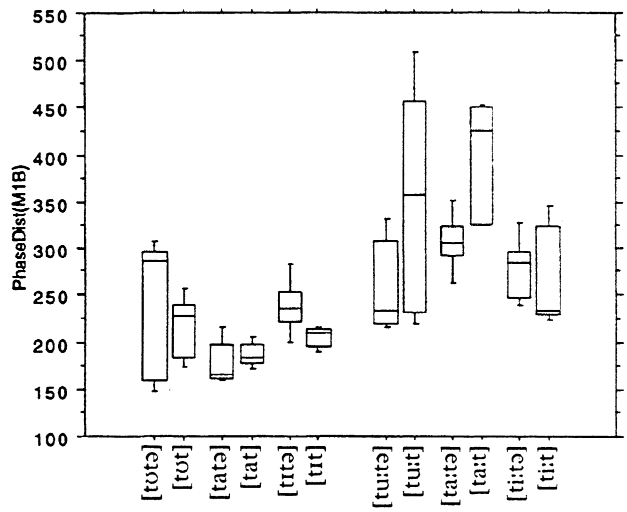


Fig. 36: Phase values (of the preceding vowel opening gesture) for the beginning of the fit interval for the [t] closing gesture.

DISCUSSION

A general problem encountered during this study is the proper delimination of the articulatory movement. As e.g. can be seen from figure 36 below the delimination of the articulatory movement in the kinematic analysis (yielding the starting values for the fit interval for the gestural model) is crucially dependent on the choice of the assumed main direction of articulatory movement. Clearly, in figure 37 phases of movements (as segmented under a rotation of 70°) are included that represent a sliding of the tongue along the hard palate that have to be excluded for the estimation of the gestural parameters of the controlled articulatory movement.

For the items of the [ʊt] gesture therefore a re-analysis was performed using a rotation angle (set to 58°) better fitting the main direction of articulatory movement. The results can be seen in figure 38.

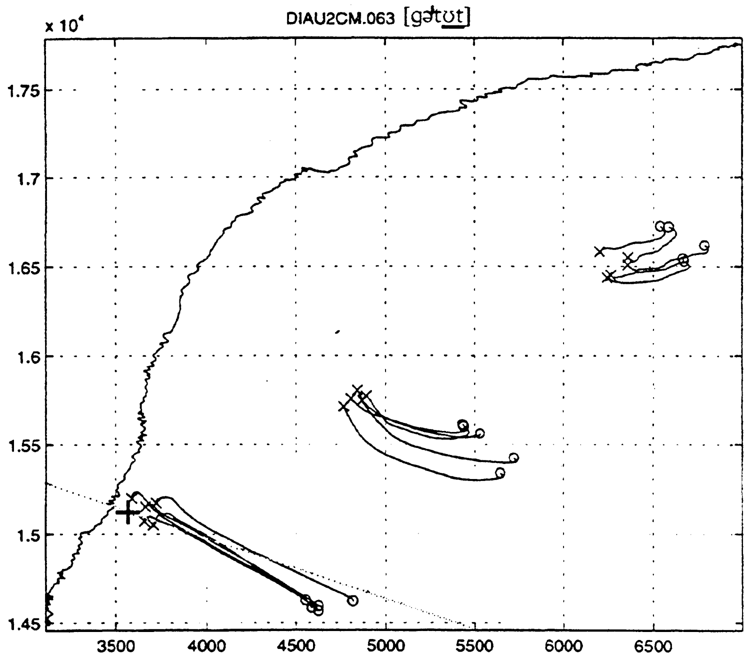


Fig. 37: Movement trajectories during the [t] closure phase measured at an angle of 70° (shown in gray): the bold cross marking the mean model derived target position.

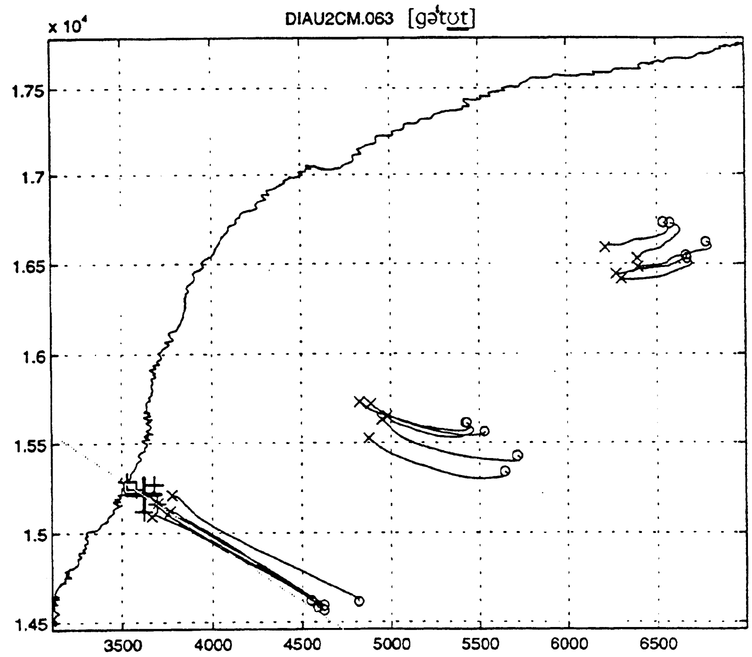


Fig. 38: Movement trajectories corresponding to the preceding figure during the [t] closure phase measured at the adjusted angle of 58° (shown in gray); crosses marking the model derived target positions.

For our future analyses (as well as for a re-analysis of the data of the present study) we therefore decided to measure the kinematic parameters of articulatory movement and the dynamic parameters as calculated by the fitting algorithm only at rotation angles appropriately adjusted for every individual articulatory movement. This will of course complicate the statistical comparison since the raw data needs a re-rotation to a mean value before. Furthermore the theoretical implications of rotating the data of different tongue coils, gestural directions and repetitions individually are not yet quite clear. In the moment this seems unproblematic for quite straight closing and opening gestures as for [t] in symmetrical vowel context.

These measurements along a main articulatory direction seem further restricted to the coil maximally near the relevant point of articulation. For the front tongue dorsum coil e.g. it was not possible to find a somewhat stable articulatory direction in the /u/ contexts for our present data.

But here for velar articulations that along these lines of reasoning have to be measured in the back dorsum coil a further problem arises since velar articulation normally proceeds in 'loops' that seem to be controlled (cf. Mooshammer et al. 1995) and not along a main direction. This behaviour leads to a permanently changing angle signal over time. Here again, the problem of delimiting the articulatory gesture correctly arises.

With respect to the question of the tense-lax opposition the present study in principle conforms to the results of Hoole et al. (1994): the timing coordination between vocalic opening and consonantal closing gestures seems to be the main parameter underlying this opposition. At the moment there is no clear indication whether there are differences between the kinematic and the different dynamic parameters (timing/phasing) with respect to this coordination.

The consonantal gesture following tense vowels seems less strictly controlled with respect to its timing/phasing as well as its inherent dynamic parameter values of target location and eigenfrequency. Regarding the latter parameters (target position and eigenfrequency) their model inherent correlation as it shows up in the [t] closing gestures (cf. figure 27 above) in tense vowel context remains an open question for further studies.

REFERENCES

Browman, C.P. (1991), Consonants and vowels: overlapping gestural organization. Proceedings of the XIIth International Congress of Phonetic Sciences, Vol. 1. Aix-en-Provence, 379-383.

Browman, C.P. & Goldstein, L.M. (1986), Towards an articulatory phonology. Phonology Yearbook 3, 219-252.

Browman, C.P. & Goldstein, L. M. (1989), Articulatory gestures as phonological units. Phonology 6, 201-251.

Browman, C.P. & Goldstein, L.M. (1990), Tiers in articulatory phonology, with some implications for casual speech. In: Kingston, J. & Beckman, M. (eds.), Papers in Laboratory Phonology: Between the Grammar and the Physics of Speech. Cambridge, 341-376.

Browman, C.P. & Goldstein, L. (1992a), Articulatory phonology: An overview. Phonetica, 49, 155-180.

Browman, C.P. & Goldstein, L. (1992b), Responses to commentaries. Phonetica, 49, 222-234.

Fischer-Jørgensen, E. & Jørgensen, H.P. (1969), Close and loose contact ("Anschluß") with special reference to North German. In: Annual Report of the Institute of Phonetics of the University of Copenhagen 4.

Fowler, C.A. (1977), Timing Control in Speech Production. Bloomington [PhD thesis]

Hoole, P. (1993), Instrumentelle Untersuchungen in der artikulatorischen Phonetik: Überlegungen zu ihrem Stellenwert als Grundlage für Entwicklung und Einsatz eines Systems zur Analyse der räumlichen und zeitlichen Strukturierung von Sprechbewegungen. München [phil.Diss.].

Hoole, P.; Mooshammer, C. & Tillmann, H.G. (1994), Kinematic analysis of vowel production in German. Proc. ICSLP 94, Yokohama, 1: 53-56.

Jespersen, O. (1904), Lehrbuch der Phonetik. Leipzig.

Kohler, K.J. (1990), Segmental reduction in connected speech in German: phonological facts and phonetic explanations. In: Hardcastle, W.J. & Marchal, A. (eds.), Speech Production and Speech Modelling. Dordrecht, 69-92.

Kohler, K.J. (1992), Gestural reorganization in connected speech: a functional viewpoint on 'articulatory phonology'. Phonetica 49, 205-211.

Kohler, K.J. (1994), Lexica of the Kiel PHONDAT Corpus Read Speech, Kiel [= Arbeitsberichte Institut für Phonetik und digitale Sprachverarbeitung Universität Kiel (AIPUK) 27 & 28].

Kohler, K.J. (1996), Articulatory reduction in German spontaneous speech. In: Proceedings of the 1st ESCA Tutorial and Research Workshop on Speech Production Modelling & 4th Speech Production Seminar. Autrans, 1-4.

Kröger, B.J. (1993), A gestural production model and its application to reduction in German. Phonetica 50, 213-233.

Kröger, B.J.; Schröder, G. & Opgen-Rhein, C. (1995), A gesture-based dynamic model describing articulatory movement data. Journal of the Acoustical Society of America 98, 1878-1889.

Kröger, B.J. (1996), Ein phonetisches Modell der Sprachproduktion. Köln [Habilschrift].

Kroos, C. (1996), Eingipflige und zweigipflige Vokale des Deutschen? Kinematische Analyse der Gespanntheitsopposition im Standarddeutschen. München. [M.A.-Arbeit].

Kroos, C.; Hoole, P.; Kühnert, B. & Tillmann, H.G. (1996), Phonetic evidence for the phonological status of the tense-lax distinction in German. Journal of the Acoustical Society of America 100, 2691 (A).

Mooshammer, C.; Hoole, P. & Kühnert, B. (1995), On loops. J. Phonetics 23, 3-21.

Ostry, D.J. & Munhall, K.G. (1985), Control of rate and duration of speech movements. Journal of the Acoustical Society of America 77, 640-647.

Perkell, J.S.; Cohen, M.; Svirsky, M.A.; Matthies, M.L.; Garabieta, I. & Jackson, M.T.T. (1992), Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. Journal of the Acoustical Society of America 92, 3078-3096.

Perkell, J.S.; Svirsky, M.A.; Matthies, M.L.& Manzella, J. (1993), On the use of electromagnetic midsagittal articulometer (EMMA) systems. In: Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM) 31, 29-42.

Pompino-Marschall, B. (1992), PHONDAT. Verbundvorhaben zum Aufbau einer Sprachsignaldatenbank für gesprochenes Deutsch. In: Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM) 30, 99-128.

Press, W.H.; Teukolsky, S.A.; Vetterling, W.T. & Flannery, B.P. (^{2}1992), Numerical Recipes in C. Cambridge.

Saltzman, E.L. & Munhall, K.G. (1989), A dynamical approach to gestural patterning in speech production. Ecological Psychology 1, 333-382.

Schönle, P. (1988), Elektromagnetische Artikulographie. Berlin.

Sendlmeier, W.F. (1985), Die Beschreibung der deutschen Vokale in betonter Stellung - ein forschungshistorischer Überblick. In: Borowsky, V. & Köster, J.-P. (eds.), Neue Tendenzen in der angewandten Phonetik. Hamburg.

Sievers, E. (^{5}1901), Grundzüge der Phonetik. Leipzig.

Trubetzkoy, N.S. (1939), Grundzüge der Phonologie. Göttingen ^{7}1989.

Vennemann, T. (1991), Syllable structure and syllable cut prosodies in modern Standard German. In: Certamen Phonologicum II. Papers from the 1990 Cortona Phonology Meeting. Turin, 211-243.