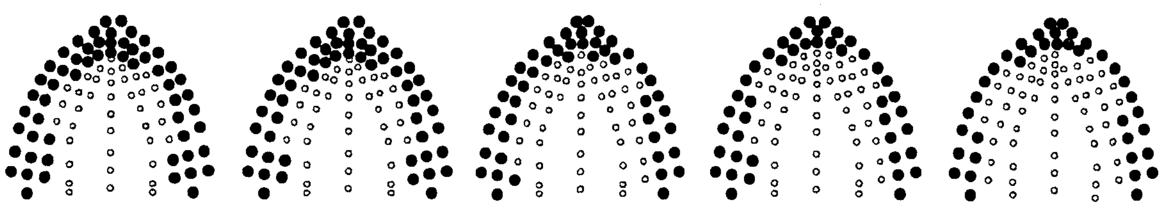
UCLA Working Papers in Phonetics



Ui IPi APi Wi Si

Number 97
July 1999

UCLA Working Papers in Phonetics

Number 97

July 1999

Table of Contents

Prosodically conditioned articulatory variations: A review

Cécile Fougeron
(translated from the French by Roger Billerey)

This review is a translation of most of Chapter I of Fougeron (1998)[*]. *It is included in this volume because it provides an extensive literature review on our theme. The paper which follows in this volume reports some of the results from Fougeron (1998).*

I. THE STRUCTURE OF THE LINGUISTIC MESSAGE

This section will illustrate how the structure of speech relates to the syntactic and prosodic structure of utterances. This will allow me to introduce the following notions: *constituent*, constituent *hierarchy*, and *prosodic position* within a constituent and in a hierarchy.

I.A. ISOMORPHIC SYNTACTIC AND PROSODIC STRUCTURES

Utterances[1] are comprised of words belonging to various syntactic classes (nouns, verbs, adjectives, etc.). The arrangement of these words follows precise syntactic rules, specified in the Grammar of the language. Words are organized into *syntactic constituents*, and syntactic constituents are organized into a hierarchical tree-like *syntactic structure*. For example, in the utterance "the little girl sings a song," the determiner, adjective and noun "the little girl" form a Noun Phrase[2], and the verb, determiner and noun "sings a song" form a Verb Phrase. The Noun Phrase and Verb Phrase then form a constituent at a higher level in the hierarchy: a Sentence.

The organization of lower-level units into higher-level units, their order, and their syntactic roles create the meaning of the utterance. Let us consider sentence (1), which is ambiguous:

(1) La belle ferme le voile
(*belle* = Adj or N: *ferme* = N or V, *le* = pronoun or determiner, voile = N or V)

Depending on the organization of the elements and on the hierarchical relations between the resulting syntactic constituents, this sentence can have two completely different meanings:

(1a) La belle ferme le voile: le voile est fermé par la belle
(the veil is closed by the beauty)
(1b) La belle ferme le voile: il est voilé par la belle ferme
(it is veiled by the beautiful farm)

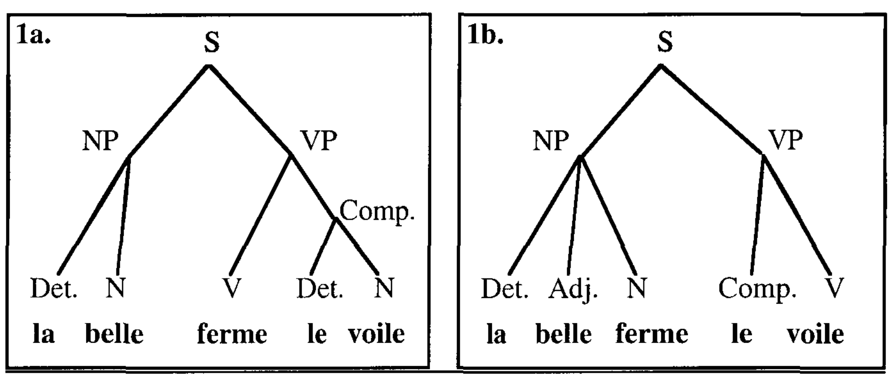
[*] Portions of the chapter referring to other parts of the dissertation have been omitted. In the translation, we have preserved original French linguistic examples, but have omitted discussion of French equivalents of technical terms.
[1] The word *utterance* here is used in its primary meaning: an instance of speech produced by a speaker W, in a place X at a time Y. This must not be confused with the Utterance, a prosodic constituent.
[2] Throughout this work, constituents standing in hierarchical relations in syntactic, prosodic or other structure will be noted with a capital letter (e.g. Syllable, Word, Noun Phrase, Accentual Phrase, etc.)

The syntactic structures that correspond to these two meanings are presented in figures 1.1a and 1.1b. In both cases, the Sentence contains a Noun Phrase and a Verb Phrase. In (1a), the Noun Phrase is "la belle" (the beauty) (determiner+noun) and the Verb Phrase is "ferme le voile" (closes the veil) (verb+complement), but in (1b), the Noun Phrase is "la belle ferme" (the beautiful farm) (determiner+adjective+noun) and the Verb Phrase is "le voile" (veils it) (complement+verb).

In parallel to its syntactic structure, a sentence is also associated to a *prosodic structure* when it is uttered. During speech[3], the speaker segments his/her utterance into *prosodic constituents* of different sizes and levels. In the acoustic domain, this segmentation is realized by particular intonation, accent and/or timing patterns, for example. These acoustic cues, among other things, will mark the *prosodic segmentation* of utterances into prosodic constituents, by grouping words together and marking boundaries between constituents.

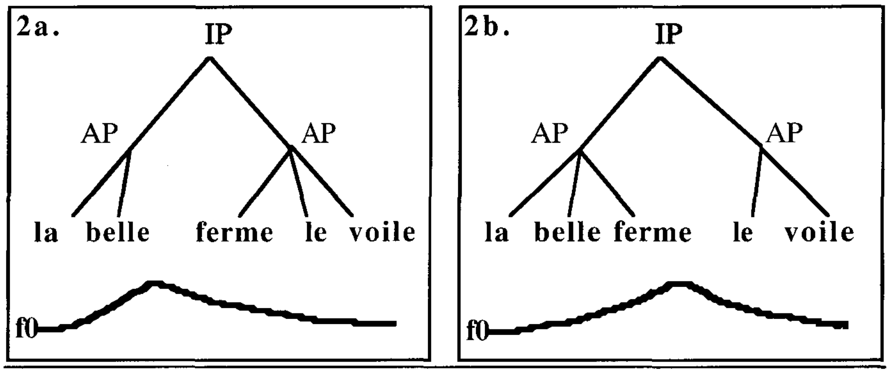
Let us go back to example (1). The written form of this sentence is ambiguous because the two possible syntactic structures are not apparent. However, when this sentence is produced by a speaker (uttered, represented mentally or read silently), the difference between meanings (1a) and (1b) will be reflected in the prosodic structure chosen by the speaker. The two possible prosodic structures are shown in figures 1.2a and 1.2b. These prosodic structures are realized acoustically by the two fundamental frequency (F0) contours represented schematically under the prosodic trees in figure 1.2. In (1a), the Noun Phrase "la belle" forms an Accentual Phrase marked by a minor continuation rise and medium final lengthening of the word "belle". In (1b), the adjective "belle" and the noun "ferme" form a single Accentual Phrase ("la belle ferme"). The two words are grouped in the same intonational contour, a minor continuation rise. The Accentual Phrase is also delimited by a medium final lengthening on the word "ferme."

Figures 1.1a and 1.1b: Syntactic structures of the ambiguous sentence "La belle ferme le voile" with meaning "le voile est fermé par la belle" in (1a) and "il est voilé par la belle ferme" (1b). (S=Sentence, NP=Noun Phrase, VP=Verb Phrase, Det.=Determiner, N=Noun, V=Verb, Comp.=Complement).



[3] In fact, it seems that prosodic structure does not require the actual utterance of a sentence. Even in silent reading or in linguistic thought processes, sentences are organized prosodically. A sentence would then only need to be "generated" or mentally represented by a speaker for it to have a prosodic structure.

Figures 1.2a and 1.2b: Possible prosodic structures of the ambiguous sentence "La belle ferme le voile" with meaning "le voile est fermé par la belle" in (2a) and "il est voilé par la belle ferme" (2b). The corresponding F0 contour is represented schematically under each prosodic tree. (IP=Intonational Phrase, AP=Accentual Phrase).



The meaning of an utterance thus depends on its prosodic organization. As Malmberg puts it,

unless one describes the prosodic structure of these two sentences, the description of their content will remain incomplete. [Malmberg 1971: 202].

Due to the prominent role given to the syntactic component in many linguistic theories, notably in the Chomskyan tradition, the definition of prosodic constituents has often been equated to that of syntactic constituents, or relegated to performance issues. However, renewed interest in spontaneous speech in the past few years has brought to the fore the notion that *prosodic structure is independent of syntactic structure*. The evidence accrued against the collapsing of prosodic structure into syntactic structure essentially lies in cases on non-isomorphism between the two structures.

I.B. NON-ISOMORPHIC SYNTACTIC AND PROSODIC STRUCTURES

In the previous examples [(1a) and (1b)], the prosodic structures shown in figure 1.2 mirror the sentences' syntactic structures. In these cases, prosodic and syntactic structures are *isomorphic* or *congruent*. It is generally accepted that the prosodic structure of an utterance is strongly influenced by its syntactic structure. However, prosodic and syntactic structure do not always match.

Before reviewing some of the factors influencing the prosodic segmentation of utterances but not their syntactic organization, let us consider two examples of non-isomorphism between the two structures.

Let us consider the syntactic structure of (2), from Nespor & Vogel (1986:57):

(2) This is the cat that caught the rat that stole the cheese
This is [the cat that caught [the rat that stole [the cheese]]]

Each bracket represents a syntactic NP node. This sentence, when uttered at a normal rate, can be produced with the following prosodic structure:

[This is the cat] [that caught the rat] [that stole the cheese]

Each bracket now represents an Intonational Phrase boundary. With this prosodic segmentation of the utterance into three Intonational Phrases, the double embedding of the relative clauses does not appear in the prosodic structure. The prosodic structure is "flatter" and does not reflect the complex syntactic structure of the sentence.

A second example of non-isomorphism between syntactic and prosodic structure is based on a fundamental difference between the two: a sentence has only one syntactic structure but may have several possible prosodic structures. Consider the following sentence:

(3) Le chat a mangé la souris (The cat ate the mouse)

This sentence is syntactically comprised of a Noun Phrase and a Verb Phrase, which itself is comprised of a Verb and its complement. This syntactic structure is fixed and it is the only possible structure. However, this sentence can be produced as two Intonational Phrases, as in 3a (the two Intonational Phrases are separated by a pause and the word "chat" is characterized by a major continuation rise in F0 and significant final lengthening) or as a single Intonational Phrase, as in 3b:

(3a) [Le chat] # [a mangé la souris] (# = pause)
(3b) [Le chat a mangé la souris]

The variability in the prosodic segmentation of an utterance is due to the fact that prosodic structure does not directly reflect syntactic structure. The prosodic segmentation of an utterance will depend on several linguistic and non-linguistic factors which do not influence the syntactic structure. Such factors include:

- The prosodic structure of a sentence depends on the *speaker* and on the *listener*.
- Prosodic structure depends on the *message*. Prosodic segmentation provides cues to the informational structure of the sentence. Emphasis given to certain elements, the theme/rheme distinction, the expression of certain attitudes or of the modality of the sentence, for example, are factors which will influence the speaker in the prosodic segmentation of the utterance. For instance, the prosodic segmentation in (3a) can be used if the speaker wants to emphasize the word "chat" in response to the question "Qui a mangé la souris?" (Who ate the mouse?): [LE CHAT] # [a mangé la souris].
- The prosodic structure of utterances also responds to general *rhythmic principles*. In particular, the length of prosodic constituents depends on the length of the sentence (number of words) and on the rhythmic weight of the elements (number of syllables). Several studies have shown that speakers tend to segment their utterances in accordance to considerations of symmetry and syllabic balance so that constituents at the same level are balanced, form a recurrent rhythmic pattern and/or follow stress alternation principles (eurhythmy) [Grosjean et al. 1979, Gee & Grosjean 1983, Dell 1984, Rossi 1985, Pasdeloup 1990, Delais 1995].
- The prosodic segmentation of utterances also depends on *speech rate*. When speaking fast, speakers tend to group more elements within the same prosodic constituents. Consequently,

constituents are heavier (they contain more elements) and the number of prosodic boundaries diminishes at faster speech rates [see Vaissière 1983, Lucci 1983, Fougeron & Jun 1995, 1998 for French; Jun 1993 for Korean; Casper & van Heuven 1991, 1995 for Dutch].

- The segmentation of the speech stream into prosodic constituents also depends on *speech style*. In read speech (as in this study), speakers tend to mark boundaries between morpho-syntactic constituents clearly and regularly. On the other hand, in spontaneous speech, the prosodic structure of utterances is mainly indicated by marking prominence. Prosodic structuring is more important than syntatctic structuring in spontaneous speech [Vaissière 1997, p.c., Lucci 1983].

These various factors affect the prosodic structure of an utterance, but not its syntactic structure, and thus make the two differ. Chomsky & Halle (1968) relegate the problems raised by non-isomorphism between syntactic and prosodic structure to performance issues. Factors that only influence prosodic structure are considered to be constraints on performance and as such not to enter in the Grammar of the language. However, as noted by Nespor & Vogel (1986: 57), the prosodic structure of sentence (2):

(2) [This is the cat] [that caught the rat] [that stole the cheese]

does not depend on performance, because the position of intonational boundaries is governed by strict rules (before each relative pronoun) and native speakers have strong intuitions about them. Prosodic segmentation is thus part of a speaker's competence and follows regular principles.

Consequently, it appears that the prosodic structures that can be taken on by an utterance cannot be correctly derived from syntactic structure alone. They are defined by linguistic (syntactic, semantic, etc.) and non-linguistic (rhythm, speech rate) information. This integrated nature of prosodic structure has led numerous authors to assume an autonomous prosodic component in the Grammar.

I.C. An autonomous prosodic component -- Prosodic theories

C.1. Introduction

A survey of the literature shows how difficult it is to define what prosody is. In fact, depending on an author's approach, this term can have multiple meanings. In short, prosody is defined:

- Either by its *physical realization* in the speech signal. Various definitions then refer to acoustic parameters such as duration, fundamental frequency (F0), amplitude, non-phonemic spectral variations, and pausing. They encompass intonation and stress phenomena.
- Or by its *utterance-structuring function*. Prosodic structure, and thus prosodic constituents, are defined as "domains" in which particular prosodic phenomena are realized. These phenomena are considered as prosodic because they do not refer to segments, but to higher-level constituents. In this sense, they are suprasegmental phenomena. Vowel harmony in Turkish, which applies beyond the word within a Clitic Group, is a good example. Lexical and post-lexical phonological phenomena are thus considered as prosodic.

Over the past twenty years, prosodic theories such as *prosodic phonology* have been proposed which redefine the position of prosody in Grammar. These theories introduce an autonomous level of prosodic representation at the morphosyntax/phonology interface [the best-known: Liberman 1975, Nespor & Vogel 1986, Selkirk 1986, Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988, Hayes 1989][4].

The introduction of an autonomous prosodic module is based on the fact that the organization of an utterance in constituents by a speaker is only imperfectly predicted by syntax (see examples in the previous section). Prosodic theories thus propose to represent an utterance as an independent hierarchical *prosodic structure*. This structure is organized in prosodic constituents defined on the basis of syntactic, morphological, semantic and rhythmic information. Thus prosodic structure refers to syntactic structure only indirectly.

The need for an autonomous prosodic structure also becomes apparent when one tries to account for phonological phenomena whose application is not determined by syntactic or morphological criteria. For example, the phenomenon of Raddoppiamento Sintattico in Italian consists of lengthening a word-initial consonant (a single consonant or a cluster without /s/) if the preceding word ends in a stressed vowel. The application of this rule does not correspond to any particular syntactic constituent. Nespor & Vogel (1986) show that the domain of application of this rule is the Phonological Phrase[5], i.e. the rule only applies to words belonging to the same Phonological Phrase and not across Phonological Phrases.

In addition, Nespor & Vogel also show that some rules apply in constituents higher than traditional syntactic constituents. For example, the linking-r rule in British English can apply within a sentence, but also between two sentences. In sequence (1) below, the rule applies between two sentences, even though they are not grouped by any syntactic constituent. In sequence (2), the [r] at the end of *mother* is elided, and thus the linking-r rule does not apply between these two sentences:

(1) There is my mothe[r]. I've got to go.
(2) There is my mothe[ ]. I've got two cats.

Syntax cannot account for the domain of application of the linking-r rule. One has to invoke semantic information governing the relationship between the two sentences: the rule only applies across phrases bearing a close semantic relationship to each other (in (1), a causal link). Nespor & Vogel propose a prosodic constituent as the domain of application of this rule: the Utterance.

These examples show that prosodic constituents form domains within which suprasegmental phenomena can be analyzed. Prosodic phonology is thus a *theory of domains* [Nespor & Vogel 1986]. Every prosodic constituent is the domain of application of specific phonetic and phonological phenomena. In a circular fashion, these phenomena allow the definition of prosodic constituents as units (with rules applying only between elements of the same constituent) and of their boundaries (with rules not applying across constituent boundaries).

---

4 For a more detailed account of prosodic theories, we refer the interested reader to, among others, Jun (1993, ch. 1), Shattuck-Hufnagel & Turk (1996), Ladd (1996). With respect to the differences between syntactic and prosodic approaches, see Selkirk (1984), Hayes (1989), Inkelas & Zec (1990), for example.

5 According to Nespor & Vogel (1986), the Phonological Phrase (PhP) is defined as a syntactic head and all the elements to its left (non-recursive side, left in French and English) and optionally a complement to its right. In Selkirk (1972, 1986), a Small Phonological Phrase includes a head and the modifiers to its left but not its complements. For example, "la petite fille" forms a PhP, and "la fille charmante" forms one PhP for Nespor & Vogel but two for Selkirk ("charmante" being the head of the Adjective Phrase complement of the Noun Phrase "la fille"). In French, the PhP is often defined as the domain of optional liaison [Selkirk 1986, de Jong 1991].

C.2. Prosodic constituent hierarchies

Descriptions of prosodic systems along several constituent levels can be found in both sequential and superpositional models, both in descriptive phonological frameworks and in applications to speech synthesis, for example. However, the proposed prosodic levels are not always clearly defined and their structural organization is not always obvious. The goal of this section is not to present all the prosodic hierarchies that have been proposed in the literature but to exhibit their commonalities and differences.

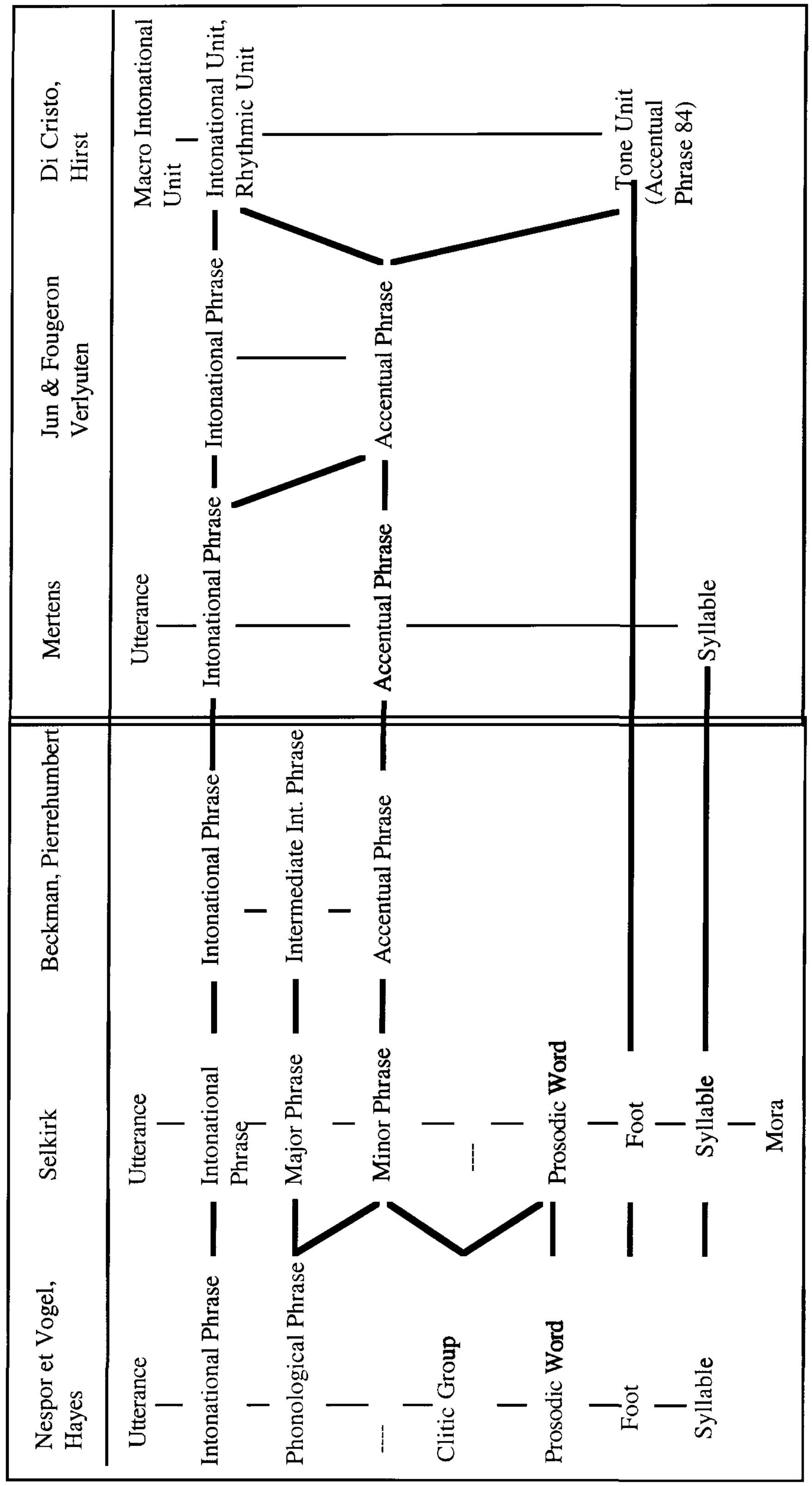
Different approaches use different criteria to define prosodic constituents, which results in multiple constituent types:

(1) Some define prosodic constituents based on morpho-syntactic and phonological criteria [Selkirk 1980, 1984, 1986, Verluyten 1983, Nespor & Vogel 1986, Hayes 1989, Post 1993].

(2) Others define prosodic constituents based on purely prosodic criteria (intonation, final lengthening, pausing, etc.) [Beckman & Pierrehumbert 1986, Mertens 1987, Pierrehumbert & Beckman 1988, Di Cristo & Hirst 1993, Jun & Fougeron 1995, 1998].

Figure 1.3 schematically presents some of the prosodic constituent hierarchies proposed in the literature for English and French. The chart is based on a chart in Shattuck-Hufnagel & Turk (1996: 206) showing hierarchies proposed for English and other languages (left part of fig. 1.3). For a complete definition and comparison of the various constituents, the reader is referred to Shattuck-Hufnagel & Turk (1996) for English and dissertations by Delais (1995) or Sabio (1996) for French.

Figure 1.3. Examples of prosodic hierarchies proposed in the literature for English and other languages (left: from Shattuck-Hufnagel & Turk 1996) and French (right).



This figure shows that different researchers have proposed various numbers and definitions of prosodic constituents in different languages, thus making a comparison of hierarchies difficult. However, three important points stand out:

(1) Prosodic constituents are organized *hierarchically*. For example, Prosodic Words are organized into Phonological Phrases, and Phonological Phrases are organized into Intonational Phrases, and so on. This hierarchical organization thus comprises high- and low-level constituents, which may vary across languages.

(2) Some researchers have proposed that the constituents are *strictly layered* [Selkirk 1986]: constituents at the same prosodic level may not overlap. A constituent X^p is obligatorily made up of one or several constituents at the level immediately below (X^{p-1}) and never includes other X^p constituents. Thus, an X^p-level constituent can only be preceded by a constituent at the same prosodic level (unless it occurs at the beginning of the speech stream). Consequently, the segment occurring at the beginning of a constituent (e.g. an Intonational Phrase) is at the boundary between two constituents at the same level. Other researchers have challenged the Strict Layer Hypothesis and proposed a structure allowing recursive prosodic constituents (e.g. an Intonational Phrase can include other Intonational Phrases) [Chomsky & Halle 1968, Ladd 1986, 1988, 1996, Ladd & Campbell 1991, Itô & Mester 1992 (for Japanese), Di Cristo & Hirst 1996 (for French)].

(3) The prosodic structure of speech is not only achieved via a *constituent hierarchy* but also via a *prominence hierarchy*. "Prominence" is assigned to certain units over others through lexical stress or phrasal stress, for example. Researchers in Metrical Theory [Liberman & Prince 1977, Hayes 1989, etc.] have proposed a finite number of prominence categories organized hierarchically on different levels. For instance, in English, these prominence categories are organized on at least 4 levels: unstressed syllable, lexically stressed syllable, pitch accented syllable, nuclear pitch accented syllable [see Shattuck-Hufnagel & Turk 1996]. The prominence hierarchy is linked to the constituent hierarchy. Each prominence category is associated to a particular constituent, in which the prominent syllable functions as a head, i.e. the most prominent element[6].

The distinction between a head (prominence) and a boundary (constituent boundary) is not always easy and sometimes confusing. In French, for example, it is difficult to integrate stress ("accent") into this dichotomy, which is set up for Germanic languages. In French, final (or primary) stress has a demarcative function. It marks the boundary of a constituent with a rise in F0 and lengthening. The syllable bearing these cues is thus prominent within the phrase compared to the other syllables which are not lengthened or do not show a rise in F0. This syllable can be compared to a pitch accented syllable in English, but it is fundamentally different from it as it primarily identifies a boundary and not merely a head. Initial stress in French ("accent secondaire," "ictus mélodique" for Rossi, "accent didactique" for Lucci, etc.) is also problematic. It is characterized by melodic prominence within a constituent (Accentual Phrase, or even Intonational Phrase) but it is not its head in the sense of "most prominent syllable." It can be a head if its domain is the foot or the tonal unit (Hirst & Di Cristo 1993). Initial stress is also an (optional) cue to the left boundary of a constituent ("initial rise," in Vaissière 1975). Section

[6] For instance, in English, the nuclear pitch accented syllable is the head of an Intermediate Phrase, and the lexically stressed syllable is the head of a foot (!). These two hierarchies (constituent and prominence) are connected but independent because some constituents do not have any particular prominence (they are headless) and vice versa. Section II.A.1 deals with these notions.

II.A.1 returns to these two hierarchies (constituent and prominence) and addresses articulatory variations observed under stress. For now, what needs to be kept in mind is that the prosodic organization of utterances is reflected in prosodic constituents, to mark either their boundaries or prominence relations within them[7].

I.D. BOUNDARIES AND PROMINENCE RELATIONS: THE MARKING OF IMPORTANT POSITIONS IN THE PROSODIC STRUCTURE OF AN UTTERANCE

Prosody has a discourse and a communicative function, and in this sense it conveys the informational (theme, rheme, focus, etc.) and pragmatic (attitude, etc.) structure of utterances.

As was shown above, prosody also serves to structure speech into *constituents* and *prominence relations*. This corresponds to Trubetzkoy's (1949) *demarcative* (dividing up the speech continuum and grouping units into "prosodic" constituents) and *culminative* (making certain units more prominent than others) functions of prosody. This study will be essentially concerned with these structural aspects of prosody.

Prosodic organization is implemented in the speech stream by marking important positions by prosodic means. These positions, according to Beckman and Edwards, are constituent boundaries and prominent units (heads):

> *... prosody is the organizational framework that measures off chunks of speech into countable constituents of various sizes... whereas segmental specifications are facts about the phonetic content of an utterance, prosodic specifications are facts about how that content is organized... we can identify two different devices for creating this structural organization: at any level of the prosodic hierarchy, the number of constituents can be indicated by marking the edges or by marking the heads. [Beckman & Edwards 1994: 8].*

Several parameters are used to mark these important positions in prosodic structure: F0, duration, amplitude, (non-phonemic) spectral variations, and pausing. In general, a boundary or prominence category is not marked by a single acoustic cue but by a set of primary cues. For example, in French, the boundary between two Intonational Phrases is marked by a wide excursion in F0 (rising or falling) and/or by substantial final lengthening and/or by a pause. The boundary between two Accentual Phrases is marked by a smaller rise in F0 and/or less final lengthening without a pause. Thus marking is *multi-parametric*. It is also *scalar*: the parameters are not binary but vary gradually (several degrees of final lengthening). Lastly, the role and possible combinations of these cues vary across languages, in particular in the number of linguistic phenomena to be marked and in the number of available phonetic parameters [Lehiste 1967, Vaissière 1989].

Boundaries and prominence relations are thus marked by prosodic characteristics which, according to Malmberg, are

> *"prosodic elements, i.e. phonemically distinctive with both a demarcative and a*

[7] For a discussion of these various models for French, see Jun & Fougeron (1995, 1998), Mertens (1993), Post (1993), Levac et al. (1993), Hirst & Di Cristo (1996), Vaissière (1996), among others.

> *unifying function." [Malmberg 1971: 213].*

This idea is elaborated on by de Jong, Beckman and Edwards, who equate these prosodic characteristics with segmental specifications:

> *...segmental specifications encode the answers to such questions as "Is this vowel segment front or back?" or "Is this tone segment high or low?" Prosodic specifications encode answers to such questions as "Does this vowel stand as the nucleus of a stressed syllable?" or "Does this tone mark the edge of an intonational phrase?" [de Jong et al. 1993: 199].*

"Prosodic features" are thus treated like phonological features. However, unlike "segmental features," prosodic specifications are realized at particular points in the acoustic signal (e.g. on a vowel nucleus or a final syllable), but these features specify domains larger than the segment (e.g. Syllable, Accentual Phrase, etc.). That is why these specifications are *suprasegmental* features.

Any comprehensive analysis of spoken material must include not only a description of its segmental specifications but also a description of its suprasegmental characteristics. As Faure points out, linguists must also study these

> *melodic contrasts, which can be distinctive too, but at the phrase level, i.e. in actual language in a discourse context, since a given melodic change occurring at a particular point in spoken discourse can cause a change in the identity of a phrase, not only in its subjective content, but also, in certain cases, in its objective notional content, i.e. in the abstract representations conveyed by the message. [Faure 1962:608].*

In Fougeron (1998, this volume) I endeavor to address this double concern by studying both the segmental and suprasegmental specifications of speech. In particular, I examine how the suprasegmental specifications of an utterance marking its prosodic organization can influence the segmental characteristics of the segments making up the utterance. In other words, I address the relations between *contenant* and *contenu*, between sound and structure [Saussure 1915, Grammont 1933].

I. THE INFLUENCE OF PROSODY ON ARTICULATION

In addition to suprasegmental elements such as F0, the prosodic organization of an utterance causes articulatory variations on segments occurring in particular positions in the utterance. These articulatory variations affect features which are generally considered as *segmental* (e.g. oral or velopharyngeal aperture). However, as they are conditioned by prosody, these variations can be included in the inventory of the suprasegmental realization of the prosodic organization of utterances.

The best-known variations are those occurring under stress. The articulation of segments occurring in a stressed syllable is different from that of unstressed segments. In the first section (II.A) I will present some of the variations observed as a function of a segment's position in relation to stress (prominence).

The articulation of segments is also influenced by their position in relation to a constituent boundary. At different prosodic levels (Syllable, Word or higher levels), it has been observed that the articulation of segments is different in initial, medial or final position within a constituent. These variations will be presented in the second section (II.B).

II.A. SEGMENTAL VARIATIONS AS A FUNCTION OF PROMINENCE

Prominence refers to the relative importance given to a unit over other units. The term "prominence" itself covers several categories of stress which are defined according to their level and function (demarcative, rhythmic, emphatic, etc.). In this study, I will not be particularly concerned with the influence of stress on articulation, but rather with variations at constituent boundaries. However, a summary of the data found in the literature is informative because variations under stress show similarities with those observed in initial position within a prosodic constituent.

In French, the issue of stress has been very extensively studied and debated [see Di Cristo & Hirst 1997 for a recent survey]. I will allude to this debate only briefly to show that it is necessary to control the type of stress being studied in order to determine its articulatory characteristics.

Research on the acoustic characteristics of stress has achieved contradictory results. This is due to the fact that the level of stress studied is not always clearly defined or controlled. For example, Fry's (1958) work on English showed that the difference between the stressed and unstressed syllable in English pairs like "PERmit" vs. "perMIT" (where CAPITALS represent lexical stress) was essentially a difference in F0. However, in his corpus, words were presented in isolation. Each word made up its own phrase, and the stressed syllable thus also carried melodic phrasal stress. Thus his results make it impossible to distinguish the variations due to lexical stress from those due to phrasal stress.

This example shows that it is necessary to control the level or the category of the stress being studied in order to observe the acoustic or articulatory characteristics of that stress. In what follows, we will see that reference to prosodic level is also necessary in the study of articulatory variations as a function of prosodic position.

A.1. Preliminary notes on stress and prominence levels

Reference to a language like English, in which lexical stress is contrastive, makes it difficult to define stress in French. This has led some researchers to consider French as a stressless language or a boundary language [e.g. Hjelmslev 1936, Rossi 1980]. However, when one considers several stress categories and stress functions across languages, comparisons are possible. In metrical theory, it is possible to distinguish different types of stress according to the metrical weight of the stressed syllable. This metrical weight is a function of the level of stress in a prominence hierarchy. We saw earlier (I.C) that this prominence hierarchy is close to the prosodic hierarchy. Thus different stress types are also defined with reference to the prosodic level of the constituent to which stress is assigned and of which it is the head [see Shattuck-Hufnagel & Turk 1996].

Prominence is always realized on a particular syllable in the speech stream (a stressable syllable [Garde 1968]). Just like the level of the prosodic constituent with which stress is associated, its acoustic realization will depend on stress type. In general, the acoustic realization of stress includes variations in amplitude, F0 and [e.g. Delattre 1938, 1939, Rigault 1962, Benguerel 1973, Rossi et al. 1981, Touati 1987, Paseloup 1990, Vaissière 1992 for French; Bolinger 1958, Lieberman 1960, Gay 1978 for English]. Stress is often marked by several parameters and the relative importance of these parameters varies across stress types and across languages [e.g. Vaissière 1983, Fant & Kruckenberg 1991, Sluijter 1995].

I will present different types of prominence (or stress levels) with French and English examples[8]. In English, four types of prominence can be distinguished: stressless (reduced) syllable, lexically stressed syllable, pitch accented syllable and nuclear pitch accented syllable. In French, two types of stress can be distinguished: final (primary, phrasal, etc.) and initial (secondary, didactic, *ictus*, etc.). In French, stress is post-lexical and, especially in the case of final stress, it is associated with the notion of constituent boundary: it is primarily demarcative. On the other hand, in English, stress is primarily a constituent-head unit. Figure 1.4 illustrates these different types of prominence by showing the prosodic constituent with which each stress type is associated.

Figure 1.4: Different prominence categories with the constituents they head. See text for details.

Prominence Hierarchy "heads"	**Constituent Hierarchy**	
	English	**French**
Lexical Stress	Foot	Ø
|	|	|
Pitch Accent	?	Accentual Phrase
|	|	|
Nuclear Pitch Accent	Intermediate Int. Phrase	Ø
Initial Accent	Ø	? (U, AP, IP)

a - Lexical stress: Lexical stress is a lexical property. In languages with lexical stress (such as English) this type of prominence is assigned to a particular syllable in the word and its position can be contrastive [Garde 1968]. Two levels of lexical stress can be distinguished in English (or more: see Chomsky & Halle 1968). The domain of lexical stress in English is the foot, defined in this language as a group of syllables including a stressed syllable and every unstressed syllable following it. A word can be comprised of several feet. Acoustically, a lexically stressed syllable differs from an unstressed syllable by the presence of a full (non reduced) vowel, which can have greater amplitude and is generally longer [Beckman & Edwards 1994]. There is no distinctive lexical stress in French.

b - Phrasal stress: The domain of phrasal stress is a constituent at a higher level than the Word. In English, phrasal stress (i.e. pitch accent) falls on a stressed syllable (i.e. carrying lexical

[8] For more details, see (among others) Bolinger 1958, Beckman 1986, Beckman & Edwards 1994, Ladd 1996 for English, and Garde 1968, Rossi et al. 1980, Vaissière 1989, Sabio 1996, Hirst & Di Cristo 1997 for French.

stress) in one of the words in an Intermediate Phrase. Its position is influenced by several factors (semantic, pragmatic, theme/rheme distinction) [see e.g. Bolinger 1958]. Acoustically, it is realized as a high or low variation in pitch (hence the name "pitch accent") and lengthening. An Intermediate Phrase can contain more than one pitch accent. Among these pitch accents, one is more prominent than all the others (in English): it is the nuclear pitch accent (*accent tonique*, sentence accent), and it is the last pitch accent in an Intermediate Phrase. Acoustically, a nuclear pitch accent is often marked by a wider pitch excursion and greater lengthening than pre-nuclear pitch accents[9] [Beckman & Edwards 1994]. The nuclear pitch accent is the head of an Intermediate Phrase. However, pre-nuclear pitch accents do not seem to be associated to any particular prosodic constituents of which they could be the heads [Beckman & Edwards 1990, Shattuck-Hufnagel & Turk 1996]. Pre-nuclear pitch accents thus differ from other types of phrasal stress only in the prominence hierarchy, where they occupy an intermediate level between lexical stress and nuclear pitch accent.

In French, the type of stress (accent) called *final, primaire, logique* or *interne* is a type of phrasal stress [Grammont 1933, Delattre 1938, etc.]. It is realized on the last syllable of a word or of a group of words forming a meaning unit (according to Grammont). I will refer to such a group as an *Accentual Phrase*. It corresponds to the notion of "prosodic word" in Martin (1987), Di Cristo (1978), Vaissière (1992). This phrasal stress has an essentially demarcative function: it marks the end of an Accentual Phrase. The use of the term "pitch accent" for phrasal stress in French is inappropriate: while French phrasal stress is indeed acoustically marked by a variation in pitch, it is mostly distinguished by lengthening of the final syllable.

<u>c- Initial stress in French</u>: French has another type of stress whose exact identity has been much debated. Initial stress (*accent d'insistance*[10], *secondaire, didactique, rythmique, ictus mélodique,* etc.) is generally realized on the initial syllable of the accentual phrase, but it can also occur on a later syllable: between the first and the third syllable [Hirst & Di Cristo 1984, Pasdeloup 1990, Jun & Fougeron 1995, 1997, Hirst & Di Cristo 1996]. It is optional and its occurrence is conditioned by several factors, such as rhythm [Pasdeloup 1990, Delais 1995], speech style [Vaissière 1975, Lucci 1983] or speaker [Vaissière 1975]. Acoustically, initial stress is essentially marked by a rise in F0. The domain of its realization is still debated: for Hirst & Di Cristo, it is a property of the tonal unit (*unité tonale,* UT); for Jun & Fougeron, it is a property of the Accentual Phrase (AP); for Rossi or Vaissière, it can be a property of the Word, of the utterance, or of the Intonational Phrase.

<u>d - Focal or emphatic stress:</u> focal stress is a "syntactic/pragmatic" type of stress [Di Cristo & Hirst 1997]. In English, it is realized as a nuclear pitch accent (but a nuclear pitch accent is not always a focal stress). It occurs on the (lexically) stressed syllable of the focalized word, but it can also occur on a lexically unstressed syllable if that syllable is contrastive (<u>ex</u>port vs. <u>de</u>port). In French, emphatic stress is often realized on the syllable bearing initial stress, but it can also be realized on the syllable bearing final stress or on both (initial- and final-stressed syllables). Acoustically, focal stress is marked (in French) by a rise in F0 and greater amplitude [Touati 1987, Pasdeloup 1990, Jun & Fougeron 1997, Hirst & Di Cristo 1999].

[9] Nuclear pitch accent is perceived as more prominent by listeners even when it cannot be distinguished from pre-nuclear pitch-accents by specific acoustic characteristics [Shattuck-Hufnagel & Turk 1996].

[10] The label *accent d'insistance* is misleading because this stress is not emphatic. Emphatic stress can be realized on the syllable bearing initial stress, but initial stress is not by nature emphatic.

A.2. Articulatory realization of stress in French

There exist very few systematic articulatory studies of stress in French compared to English. I will give a detailed account of some of these studies. The type of stress studied is essentially final stress or emphatic stress. I do not know of any study comparing the articulatory characteristics of different types of stress (final, initial emphatic) in French.

Straka (1963) did not study systematically articulatory differences between stressed and unstressed syllables. He mentions the differences observed under stress when he presents the characteristics of reinforced speech (which I address below). In addition, he does not specify what he meant by stress. In general, he collapses the variations observed under some type of (probably) emphatic stress with those observed under phrase-final stress. To him, the differences between stressed and unstressed syllables are due to a difference in articulatory force. Stressed syllables, articulated more forcefully, maximize the aperture differences between vowels and consonants. Articulatory gestures are longer and more extreme. Consequently, consonants exhibit greater closure and vowels are more open. He shows X-rays of a stressed final /a/ and of an unstressed /a/, in which the jaw and tongue position are much lower in the case of stressed /a/.

Simon (1967) also observes the effects of stress only indirectly in her X-ray study of the articulatory characteristics of French consonants. Her corpus of French utterances is very varied and does not allow her to compare every syllable in comparable positions distinguished by a stressed difference only. However, she devotes a small section to the characteristics of Intonational-Phrase-final stressed syllables. For consonants, she observes that the duration of articulatory steady-state varies as a function of stress. Onset consonants in stressed syllables have longer closure durations than the same consonants in unstressed syllables (p. 210). The ratio of unstressed to stressed consonant closure duration is generally 1:2 or less, but never less than 2:3. Vowels also exhibit greater duration of their steady-state: while unstressed vowels are characterized by continuous articulatory movement (no steady-state), stressed vowels are characterized by a marked steady-state period. The difference in the realization of stressed and unstressed vowels is thus very clearly marked. Simon concludes that the effect of stress on steady-state duration is stronger for vowels than for consonants. Unfortunately, she does not provide data on the amplitude of articulatory gestures under stress.

In her study of the articulatory characteristics of French vowels, Brichler-Labaeye (1970) examines mostly vowels bearing final stress. However, she also studies some of the pre-tonic (or ante-pre-tonic) vowels in her corpus to compare them to final stressed vowels. Like Simon, she observes that unstressed vowels are characterized by shorter duration and especially by an absence of an articulatory steady-state. Unstressed vowels are articulated more cursorily (1970:104). She only gives a detailed study of the open vowel /a/ in unstressed position and shows that it is less open than stressed /a/. For close vowels, she cites the works of Chlumsky (1938) or Rousselot (1901) which showed that the vowel /i/ is more closed in stressed position than in unstressed position.

Unlike the previous studies, Giot (1977) studied systematically the effects of final stress on the articulation of syllables in French. His corpus is comprised of CV syllables where V is one of the four front unrounded vowels of French /i e — ɛ a/. His study is based on X-ray data and spectrographic analyses. He observes that in stressed position, the open vowel /a/ exhibits more aperture, together with a higher F1 and a lower F2, greater duration and a steady-state part not observed in unstressed /a/. Non-low front vowels /i e ɛ/ in stressed position show less aperture under the palate (or under the alveolar ridge for /i/) than in unstressed position. This shrinking of the front resonator is accompanied by a lowering of F1 for /i/ and /ɛ/. For /e/, F1 is stable or lower in most cases, but higher in others. All three vowels also have a higher F2 in stressed position. In sum, compared to their unstressed counterparts, stressed non-low front vowels are less open and fronter and the low vowel /a/ is more open when stressed.

Vatikiotis-Bateson (1988) and Fletcher & Vatikiotis-Bateson (1994) studied the movement of the lip/jaw unit as a function of stress in French, based on a corpus of reiterant speech. The authors collapse syllables bearing initial and final stress. They show that gestures exhibit greater duration, amplitude and velocity in stressed syllables. In the later study (1994), the authors compare Intonational-Phrase-final stressed syllables (before a pause) and Intonational-Phrase-medial stressed syllables (bearing final or initial stress). They show that the gestures of final stressed syllables are even longer, wider and faster than those of non-final stressed syllables. They conclude that the distinction between stressed and unstressed syllables results from a difference in the specification of gesture-internal kinematic parameters (intra-gestural differences), whereas the difference between Intonational-Phrase-final and non-final stressed syllables is due to a difference in the phasing of the closing and opening gestures (inter-gestural differences).

Meynadier et al. (1998) show that focal stress also alters the lingual articulation of consonants in French. They use electropalatography to show the variations in the lingual articulation of up to four-consonant clusters (/kskl/). Sequences are produced with contrastive focal stress placed on the vowel preceding or following the consonant cluster and are compared to unstressed sequences. The authors show that the presence of focal stress lengthens the overall duration of the consonant cluster and of each individual consonant. The time interval between the release of a closure and the beginning of the following closure is longer under stress. This lengthening does not result from shorter overlap between consonantal gestures but from a lengthening of the non-overlapping closure and release phases of each consonant.

In summary, these studies of French show that, in a stressed syllable (bearing final, initial or emphatic stress), the articulatory gestures of consonants and vowels generally have greater amplitude, duration and (sometimes) velocity than in an unstressed syllable. Stressed vowels have a more peripheral tongue position than unstressed vowels.

A.3. The articulatory realization of stress in English and other languages

The articulatory realization of stress in English has been extensively studied, and so I will only present a few studies in this section. First, I will review some of the studies that showed that articulatory variations under stress make it possible to distinguish several degrees of stress in English. Then I will present studies showing different results with respect to the effects of stress as a function of the particular segments under study.

A3.a. Different articulatory realizations depending on stress type?

Several types of stress can be distinguished in English according to their articulatory correlates.

Stone (1981) examined the articulatory correlates of different types of stress in English by comparing the displacement and velocity of jaw opening gestures. She found that the velocity of the jaw opening gesture makes it possible to distinguish three degrees of stress: no stress, lexical stress and phrasal stress ("prominently stressed, beat syllable," which corresponds to the nuclear pitch accent). The velocity of the jaw opening increases gradually from one stress type to the next. However, the displacement of the jaw appears to be more variable and depends on phonemic variations, speech rate and position in the sequence. The velocity of gestures thus appears to be a more reliable correlate of the degree of stress than the movement of the articulators as a whole.

Beckman & Edwards (1994) compared the articulatory correlates of three degrees of prominence in English: no stress (reduced syllable), lexical stress and contrastive nuclear pitch accent. They examine the movement of the lower lip and jaw for the syllable [pɑ]. The most remarkable articulatory effect is that which distinguishes a lexically stressed syllable from an unstressed syllable. The stressed syllable contains a full vowel, articulated with much longer, wider and faster jaw opening gesture than the reduced vowel in the unstressed syllable. The stressed syllable is also longer. However, when the authors compare syllables containing full vowels at different prominence levels (lexical stress vs. nuclear pitch accent), the articulatory distinction is less obvious. Nuclear-pitch-accented syllables are generally longer (3 speakers out of 4) than lexically stressed syllables. This lengthening can be associated with a slightly wider and faster opening gesture, but these characteristics are not shown by all speakers or at every speech rate. The authors conclude that the specification of higher prominence levels, such as phrasal stress and nuclear pitch accent, is mainly achieved as an F0 variation and optionally as an articulatory variation (essentially lengthening). The greatest articulatory variation is observed between full and reduced vowels.

In their 1992 study, Beckman et al. examine in more detail the articulatory factors causing the lengthening of nuclear-pitch-accented syllables. They compare syllables bearing non-contrastive nuclear pitch accents to unaccented but lexically stressed syllables, examining the kinematic properties of jaw movement. The lengthening of nuclear-pitch-accented syllables is due to longer and larger (+2mm) jaw opening and closing gestures. The jaw opening gesture is faster in pitch-accented syllables. However, the lengthening of the syllable is not due to an increase in the stiffness of the gestures but to a difference in the phasing of the opening and closing gestures. In nuclear-pitch-accented syllables, the vocal tract remains open longer because the jaw closing gesture begins later in relation to the opening gesture. This lag between closure and opening can account for the greater displacement of the opening gesture, which is not truncated by the closing gesture as is the case in non-accented syllables.

In summary, these studies show that in English it is sometimes possible to distinguish three degrees of prominence on the basis of their articulatory characteristics. The most striking distinction is the difference between reduced unstressed syllables and lexically stressed syllables, whose vowels are full, longer and more open. The distinction between lexical stress and nuclear pitch accent is less clear. A nuclear-pitch-accented syllable is generally longer and the jaw position is more open (which may be due to greater velocity of the opening gesture or later phasing of the closing gesture). A greater jaw opening has also been reported in the case of emphatic nuclear pitch accents by Erickson & Fujimura (1992).

A.3.b. Different articulatory variations depending on segment type?

The studies presented above show that stress is marked by variations in jaw movement. Several authors have noted that the position of the jaw is particularly sensitive to variations under stress [Kent & Netsell 1971, Giot 1977, Beckman et al. 1992, 1994]. This has led some authors to consider the jaw as the articulator in charge of marking prosodic organization [Macchi 1988, Beckman et al. 1992]. However, while the position of the jaw does contribute to the articulatory distinction between vowels and consonants, it contributes very little to the phonological specification of various phonemes. Thus, the jaw may vary more freely as a function of stress than other articulators, such as the tongue, the velum or the lips, whose position directly contributes to the specificity of a segment's articulatory features. In what follows, we will examine whether these other articulators are affected by stress.

This issue is particularly interesting when comparing segments with different apertures. In most of the above studies, the syllables studied are made up of an obstruent followed by an open vowel [Beckman et al. 1992, 1994, Erickson & Fujimura 1994, 1996 (but /a/ and /e/ in Stone 1981)]. Greater jaw opening under stress thus contributes to reinforcing the distinctive character of open vowels. We will see what happens in the case of segments whose articulatory specifications conflict with an increase in opening, in particular close vowels.

α. Lingual articulation in relation to stress

The direction of variations in the tongue position of different vowels varies across studies.

Houde (1967) observes that the tongue is lower for all stressed vowels in English.

On the other hand, Kent & Netsell (1971) show that, in English, the effect of lexical stress on the position of the tongue depends on the vowel in question: variations go towards the "presumed target" of the vowel. The tongue for a stressed /i/ is higher and more anterior, while it is lower for a stressed /æ/.

Macchi (1985) also observes that the effect of lexical stress on the tongue/jaw unit varies across vowels in English. The jaw is lower for all stressed vowels, and consequently the front oral cavity is more open for all vowels. However, the opening of the jaw is associated with a lowering of the tongue in low vowels /ɑ/ and /ʌ/ but not in non-low vowels /u i ɛ/, in which case the tongue keeps its regular oral constriction and is not affected by stress.

De Jong (1995) shows similar results in his study of the articulatory correlates of contrastive nuclear pitch accent for various segments in English. On the whole, he finds that prominence is marked by an increase in the articulatory movement of the jaw, lips and tongue. With respect to the tongue, variations under stress depend on the type of vowel and contribute to making the vowels more peripheral. The body of the tongue is lower in low vowels and higher in non-low vowels (e.g. "toes", "toast") and backer in non-low back vowels (e.g. "put").

Farnetani & Vayra (1996) confirm the variability of the effect of stress on different segments in Italian. They observe lexically stressed and unstressed syllables. Their electropalatographic data show that variations in lingual articulation under stress reinforce the segments' features: /a/ is more open, /i/ is closer, /u/ more posterior (and the constriction area of /t/ is greater) under stress.

Harrington, Fletcher & Beckman (1998) compare the effect of stress on jaw and tongue gestures in close and open vowels in Australian English. They study syllables bearing nuclear pitch accents and compare them to unaccented, lexically stressed syllables. The presence of a nuclear pitch accent lowers the jaw for open vowels but also for close vowels. This jaw lowering is accompanied by an increase in the vowel's energy. However, the effect of stress on tongue position varies across speakers. In the case of stressed /i/, the tongue is more anterior and lower for one speaker but higher for the other speaker. Acoustically, these articulatory variations cause a higher F1 for one speaker and a higher F2 for both speakers. The authors conclude that the speakers use two different strategies aiming at reinforcing the distinctive characteristics of close vowel /i/ under stress: one speakers produces a closer vowel (higher tongue position), and the other produces a fronter vowel (advanced tongue position). But for both speakers, these variations are accompanied by an opening of the jaw.

β. Labial articulation in relation to stress

Only one speaker in Kent & Netsell's (1971) study exhibits more lip protrusion for rounded vowels in English under stress. Their other speaker shows little protrusion in all cases and it is not affected by stress. De Jong (1995) shows that in the case of the labial consonant /p/ in English, the lips are more protruded and more tightly approximated when it is nuclear-pitch-accented. The vowel /ʊ/ in "put" also exhibits more protrusion and a lowered lower lip in stressed position for two of his three speakers.

Harris et al. (1968) and Slis (1971) show that the electromyographic (EMG) activity of the lips is greater under stress. In Dutch, Slis observes a 16% increase in EMG activity for /p/ in lexically stressed syllables and a 20% increase in emphatically stressed syllables compared to unstressed syllables.

γ. Nasal articulation in relation to stress

Vaissière (1988) studied velum gestures in English as a function of the presence or absence of lexical stress. She observes an effect of stress on oral and nasal consonants, but this effect depends on the position of the consonant within a word (and a syllable). For both oral and nasal consonants, the velum is higher in initial position and lower in coda position in stressed syllables compared to unstressed syllables. Krakow (1989, 1993) confirms the higher position of the velum under stress in English oral consonants. In the case of stressed nasal consonants in CVN or NVC syllables, she notes that the velum remains open longer than in unstressed syllables. Slis (1971) citing Fritzell (1969) notes that the velum muscles show greater activity under stress in English.

As for vowels in English, Krakow (1993) observes that the effect of stress on velum position in oral vowels varies across speakers. One speakers emphasizes the differences in intrinsic velum positions: the velum is higher in stressed /i/ and lower in stressed /a/. However, the other speaker has a systematically lower velum in stressed vowels regardless of vowel height.

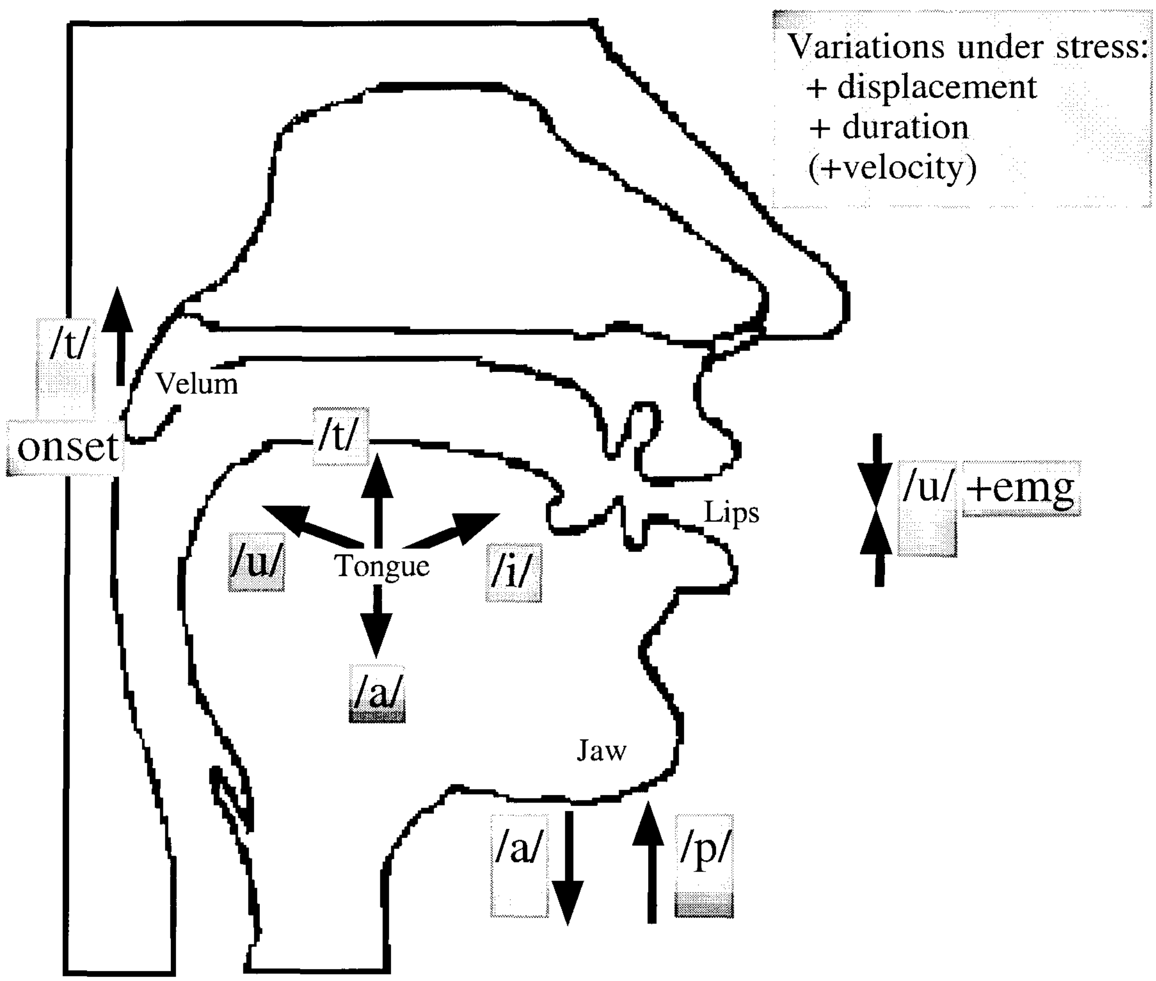
A.4. Articulatory variations in relation to stress: Conclusion

To summarize the results presented in this section, the presence of any type of stress affects the supraglottal articulation of segments. The kinematic characteristics of the jaw as well as the other articulators (tongue, velum, lips) are altered by the presence/absence of stress. This has been shown in French, English and a few other languages. Stress affects the whole stressed syllable: the articulation of both vowels and consonants is altered.

Figure 1.5 summarizes the articulatory variations observed under stress, for each articulator and segment type. Oral obstruents (e.g. /t/) exhibit longer closure under stress (see e.g. Simon for French), a lower jaw position (see e.g. De Jong for English), a higher tongue position (see e.g. Farnetani & Vayra for Italian) and a higher velum position (see e.g. Vaissière (only for initial consonants) and Krakow for English). The jaw is lower under stress for all vowels

(Macchi, Beckman, Stone, Harrington et al., Straka, etc.). However, the position of the tongue under stress varies in different directions depending on the vowel and contributes to making vowels more extreme (peripheral). The direction of articulatory variations under stress is thus articulator-, segment- and sometimes speaker-specific.

Figure 1.5: Summary of articulatory variations observed under stress. The variations are illustrated for each articulator with a segment representing its category (e.g. /t/ for oral obstruents, /a/ for open vowels).



II. B. Segmental variations as a function of a segment's position within a prosodic constituent

We have seen what kinds of articulatory variations have been observed in stressed position. We now examine how the articulation of a segment varies according to its position within a prosodic constituent, i.e. as a function of its *prosodic position.*

Three positions can be distinguished within a constituent: *initial*, *medial* and *final.*

The final and initial positions are *boundary positions.* Segments in initial and final position are both at the boundary of a given constituent and at the boundary between two constituents. The *prosodic level of the boundary* between constituents will depend on the level of these constituents within the prosodic hierarchy. The higher-ranked the constituents in the hierarchy, the heavier the prosodic weight of the boundary. For example, the boundary between two Intonational Phrases has a greater weight than a boundary between two Words.

Thus, as in the case of stress, if one wants to study articulatory variation as a function of prosodic position, the prosodic level of the constituent studied must be specified. However, in a number of studies, the hierarchical level of constituents (and thus of their boundaries) is not controlled. For instance, some studies mention variation according to position within the Syllable, while the items being studied are monosyllabic or isolated words. But in a monosyllabic word, positions within the Syllable are merged with positions within the Word. Besides, when the items are produced in isolation, they form an Accentual or Intonational Phrase by themselves, and thus the positions studied are hierarchically higher.

In this review, I will first present the articulatory variation observed within lower-level constituents: the Syllable and the Word (B.1). Most studies are generally concerned with comparing the initial and final position within a constituent.

Second (B.2), I will deal with articulatory variations at levels higher than the Word level (e.g. Phonological Phrase, Intonational Phrase). These less common studies are part of a relatively recent trend that concerns itself particularly with the prosodic level of the constituent being studied. This addresses the issue of whether *segmental variation can distinguish several levels of prosodic boundaries.* These studies compare the initial position in constituents at different prosodic levels (e.g. Word-initial vs. Intonational Phrase-initial).

Articulatory variations as a function of prosodic position have been studied in several languages. We will see that English data dominate the literature, but that isolated observations have been made on less-studied languages (e.g. Estonian, Korean, Taiwanese). Few studies have been conducted on French (hence the relevance of the study in Fougeron 1998, this volume). The following section presents a summary of research findings based on observations made on different languages, grouped by articulator. The universality of the phenomenon is discussed in Fougeron (1998), Chapter 6.

<u>B. 1. Articulatory variations according to position within a lower-level constituent: Syllable and Word</u>

The existence of positional allophones (initial, medial and final) has been studied most extensively at the Syllable or Word level. For a long time, researchers have been concerned with the Syllable as a unit of the motor organization of speech [Delattre 1940a, Malmberg 1950, 1971, Stetson 1951, Kozhevnikov & Chistovich 1965, etc]. These works focus on the identification of cues to syllable boundaries at the physiological and/or acoustic level. Their goal is to define the syllable, not only as a unit of linguistic representation, but as a structural unit of production.

At the Word level, the identification of positional allophones also addresses the issue of finding cues to boundaries between lexical units in continuous speech. The identification of a phonetic unit corresponding to the Word would then make it possible to understand the segmentation and lexical recognition processes used by the listener. For example, Lehiste (1960, 1961, 1964)'s remarkable research program investigates the acoustic realizations of syllable, word, and morpheme boundaries, and in particular the allophonic variations of certain segments according to their position. She claims that it is possible to distinguish segmentation indicators in the speech continuum, such as the presence of certain positional allophones. For instance, based on the formant spectrum of /l/, she distinguishes an initial and a final allophone. The presence of an initial allophone in the middle of a word allows her to identify a morpheme boundary, which for example distinguishes the morpheme *–ly* in the word *highly* from *wily*, which is monomorphemic.

<u>Preliminary note on the distinction between "onset" and "coda": not only a positional difference, but also a difference in the combination of segments</u>
Most of the studies that we surveyed are concerned with differences in articulation between initial and final consonants. Few studies show a comparison with the medial position and very few results are given for variations in the articulation of vowels.
However, in a Syllable or a monosyllabic Word, consonants in initial and final position are not only distinguished by a positional difference but also by a *difference in segmental combination.* In a CVC syllable, the onset consonant is a pre-vocalic consonant, while the coda consonant is a post-vocalic consonant. Thus, the direction of articulatory movements for these two types of consonants is different. For an onset consonant, the articulators effect an opening gesture, from a consonantal stricture to a wide vocalic opening. For a coda consonant, the articulators close, from a vocalic opening to a consonantal stricture. The aerodynamic forces, the tension of the articulators and the coordination of the various movements are thus not the same. It follows that these segments have different acoustic characteristics [see for example Malmberg 1950, Delattre et al. 1955, Fujimura et al. 1978, Gow et al. 1996].

In this section, for ease of exposition and comparison with the following sections, I will present the findings with an emphasis on the differences observed in initial position. However, in the literature, the articulatory variations observed are most frequently presented with an emphasis on the reduction of articulatory movements in coda position. The articulatory variations are listed by articulator.

α. Syllable-Word: Glottal articulation

In several languages, the presence of a Syllable or Word boundary can be associated with a change in phonation type: laryngealization, aspiration, glottal stop insertion. This change affects both consonants and vowels in initial position.

<u>Initial consonants:</u> aspiration in a voiceless consonant in Syllable-initial position in

English is the best-known positional allophonic variation [see for example Kahn 1976]. This aspiration is an important cue to the syllabic segmentation of a VCV sequence: aspiration in the consonant indicates a /V.CV/ syllabification (where the period marks a syllable boundary) [Lehiste 1960]. Garding (1967) notes that aspiration has a similar function in Swedish. Cooper (1991) has shown with transillumination data that the glottal opening gesture is wider and longer for Word-initial consonants in English compared to Syllable-initial consonants in Word-medial position. This gesture, which is wider in initial position, also begins earlier than the oral closure (lingual or labial).

In contrast, the glottal gesture is often reduced Word-finally. Lisker & Baer (1984) shows that the glottal opening gesture is often absent in this position (the glottis remains constricted). Its trace is still present in EMG data, but the muscular activity associated to the glottal opening gesture has a reduced amplitude Word-finally.

<u>Initial vowels:</u> With respect to vowels, modifications in glottal articulation are generally a cue to Word-level boundaries. In English and German, for instance, vowels in Word-initial position are often glottalized [Umeda 1978, Gimson 1980, Kohler 1994]. Lehiste (1964) also notes a brief period of laryngealization or glottal stop insertion in Word-initial vowels in Finnish and Czech. This laryngealization, which Lehiste calls a "boundary segment," makes it possible to distinguish a sequence of vowels /VV/ separated by a Word boundary (/V#V/) from a sequence of vowels separated by a Syllable boundary (/V.V/). The glottalization of an initial vowel can also be a cue to a morpheme boundary within a polymorphemic word [Kohler 1994]. It is also a cue to syllabification: in Swedish VCV sequences, the presence of a glottal stop or glottal constriction before the second vowel induces syllabification as VC.V [Garding 1967].

By contrast, in French, Word-initial vowels are rarely glottalized when they are not also initial in a higher-level constituent. When a Word-initial vowel is not glottalized, linking with the preceding consonant or a hiatus with the preceding vowel occurs. [e.g. Delattre 1965].

β. Syllable-Word: Jaw articulation

In contrast to what we saw in the case of variation under stress, the effect of prosodic position on jaw articulation has been extremely understudied.

Macchi (1988) notes a difference in jaw height between an onset /p/ and a coda /p/ in English. However, her two speakers present opposite variations: one speaker's jaw is lower for onset /p/ than for coda /p/, whereas the other speaker exhibits the opposite behavior. Stone (1981) noticed that, for a constant level of stress, the jaw is higher in the initial /d/ in a sequence of 8-9 CV syllables than in the following /d/'s.

γ. Syllable-Word: Labial articulation

In the case of labial articulation, few data are available about the influence of prosodic position. However, the few data presented on the EMG activity of the lips are very informative.

Fromkin (1965) shows that in English, the muscular activity of the Orbicularis Oris is greater and longer for a labial consonant in Word-initial position. McAllister et al. (1974, quoted in van Lieshout et al. 1995) also observed that the EMG activity of the lips is longer and greater for Swedish rounded vowels Word-initially than Word-medially. In French, Straka (1963), based on photographs, notes that the lips are pressed more tightly in labial stops Word- or Syllable-initially. Bonnot et al. (1986) note in the case of French that in nonsense words of the form CVCVCV, the muscular activity involved in lip closure, the position of the peak of maximum activity and the total duration of muscular activity make it possible to distinguish the articulation of consonants /p/ and /m/ in initial position. It thus appears that vowels and consonants exhibit stronger muscular activity in initial position.

With regard to lip movement, Macchi (1988) does not observe any differences in the movement of the lower lip (from which she subtracted jaw movement) between an onset /p/ and a coda /p/. Krakow (1989) also notes that the position of /m/ within the Word or the Syllable rarely affects lip movement. However, when this closure movement does vary, it varies towards increased amplitude and duration.

δ. Syllable-Word: Lingual articulation

The influence of prosodic position within a Syllable or a Word has been studied most extensively with respect to lingual articulation. The data collected from the literature include variations in spatial tongue movement as well as variations in pressure of the tongue against the palate.

<u>Spatial variations in consonants:</u> the variations in lingual articulation according to the position within a Word or a Syllable were observed long ago in French by l'Abbé Rousselot (1901). He notes from palatography data that the area of contact of the tongue against the palate is wider in initial than in final consonants. These results are confirmed by the palatography data in Straka (1963). Straka notes that this widening of linguopalatal contact corresponds to a higher raising of the tongue apparent in cineradiographic data. In initial position, the tongue is widely pressed against the palate. Palatograms show that this contact is widened both in the center and on the sides of the palate. This raising of the tongue not only affects stops but also fricatives. In her palatographic study of apical and laminal articulations, Dart (1991) notes a difference in the articulation of French and English anterior consonants in Word-initial and Word-final position. The variations between initial and final position affect mainly the place of articulation and the width of constriction. Unfortunately, the author does not describe precisely the nature and the direction of the variations observed. She only provides the number of speakers (out of 21 French and 20 English speakers) who do exhibit variations between these positions for every consonant. In French, the consonants /t d n l s z/ vary according to their position for 30 to 50% of speakers. In English, non-fricative consonants vary in similar proportions according to their position, but fricatives are much less affected than they are in French. English /l/, however, varies much more than French /l/.

Other studies of English show that variations in consonants according to their position within a Word or Syllable consist in a higher raising of the tongue in initial position, as in French. This has been observed Syllable-initially for the tongue tip for /l/, /t/ and /n/ and for the tongue dorsum for /k/ [Browman & Goldstein 1995]. In electropalatography data (EPG), this raising of the tongue causes a widening of the area of linguopalatal contact, which is often associated with a lengthening of closure duration [Byrd 1994, 1996, Wright 1994, Keating & Wright 1994, Keating 1995]. For the fricative /s/, Byrd (1994) does not observe any difference between initial and final position spatially (as also found in Dart 1991) but she notes that constriction is longer in initial /s/. Farnetani and Vayra (1996) also observe an increase in linguopalatal contact Word-initially in Italian. In that study, the authors compare the articulation of /t/ Word-initially to that of /t/ Syllable-initially but Word-medially in CVCVCV sequences. They show that /t/ exhibits more linguopalatal contact Word-initially than Word-medially. Farnetani (1986) also shows that linguopalatal contact for /n/ in Italian is longer and greater Word-initially (after a pause) than Word-medially.

Pressure variation in consonants: a higher raising of the tongue is accompanied by an increase in tongue pressure against the palate. Rousselot (1901) claims that onset consonants are articulated more strongly, which he deduces in palatographic data from the clear delineation of the contour, the size of the linguopalatal contact area and the speed at which the stain left by the tongue disappears. This idea was later developed with pressure sensors, which made it possible to quantify the pressure exerted on the palate by the tongue more precisely. McGlone & Proffit (1967) and McGlone et al. (1967) show with a few sensors that the difference in tongue pressure against the palate makes it possibleto distinguish a CV from a VC syllabic structure. The pressure of the tongue against the incisors is stronger in onset than in coda consonants [McGlone et al. 1967]. This appears in the case of different obstruents (/n t d l/) but not for the fricative /s/. The positional difference is almost double for /n t/, slightly less for /d/ and more than double for /l/. By contrast, the pressure difference between onset and coda position does not appear on the lateral pressure sensors located on the molars. It thus seems that the increase in pressure is effected at the tongue tip. The authors do not provide data for non-anterior consonants articulated with the tongue body. The difference between onset and coda consonants also implies a difference in temporal alignment between the pressure peak and the beginning of phonation: for onset consonants, the pressure peak is reached at the beginning (and sometimes before) the acoustic beginning, whereas for coda consonants, that peak is reached towards the end of the consonant. McGlone and Proffit (1967) note that the pressure differences according to the consonant's position within a Syllable are more clear-cut than the pressure differences between different consonant categories.

Lingual articulation variation in vowels: Compared to consonants, few studies have been devoted to articulatory variations in vowels according to their position. However, the articulation of vowels also appears to be affected by their position, when one observes the articulation variations indirectly, through their (acoustic) effect on the spectrum. Lehiste (1964) presents data on Finnish, in which she notes that a Word-initial vowel is distinguished from a Word-medial, Syllable-initial vowel by less centralized F1 and F2 formants. Such a vowel, longer and more peripheral, makes it possible to distinguish /VV/ sequences separated by a Word-boundary (in a compound word, e.g. *lintu-ansa*) from /VV/ sequences separated by a Syllable-boundary only (e.g. *lintuansa*). She observes this in the case of several Finnish vowels /i u ä A/ (for /i/, higher F2 and lower F1; for /u/, lower F2 and F1; for /ä A/, higher F1 mostly). If F1 and F2 variations are interpreted as a consequence of a difference in lingual articulation, it would seem that the tongue position in Finnish vowels is more peripheral word-initially. However, the author notes that this variation in vocalic quality Word-initially may be language-specific, because it does not appear in Czech [Lehiste 1967]. Straka (1964) does not provide articulatory data for initial vowels, but notes that the initial position in a Syllable is the "strongest" for a vowel (e.g. in *haïr*, *péage, arbre*). Conversely, Syllable-final (pre-consonantal) or Word-final vowels are in a "weaker" position and may be weakened (we return to the distinction between "weak" and "strong" in Section IV).

The literature often presents information on the articulation of vowels which are described as "initial." However, in these studies, these vowels occur within a CV initial syllable.

For example, Farnetani and Vayra (1996) study the characteristics of linguopalatal contact for the vowels /i a u/ in CVCVCV sequences. They note that the vowel which is placed in the initial Syllable of the Word is more open than the vowels in following syllables. In Fougeron (1998), I do not take vowels in an initial CV Syllable to be initial (neither in the Syllable nor in the Word).

ε. Syllable-Word: Nasal articulation

In several languages, it has been observed that the velum is higher in Syllable- or Word-initial position, in both nasal and oral consonants.

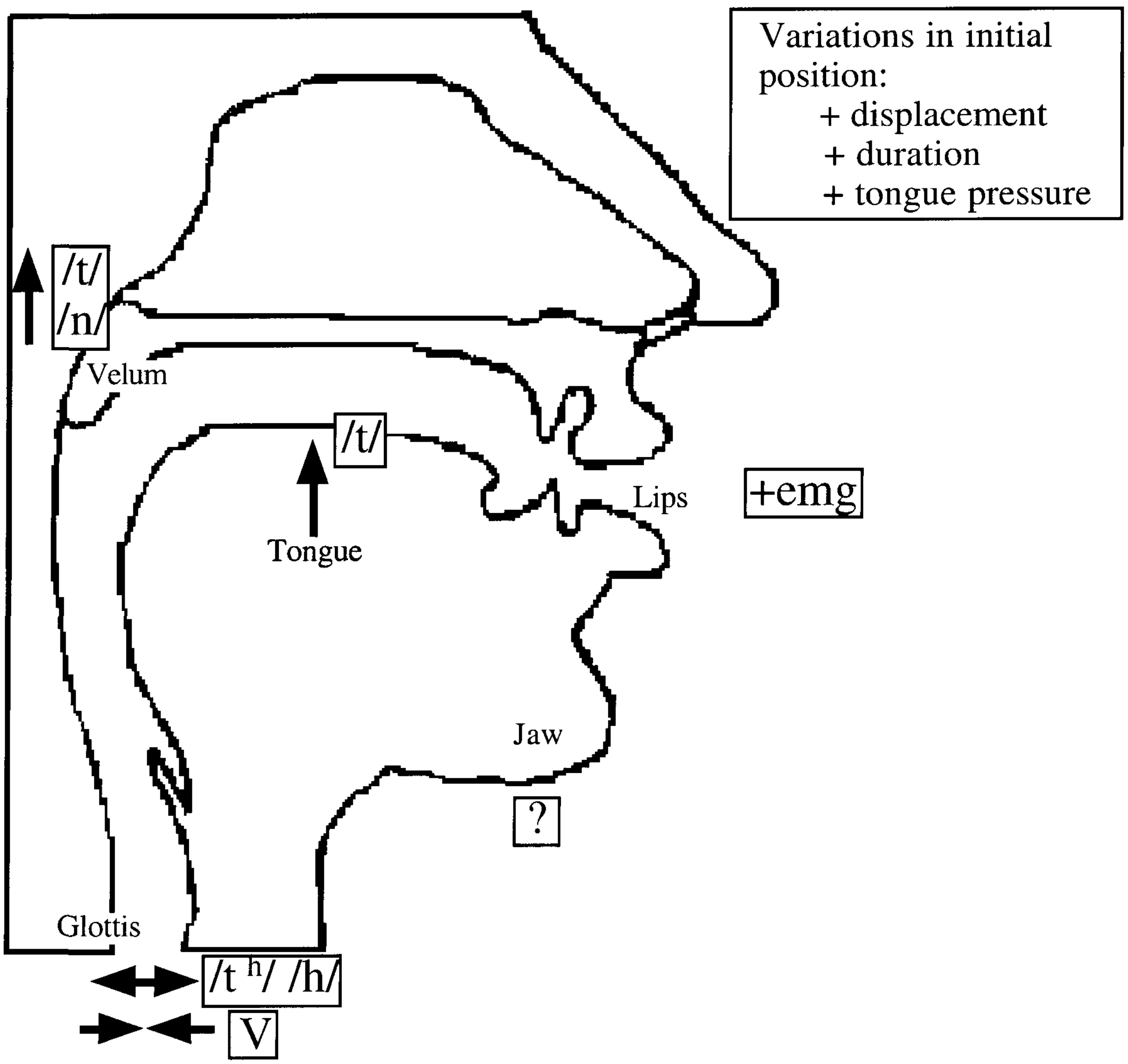
In the case of oral consonants, a CVC sequence in French reaches a maximum in velum height for each oral consonant, but the first maximum is always higher than the second one [Benguerel 1977]. Additionally, Simon (1967) shows that oral consonants in French (stops, fricatives or laterals) exhibit a narrower constriction of the velum against the pharyngeal wall and/or a higher velum in Word-initial position than in intervocalic position. Vaissière (1988) observes the same phenomenon in English: the velum is higher in Word-initial oral consonants.

In the case of nasal consonants, the velum is also higher in onset than in coda position. This has been observed in English and Japanese [Fujimura 1977, Fujimura & Lovins 1978, Fujimura 1990, Krakow 1989, 1993], but not in French (to my knowledge). In English, Krakow (1989) adds that the movement of the velum is smaller and that the velum remains lowered for a shorter period in the case of a Word-initial /m/. Comparing Syllable-initial /m/ to Word-initial /m/, Krakow finds very little difference. She concludes that the effect she observes is linked to the position within the Syllable. Manuel (1991) interprets variations in the position of the velum as a characteristic of the initial position, which aims at decreasing the sonority of an initial nasal consonant (a raising of the velum causes a decrease in nasal airflow, and thus in overall acoustic energy). The less sonorous consonant will thus be less distinct in its context in terms of acoustic energy. The nasal murmur is also shorter in an onset consonant than in coda position [Fujimura & Erickson 1997].

ζ. Variations within the Syllable and the Word: Conclusion

The glottal and supraglottal articulations of segments in Syllable- or Word-initial position can be distinguished from those of medial or final segments. Figure 1.6 summarizes the results found in the literature. In initial position, the glottal opening gesture for consonants is longer and greater. Vowels are glottalized or preceded by a glottal stop. Labial muscular activity in initial consonants and vowels is greater. The velum is higher in initial oral and nasal consonants. The tongue is higher and linguopalatal pressure greater in consonants. The few spectral data available about vowels suggest that the tongue has a more peripheral position word-initially in some languages.

Figure 1.6: Summary of articulatory variations observed in Syllable- or Word-initial position. These variations are illustrated by one articulator with an example of a segment representing its category (e. g. /t/ for an oral plosive, V for a vowel).



B. 2. Articulatory variations as a function of a segment's position within constituents higher than the Word

We have seen that articulatory variations according to the position within a Word or a Syllable have most frequently been studied as differences between initial and final position. However, these two positions are distinguished by a difference in the combination of segments, which also has an influence on their articulation.

The special status of the initial position appears more clearly in the following studies, in which the articulation of segments is compared in initial position *at different levels.*

A comparison between these positions highlights the difference between the *initial* and *medial* position, as segments occurring at the beginning of a lower-level constituent are in medial position in a higher-level constituent. For example, the consonant /m/ in *clémentine* is in initial position in the Syllable /mã/ and in medial position in the Word. By contrast, /k/ is in initial position both in the Syllable and in the Word. Furthermore, this kind of comparison makes it possible to observe whether the articulation of a segment varies according to the prosodic level of the constituent in which it is in initial position, and thus *whether segmental articulation reflects the prosodic organization of speech into different levels of constituents.*

As early as 1901, Rousselot noted that the nature of morpho-syntactic groups and their boundaries exerts a constraint on the articulation of segments, and that one must take into account the

> "*many variations incurred by a single consonant because of its position within the word or the sentence*" [Rousselot 1901:601, our translation, emphasis added].

This section presents the few studies which are concerned with the articulatory characteristics of initial segments above Word-level, and which compare different levels of constituents. The number and nature of the constituents compared vary depending on the studies and the languages studied. They include Accentual Phrases, Phonological Phrases, Intermediate Intonational Phrases and Intonational Phrases. This presentation is organized by articulator.

α. High-level constituents: Glottal articulation

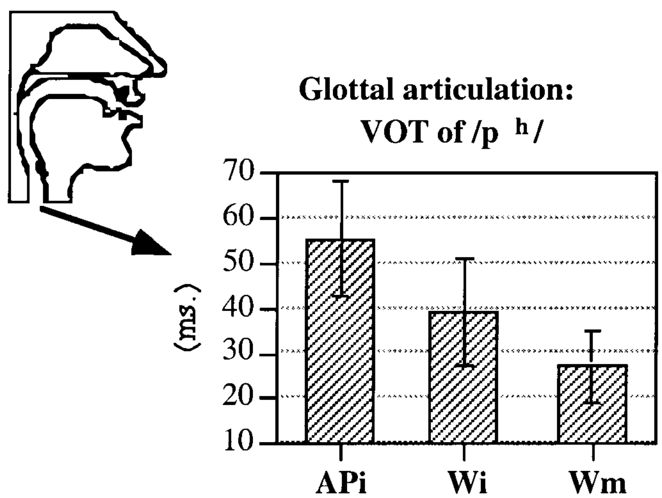
Changes in the glottal articulation of segments have been observed as cues to prosodic boundaries in several languages. The results found in the literature are generally based on the acoustic correlates of these articulatory variations (amplitude of the aspiration noise, VOT duration, aperiodicity, etc.)

Pierrehumbert and Talkin (1992) show that the glottal articulation of the aspirated consonant /h/ in English is different depending on whether it is found at the beginning of an Intonational Phrase or Word-initially. The aspiration noise of this consonant (RMS energy compared to the energy of the following vowel) is weaker and longer at the beginning of an Intonational Phrase. The consonant /h/ is articulated with an abduction of the vocal folds, and thus has less energy than the following voiced vowel. At the beginning of an Intonational Phrase, the decrease in energy of /h/ is thus interpreted by the authors as a reinforcement of its consonantal character. Pierrehumbert and Talkin also provide data for /h/ Word-initially and Word-medially but do not compare them directly. Their charts (4.6-4.7 v. 4.8-4.9) show that the aspiration noise in Word-initial /h/ is weaker than in Word-medial /h/ (whose energy is closer to the vowel's). This can be observed essentially in the case of stressed /h/ (which goes against Goldstein (1992)'s conclusions in his commentary on the article). The aspiration noise of /h/ thus appears to decrease progressively between the middle of a Word (Syllable-initially), the beginning of a Word and the beginning of an Intonational Phrase.

The duration of the VOT of voiceless consonants also varies according to the prosodic level of the constituent in which they are in initial position. Whereas Lisker & Abramson (1967) do not observe any difference in VOT between the beginning and the middle of an Utterance, Pierrehumbert & Talkin (1992) show that the VOT of the aspirated consonant /t^h/ in English is longer at the beginning than in the middle of an Intonational Phrase. Jun (1993) also observes variations in the VOT of aspirated consonants in Korean according to their prosodic position. Her results are shown in figure 1.7. The duration of VOT increases gradually from the Word-medial position (Wm) to the Word-initial position (Wi) to the beginning of an Accentual Phrase (APi). In their study of Taiwanese, Hsu & Jun (1997) make similar observations. The VOT of the aspirated consonant /k^h/ (positive VOT) and that of the voiced consonant /b/ (negative VOT) vary according to their prosodic position. VOT makes it possible to distinguish between three positions and increases from the Syllable-initial (Si) position to the Word-initial position (Wi) to the Intonational-Phrase-initial position (IPi). In contrast, the VOT of the unaspirated voiceless consonant /t/ is not affected by the prosodic position of the consonant in Taiwanese.

Variations in VOT according to prosodic position can result from a change in the glottal gesture both spatially and temporally. A longer VOT may be the consequence of a wider glottal gesture, but also the consequence of an increased lag between the glottal opening gesture and oral closure gesture [see Cooper 1991, Goldstein 1992].

Figure 1.7: VOT duration for the aspirated consonant /p^h/ in Korean Accentual Phrase-initially (APi), Word-initially (Wi) and Word-medially (Wm). This summarizes the results in Jun (1993) based on her figure 6.2.



Prosodic position also affects the glottal articulation of vowels. As seen in the previous section, Word-initial vowels may be glottalized in English. This glottalization is generally considered to be an optional allophonic variation. More recently, works inspired by a theory of prosodic structure have shown that there exists a strong link between prosodic structure and the frequency of glottalization. For instance, Pierrehumbert & Talkin (1992) show that a Word-initial vowel (regardless of stress) is more frequently glottalized when it also occupies the initial position in an Intonational Phrase than when it occurs inside the Intonational Phrase. Dilley et al. (1996) study the effect of prosodic position on the occurrence of glottalization in a large corpus of continuous speech taken from a radio program database in English. The corpus is comprised of five speakers and contains 3709 cases of glottalization of initial vowels. The prosodic structure of the utterances is transcribed using the ToBI system [see Silverman et al. 1992]. Three types of prosodic positions emerge: beginning of a Full Intonational Phrase (marked by a boundary tone and a break index of 4, following the ToBI notation); the beginning of an Intermediate Intonational Phrase (marked by a break index of 3 and no boundary tone); the beginning of a Word (within an Intermediate Phrase). The presence and position of the phrase accent and lexical stress are also considered as factors of variation. One of the questions raised in this study is whether the frequency of glottalization is higher in important prosodic positions (e.g. at the beginning of an Intonational Phrase or on the Word bearing Phrase accent). Their results confirm Pierrehumbert & Talkin (1992)'s observation that vowels are more frequently glottalized when they occur at the beginning of an Intonational Phrase (and when they bear phrase accent). Vowels occurring at the beginning of an Intermediate Intonational Phrase are also more frequently glottalized than vowels occurring at the beginning of a Word. Only unstressed vowels occurring at the beginning of an Intonational Phrase are more frequently glottalized than vowels occurring at the beginning of an Intermediate Intonational Phrase. When they bear lexical stress, the difference between these two positions is neutralized. The authors conclude that the occurrence of glottalization in initial vowels is conditioned by the prosodic structure of an utterance.

The occurrence of glottalization as a cue to higher prosodic boundaries than the Word has been observed in other languages than English. For instance, Huber (1988) distinguishes different types of laryngealization in Swedish, the occurrence of which depends on the nature of the boundary: (1) creakiness and creaky voice, which occur mostly at the end of an Intonational Phrase, especially with female speakers; (2) diplophonia (the alternation of strong and weak glottal pulses), also used by female speakers, either at the end of an Intonational Phrase or Word-finally inside an Intonational Phrase; (3) glottalization, which occurs only at the initial boundary, regardless of the speaker's gender. This glottalization is used to mark the beginning of constituents which are internal to the utterance, either the beginning of an Intonational Phrase or the beginning of internal constituents like a Clause. Lehiste (1964) also notes the occurrence of glottalization of an initial vowel in Serbo-Croatian to mark an initial boundary in an Accentual Phrase (an accentual unit comprised of a proclitic and a noun).

In summary, whether one considers variations in VOT or in aspiration noise, glottal opening in consonants seems wider at the beginning of a prosodic constituent and this opening increases with the hierarchical level of the constituent. In the case of initial vowels, the occurrence of glottalization also increases with the hierarchical level of the constituent.

β. Higher-level constituents: Labial articulation

The variations in labial articulation in higher-level constituents have been studied in English by Byrd & Saltzman (1996, 1998) for the consonant /m/. They study sequences of the form /mV#m/ in which different types of syntactic boundaries are inserted before the second consonant (word boundary, boundary between items in a list, post-vocative boundary, phrase boundary). They observe that the three speakers distinguish between at least two or three boundary levels with the articulation of the initial consonant: the duration and amplitude of the labial closure tend to increase in higher positions. The duration of the labial closure is also more variable at the beginning of higher-level constituents. In Tamil, Byrd et al. (1996) show that labial closure is longer in an /m/ occurring at the beginning of the highest constituent among those studied ("Large Phrase" vs. "Small Phrase" and "Word"). However, in Tamil this lengthening is not linked to a change in amplitude of the gesture.

With respect to vowels, van Lieshout et al.(1995) have shown a variation in labial muscular activity in Dutch rounded vowels according to their prosodic position. These vowels do not occur in absolute initial position but in the initial syllable of a CVC word occurring either at the beginning or at the end of an Utterance. The authors observe that the muscular activity of the lips is greater and longer Utterance-initially. The muscular activity is greater not only at the beginning of the vowel but throughout the CVC word.

γ. Higher-level constituents: Lingual articulation

In a previous study of English [Fougeron & Keating 1995, 1997] we observed that the lingual articulation of /n/ varies according to the prosodic position of the consonant. The corpus, using reiterant speech with the syllable /no/, is modeled after algebraic expressions of the type "(x+x+x)*x=y", in which the position of parentheses and operators is variable. The numbers imitated ("x") are trisyllables in which lexical stress falls on one of the three syllables (*eighty-nine*, *seventy* and *one-hundred*). Based on a prosodic transcription of tones and boundaries, five prosodic constituents appeared: the Syllable (S), the Word (W), the Phonological Phrase (PP), the Intonational Phrase (IP) and the Utterance (U)[11]. The /n/s occurring in the reiterant syllable /no/ are coded according to their initial position in these constituents (Si, Wi, PPi, IPi and Ui). In parallel, the /o/s occurring in the syllable /no/ are coded according to their final position in these constituents (Sf, Wf, PPf, IPf and Uf). The degree of linguopalatal contact (determined by electropalatography) for the consonants /n/ and for the vowels /o/ is compared according to the prosodic position of the segment for three American speakers.

In a first analysis, from a syntagmatic perspective, we compared the articulation of segments according to their position (initial, medial and final) at each constituent level. The results, summarized in figure 1.8, show more linguopalatal contact in consonants in initial position than in medial or final position. The increased contact in initial position appears at all constituent levels and for all speakers, except at the Word level for speaker 1. In the case of vowels (not shown), one observes an increase in aperture/backness of /o/ in final position compared to the medial position at all prosodic levels.

[11] The reader is referred to the 1997 article for a detailed explanation of the prosodic coding of these constituents.

Figure 1.8: Degree of linguopalatal contact for /n/ in English according to its position: initial (i), medial (m) and final (f) in the Utterance, the Intonational Phrase, the Phonological Phrase and the Word. Results for 3 American speakers. From Fougeron and Keating (1997), figure 3.

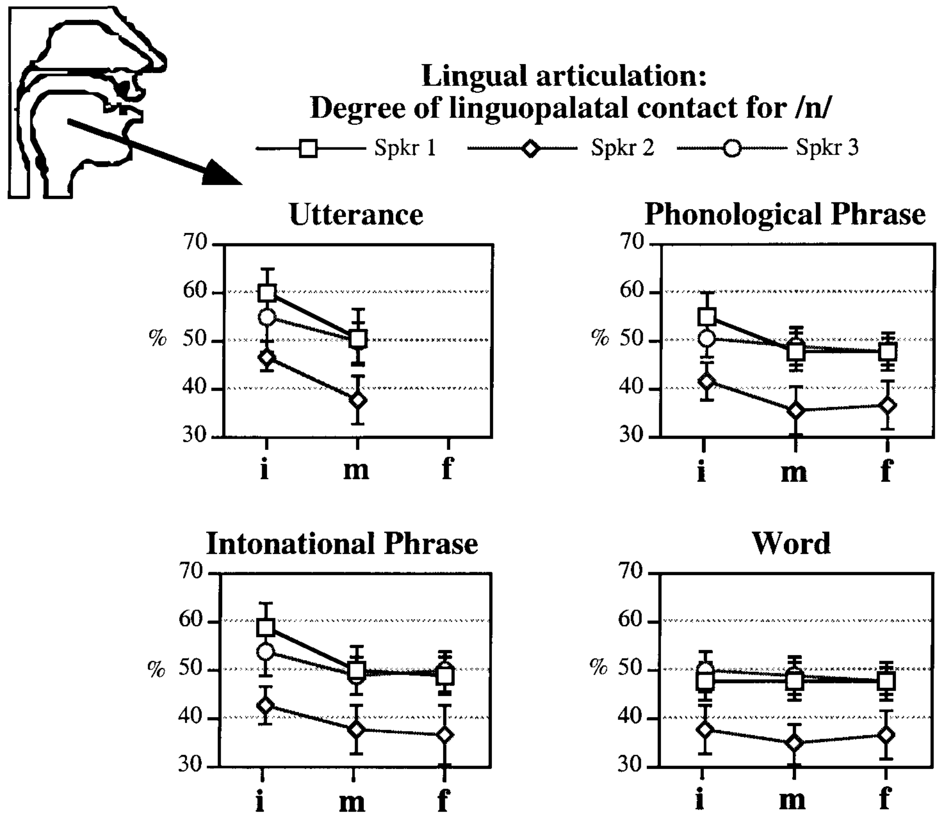
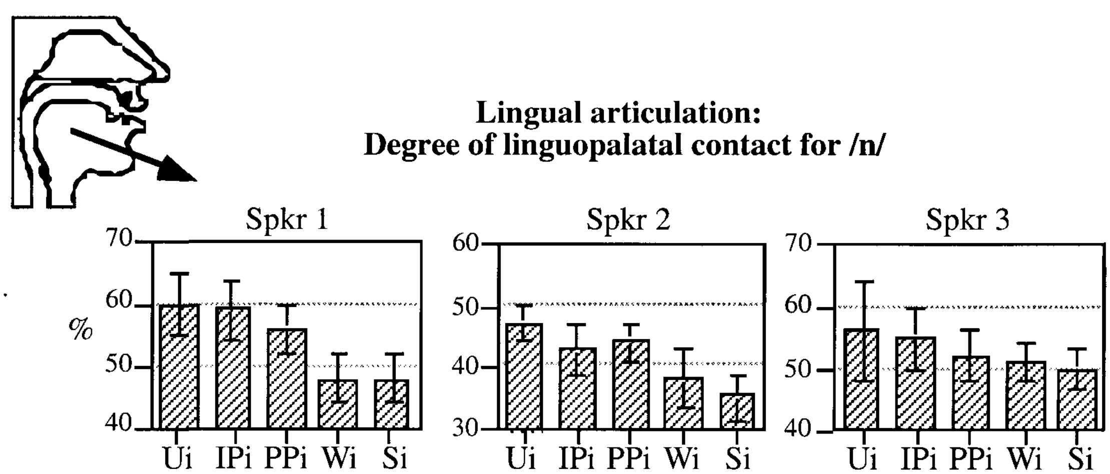


Figure 1.9: Degree of linguopalatal contact for /n/ in English in initial position at different prosodic levels: Syllable (Si), Word (Wi), Phonological Phrase (PPi), Intonational Phrase (IPi) and Utterance (Ui). Results for 3 American speakers. From Fougeron & Keating (1997), figure 4.



In a second analysis, from a paradigmatic perspective, we compared the initial position (for /n/) and the final position (for /o/) between the five constituent levels. The goal of this analysis was to determine whether the effect of the prosodic position is hierarchical and progressive along the constituent hierarchy. Thus consonants occurring at the beginning of a Syllable (Si), of a Word (Wi), of a PP (PPi), of an IP (IPi) and of an Utterance (Ui) were compared. Final vowels (CV) are compared in a similar way in all constituents. Only those phrases fashioned after algebraic expressions that contained the number *eighty-nine* (final stress) were included in the comparison. The results show that linguopalatal contact in the consonant increases with the hierarchical level of the prosodic constituent. However, the number and nature of the constituents distinguished by this increase in contact vary according to the speaker. These results are shown in figure 1.9 for each speaker. The positions distinguished by increased contact (">") are as follows:

Speaker 1:	IPi>PPi>(Wi, Si)	(with Ui=(IPi, PPi))
Speaker 2:	Ui>(IPi, PPi)>Wi>Si	
Speaker 3:	(Ui, IPi)>(PPi, Wi)>Si	

In summary, for all three speakers, three or four prosodic positions are distinguished by increased linguopalatal contact in the consonant. The three speakers distinguish at least between the highest (Ui-IPi) and the lowest constituents (Wi-Si). In the case of final vowels, the variations in linguopalatal contact observed are constrained by the experimental technique: beyond a certain degree of aperture, linguopalatal contact does not appear and electropalatography does not make it possible to distinguish wider degrees of aperture. For two of the three speakers, an increase in the aperture of the final vowel distinguishes three constituent levels (IPf>PPf>(Wf, Sf) for speaker 2 and (IPf, PPf)>Wf>Sf for speaker 3).

These results show that the degree of linguopalatal contact is not a particular cue linked to a specific prosodic position. The variations in the amount of contact are, on the contrary, progressive and vary with the speaker, albeit in a consistent direction: contact tends to increase in initial consonants in higher prosodic constituents. To a lesser extent, vowels tend to be more open at the end of higher prosodic constituents. Consequently, prosodic constituents are bounded to the left and to the right by more open vowels and consonants with a narrower closure.

In English, Keating (1995, 1997), using a non-reiterant speech corpus, confirms the increase in linguopalatal contact for various consonants occurring Utterance-initially compared to consonants occurring Utterance-medially. Variation in linguopalatal contact at different prosodic levels in Taiwanese for the consonant /t/ is also shown by Hsu & Jun (1997), who compare the beginning of the Intonational Phrase (IPi), of the Word (Wi) and of the Syllable (Si). One of their two speakers distinguishes the three positions Si, Wi and IPi by a gradual increase in contact, whereas the other speaker only distinguishes two positions, with IPi and Wi both exhibiting more contact than the Si position. In Korean, Cho (1998) also shows a gradual increase of linguopalatal contact between the initial position in a Syllable, a Word, an Accentual Phrase, an Intonational Phrase and an Utterance. This distinction is made by three speakers for unaspirated consonants /t/ and /n/, except for the distinction between Wi and APi for one speaker. In Tamil, Byrd et al. (1996) observe a lengthening of the lingual closure (closure gesture) for an /n/ occurring at the beginning of a higher-level prosodic phrase ("Large Phrase" vs. "Small Phrase" & "Word"). This lengthening is accompanied by a decrease in stiffness and a small increase in the amplitude of the lingual closure gesture.

Position within a higher-level constituent also affects the articulation of coda consonants. In English, coda consonants occurring Utterance-finally exhibit more contact than those occurring Utterance-medially. Consequently, the difference between onset and coda is smaller Utterance-finally [Keating 1995, 1997].

δ. Higher-level constituents: Nasal articulation

In the case of nasal and oral consonants, we have seen that the velum is higher in Syllable- or Word-initial position than it is in medial or final position. Few data can be found in the literature about the behavior of the velum in higher-level constituents. For example, Krakow et al. (1991, 1994) have shown with the Velotrace that the velum is also higher for an oral consonant occurring Utterance-initially.

Gordon (1996) compares the nasal articulation of /n/ in Estonian at different constituent levels. The variations observed are inferred indirectly from the nasal airflow during the consonant. Four positions are studied: Utterance-initial (Ui), the beginning of an Intonational Phrase (following a pause) (IPi), Word-initial (Wi) and Syllable-initial (Si), with three or four speakers. Gordon observes a progressive decrease in nasal airflow from Si to Wi to IPi. The Ui position, however, shows this tendency with one speaker only, whereas the nasal airflow is greater Utterance-initially than at the beginning of an Intonational Phrase for the two other speakers. With the exception of the Utterance-initial position, Gordon shows that the decrease in nasal airflow at the beginning of a constituent follows constituent hierarchy. If the decrease in nasal airflow is interpreted as a decrease in the velopharyngeal opening due to a raising of the velum, these results suggest that the velum is higher at the beginning of a higher-level constituent.

ε. Variations in higher-level constituents: Conclusion

The observations mentioned above show that the articulation of segments is also influenced by their position in higher prosodic constituents than the Word or the Syllable. The initial position in these constituents is generally marked articulatorily in comparison with non-initial positions at the constituent level being examined. Additionally, when a comparison is made between the different constituent levels, the articulatory variations in initial position (frequency of glottalization, glottal opening, labial or lingual constriction or velum raising) tend to increase with the hierarchical level of the constituent. However, not all prosodic constituents are uniformly marked, and this marking may vary with the speaker.

B.3. Variations in the coordination of articulatory gestures according to prosodic position

We have seen that prosodic position affects the spatial characteristics of articulatory gestures. We will now see that it also affects the relative timing of these articulatory gestures.

B3.a. Coordination in the Syllable or the Word

The syllable is often considered to be the basic linguistic unit in which the coordination of articulatory gestures is specified [Kozhevnikov & Chistovich 1965, Browman & Goldstein

1995]:

> *syllable structure is a characteristic pattern of coordination among gestures, certain types of variation are automatic consequences of this pattern of coordination, and therefore, necessary correlates of syllable structure. [Browman & Goldstein, 1995 : 20]*

According to the position of a segment within a Syllable or a Word, several studies have shown that the coordination of articulatory gestures varies within segments (3.a.α) or between segments (3.a.β).

α. Coordination of different articulators within segments

Krakow (1989) shows that the coordination of the velar and labial gestures for a labial nasal consonant in English is affected by the position of the consonant. It is precisely this difference in coordination that most distinguishes an initial (onset) and a final (coda) consonant in her corpus. Word-initially, the oral and the nasal gestures are synchronized: the end of the velum lowering gesture and the end of the raising of the lower lip appear at the same time. However, Word-finally, the end of the velum lowering gesture is synchronized with the beginning of the raising of the lower lip. Thus, Word-finally, the velar gesture is not only greater but earlier relatively to the oral gesture. Farnetani (1986) finds similar results in Italian, comparing not Word-initial and final consonants but Word-initial and medial onset consonants. Word-medially, the increase in nasal airflow for /n/ begins before the oral closure (it occurs during the preceding vowel). However, Word-initially (following a pause), the nasal airflow begins to increase after the lingual closure is completed. The vocal fold vibration begins at the same time.

Another well-known variation in the coordination of articulatory gestures occurs in the English consonant /l/. In post-vocalic and pre-consonantal final position, /l/ appears as a "dark" allophone (less fronted tongue body), whereas in initial and post-consonantal position, /l/ appears as a "light" allophone. The difference between these two allophones is due to a difference in the coordination of the tongue tip gesture and the tongue body gesture. In initial position, the retraction of the tongue body closely follows the raising of the tongue tip. On the other hand, Word-finally, the retraction of the tongue body precedes the tongue tip closure [Sproat & Fujimura 1993]. The lag between the two gestures varies according to the prosodic boundary: the lag is longer for Word-final /l/ (as well as at the end of an Intonational Phrase) than for /l/ in coda position. Recasens & Farnetani (1990) confirm that the velarization of /l/ in final position is not categorical, but a gradient phenomenon which varies according to the language (greater in Catalan and American English than in Italian) and the segment's position.

β. Coordination of gestures between segments

Position within a Word or Syllable also affects the coordination of gestures associated with different segments.

Browman & Goldstein (1995) define syllabic organization in terms of alignment of consonantal gestures with vocalic gestures. Initial consonants are synchronized with the beginning of the vocalic gesture and final consonants are synchronized with the end of the vocalic gesture. This develops an idea in Kozhevnikov & Chistovich (1965), who claim that the articulatory syllable is CV. Coarticulation within a syllable thus gives it its unity. For example, Lehiste (1964) shows that progressive nasalization in an NV sequence is stronger if the vowel is in the same syllable as the nasal segment (e.g. stronger in /V.NV/ than in /VN.V/).

Other studies emphasize the fact that the coordination of vocalic and consonantal gestures can also define the rhyme (VC) as a domain with stable timing [e.g. Lehiste 1970 (in English) or Slis 1971 (in Dutch)]. This idea is shared by Fujimura & Erickson (1997) who explain that the articulatory gestures of sonorant consonants in coda position are more closely linked to the tautosyllabic vowel than with the onset consonant. In initial position (e.g. *lap, lip*), they show a discontinuity in the spectrogram between /l/ and the following vowel, with faster formant transitions. In final position, on the other hand, this discontinuity does not appear and the formants in the stable portion of the vowel are strongly affected by the characteristics of /l/. Temporal compensation phenomena within the rhyme also highlight the coordination of the nucleus and coda gestures. For example, Lehiste (1970) shows several cases of temporal compensation between segments. She attempts to establish what the domain of the temporal compensation is, assuming that it is an "articulatory unit," i.e. the domain of the articulatory program. She shows that in English this domain corresponds to the VC syllable: the duration of a vowel and that of the following consonant in a monosyllable are strongly correlated; if the vowel is lengthened, the consonant is shortened. In Icelandic, Norwegian and Swedish, there is also a negative correlation between the duration of the vowel and that of the consonant in a VC syllable.

These results show that the overlap of articulatory gestures highlights the temporal cohesion of the gestures associated with the segments in a syllable. However, these results do not allow us to determine whether there is more cohesion between an initial consonant and the following vowel or between a coda consonant and the preceding vowel. If such is the case, the distinction between these two types of cohesion (C-V and V-C) does not merely reflect a positional difference (initial vs. final) but most importantly a difference in the direction of coarticulation: anticipatory-perseverative [see Fougeron 1993 for a discussion of the mechanisms involved in these two types of coarticulation].

The results found in Byrd (1994, 1996) provide more insight into the specificity of temporal organization Word- or Syllable-initially. In the case of English, she shows that there is less overlap of consonantal gestures in a CC sequence when the two consonants are the onset of a syllable (CCV) than when they are in coda position (VCC) or heterosyllabic (C.C). Additionally, the coordination of these gestures is much less variable for an onset consonant cluster. She concludes that the timing relations for these consonants are more precise and stable in this position. Gendron (1966) suggests that there is less coarticulation between C and V in Word-initial position. He explains in these terms the small frequency of devoicing in Québec vowels preceded by a voiceless consonant when they occur in the first syllable of a Word (20% of cases) compared to a Word-internal syllable (50% of cases).

B3.b. Gestural timing in higher-level constituents

Numerous studies have shown that articulatory overlap between segments is strongly affected by the weight of the prosodic boundary between them. As I am more particularly concerned with initial segments, I present studies showing that the anticipation of articulatory gestures for an initial segment depends on the level of the preceding prosodic boundary.

In a cinefluorographic study, MacClean (1973) shows that the level of the syntactic boundary between the two vowels of a CVVN sequence in English affects the temporal alignment of the velum gesture. In the case of major boundaries (prosodically marked, between phrases or clauses), the beginning of the lowering of the velum for /n/ is delayed with respect to the preceding vowel: it coincides with the beginning of the second vowel (CV_1V_2N). Thus there is less nasal anticipation. In the case of minor boundaries (not prosodically marked between words, phrases and clauses), the lowering of the velum begins at the beginning of the articulatory gesture of the first vowel. Anticipation of the nasal gesture thus extends to the two preceding vowels, and this is independent from the duration of the CV sequence. The data in Vaissière (1988) also show a decrease in nasal anticipation for /n/s that are separated from an oral consonant by a pause (C#N). In this position (after a pause, beginning of an Intonational Phrase), the lowering of the velum for the initial nasal consonant does not begin before the beginning of the oral constriction. Thus, compared to other positions, there is no anticipation of the lowering of the velum in the preceding consonant.

The effect of boundary level has also been observed on the relative timing of tongue gestures. Hardcastle (1985) looks at the relative timing of tongue gestures for the sequence of consonants /kl/ according the the level of the syntactic boundary between the two consonants. Even though there is a rather high variability between his four speakers, he concludes that the least favorable condition for overlap between gestures for /k/ and /l/ is when the consonants are separated by a major syntactic boundary (and when the speech rate is slow). Holst & Nolan (1995) also show that assimilation of place of articulation between the fricatives /s/ and /ʃ/ is reduced when these consonants are separated by a "major clause boundary" compared to a Word boundary.

The decrease in overlap between gestures across a major boundary is not due only to the presence of a pause. Holst & Nolan (1995) observe an absence of total assimilation between /s/ and /ʃ/ even with a major boundary that is not marked by a pause (assimilation is partial). Byrd et al. (1996) also show a decrease in overlap between labial and lingual gestures for a /n+m/ or an /m+n/ sequence when the consonants are separated by a major boundary as opposed to a lower-level boundary when none of the boundaries is marked by a pause. In addition, the presence of a pause does not prevent overlap between segments. For example, Lewis et al. (1975) show that the dental fricative /θ/ triggers a fronting of /n/ even when the consonants are separated by a pause.

B3.c. Variations in gestural timing: Conclusion

These studies show that there is greater synchronicity between the gestures involved in the articulation of a segment (velum and lips for /m/ or tongue body and tip for /l/) when this segment is Word- or Syllable-initial (Browman & Goldstein 1992, 1995). In addition, relative gestural timing in a Word-initial sequence of segments (CC or CV) appears to be governed by a stable temporal template which is more constrained than in other positions. The relative timing of articulatory gestures also varies in higher-level constituents. The degree of overlap of gestures is negatively correlated with the level of the boundary separating them (Browman 1995). Anticipation of the gestures of an initial consonant decreases when the consonant is preceded by an major boundary (i.e. when the consonant is initial in a high-level constituent), whether this boundary is marked by a pause or not. Thus, the greater synchronicity between gestures for an initial segment observed at the Word level also appears to affect initial segments in higher-level constituents since there is less anticipation in their gestures.

B.4. Variation in acoustic duration as a function of a segment's position in prosodic constituents at different levels

We have seen that a segment's position in a prosodic constituent affects its articulation. Articulatory variations are observed in the amplitude and the duration of articulatory gestures as well as their relative timing. Consequently, the acoustic duration of segments will also vary according to their prosodic position. I will briefly sketch the best-known lengthening phenomenon due to prosodic position: final lengthening (B.4.a). Then I will survey a few studies showing the existence of initial lengthening of segments and the fact that this lengthening may depend on the prosodic level of the constituent.

B4.a. Lengthening in constituent-final position

Lengthening of final units in a sequence seems to be a natural tendency of any (high-level) motor activity (see e.g. Vaissière 1983). It appears in music (Cooper 1976) and in sign language (Wilbur & Zelaznik 1997).

At the Word level, final lengthening of consonants and vowels has been observed in numerous studies and languages (see, e.g., Delattre 1966 for French, Spanish, German and English, Lindblom 1968, Lindblom & Rapp 1973 for Swedish, Lehiste 1972, Klatt 1975, Crystal & House 1990 for English).

In higher-level constituents, final lengthening of the last syllable in the constituent tends to increase gradually as a function of the hierarchical level of the constituent, i.e. as a function of the level of the following boundary (see, among others, Delattre 1966, Grosjean & Deschamps 1972, Crompton 1980, Rietveld 1980, Duez 1987, Pasdeloup 1990, Fletcher 1991 in French, Klatt 1975, Cooper & Paccia-Cooper 1980, Wightman et al. 1992 in English). According to these studies, several degrees of boundary are specified by final lengthening. In a recent study, Wightman et al. (1992) show that in English four prosodic levels can be distinguished by lengthening of the final vowel (the Word and three higher-level constituents). In Ladd & Campbell 1991, final lengthening distinguishes at least 4 constituent levels above the Word. In French, three prosodic levels are generally distinguished with final lengthening: the Word, the Accentual Phrase (prosodic word or rhythmic group) and the Intonational Phrase.

B4.b. Constituent-initial lengthening

As in the case of final lengthening, the duration of initial segments is also affected by their prosodic position.

At the Word and Syllable level, we have seen that the amplitude and duration of articulatory gestures of consonants are greater in initial than in non-initial position. Acoustically, the duration of initial consonants is also longer. This Word-initial lengthening has been observed in several languages (see, among others, Vaissière 1977, 1989 in French, Kohler & Hardcastle 1974, Kohler 1976 in German, Nooteboom 1972, Quené 1992 in Dutch, Carlson & Grandström 1973, Lindblom & Rapp 1973 in Swedish, Nespor 1977 in Italian, Lehiste 1960, Oller 1973, Umeda 1977, Nakatani & Schaffer 1978 in English, Lehiste 1964 in Czech). For example, in English, Oller (1973) studied the variation in vowel and consonant duration according to their phrasal position and stress. He shows that, in addition to final lengthening, word-initial consonants are lengthened, whether the syllable is stressed or not. For unstressed consonants, Word-initial lengthening is greater (30ms) than Word-final lengthening (20ms). Stressed consonants only exhibit initial lengthening and no final lengthening. Vowel duration (CV or CVC) is only affected by stress and final position. Vaissière (1988, 1989) shows that the initial consonant of a prosodic word is lengthened in French. Compared to final syllables (which are long, or extra-long phrase-finally) and medial syllables (which are short), initial syllables have an average duration (which is lengthened by the consonant only). In Dutch, Quené (1992) shows that the lengthening of an initial consonant is an acoustic cue to a lexical boundary (#) in VC#V (short C) and V#CV (long C) sequences.

In our study of English (Fougeron & Keating 1997, summarized in B2.c), we observed that the lengthening of an initial consonant can also be a function of the constituent level and therefore of the preceding prosodic boundary (as in the case of final lengthening). Two out of three speakers distinguish 4 prosodic levels by lengthening of initial /n/. Lengthening increases from Syllable initial (Si) to Word-initial (Wi) to Phonological Phrase initial (PPi) to Intonational Phrase initial position (IPi). The third speaker only distinguishes 3 prosodic levels by progressive lengthening from Si to Wi to PPi-IPi position. Consonant duration utterance-initially shows the most variability. In Taiwanese, Hsu & Jun (1997) also observe progressive lengthening of the acoustic duration of /t/ closure from Si to Wi to IPi for both speakers.

B4.c. Variations in acoustic duration: Conclusion

Final lengthening is a cue to prosodic demarcation between constituents, which varies according to the weight of the following prosodic boundary. Lengthening is also observed in constituent-initial segments. This lengthening has essentially been described at the Word level for consonants, but a few results suggest that initial lengthening can also be a cue to a prosodic boundary, being a function of the boundary's weight.

B.5. Variation as a function of linear position within the phrase: articulatory declination?

Independently of prosodic position, some authors have suggested that the articulation of segments is a function of their "linear" position in the phrase. Segments would have a more extreme articulation early in the phrase, which would gradually decline until the end of the phrase. This tendency is similar to the phenomenon of declination observed in fundamental frequency or sub-glottal pressure (see, e.g., Ladd 1984 or t'Hart et al. 1990 for a survey). It has been described as "articulatory declination" or "supra-glottal declination" (Vayra & Fowler 1992, Krakow et al. 1994).

Vaissière (1986b) notes that the velum is generally very high utterance-initially in English. She suggests that this is due to increased tension in all articulators at the beginning of a phrase. Later articulations would then undergo a general weakening or decrease in energy until the end of the phrase. Vatikiotis-Bateson & Fowler (1988) for English and Vayra & Fowler (1992) for Italian show articulatory data supporting a declination of articulatory gestures in 2- or 3-syllable sequences. Krakow et al. (1994) tested this hypothesis on a corpus of reiterant speech that includes longer and more syntactically complex phrases. The position of the velum for the consonant /t/ is tracked by a velotrace at several points within the phrase. The authors find that the position of the velum is always higher phrase-initially than phrase-finally. In lexically stressed syllables, the velum goes down gradually in successive positions within the phrase. In lexically unstressed vowels, little or no lowering is observed. The authors conclude that articulatory declination does occur from the beginning to the end of a phrase and that it mainly affects stressed syllables.

However, a detailed examination of the data in Krakow et al. (1994) shows that, if declination is present from the beginning to the end of a phrase, local variations can be observed throughout the phrase in medial syllables. In a previous study [Fougeron & Keating 1996] we tried to replicate Krakow et al.'s findings with EPG and to infer the prosodic organization of the phrases used by these authors. Our results suggest that these authors' observations result from the prosodic position of segments in their phrases. In their corpus, segments appearing early in the phrase are also initial in a higher prosodic constituent. The segments appearing later in the phrase are also initial (or medial) in a lower prosodic constituent. Thus variations in velum height can not only be the consequence of the segment's linear position within the phrase but also be the consequence of the segment's position within different prosodic constituents.

These results do not completely invalidate the hypothesis of articulatory declination. Conversely, the existence of articulatory declination does not invalidate the variations observed according to a segment's prosodic position. Some articulatory declination may occur within certain prosodic constituents and be reset at the beginning of a new constituent (in a manner similar to F0 reset). The constituents delimiting the domain of this declination yet have to be defined.

II.C. ARTICULATORY VARIATIONS AS A FUNCTION OF PROSODIC POSITION: CONCLUSION

The results shown in this section suggest that a segment's articulation is influenced by prosodic structure. A segment's position within a constituent and its stress as well as its position relative to stress, influence its articulation.

While most studies in the literature are concerned with segmental variations in final position, we have seen that segments in initial position also articularily have particular properties. The few studies comparing different prosodic levels show that *this effect appears not only at the Syllable and Word level, but also in higher-level constituents.* Thus initial segments are different from non-initial segments within a constituent, and from initial segments within lower-level constituents.

Articulatory variations in initial position converge: *any articulatory gesture at the beginning of a Syllable or a Word tends to strengthen at the beginning of higher-level constituents, following prosodic hierarchy.* Such a gesture may be a feature of the segment (e.g. raising of the tongue for stops) or an "unexpected" articulatory feature (e.g. raising of the velum in nasals or glottalization of vowels). Prosodic position also affects the relative timing of articulatory gestures as well as the acoustic duration of segments.

The articulatory variations according to prosodic position that we have seen so far are subtle phonetic details that do not change the nature of the segment. For example, a stop remains a stop in all positions even if the degree of stricture increases. However, as we shall see in the next section, some of these positional characteristics may have been phonologized. The integration of these articulatory variations into a language contributes to the special linguistic status of the initial position.

III. CONSTITUENT-INITIAL POSITION: A SPECIAL LINGUISTIC STATUS?

The articulatory variations surveyed so far suggest that the initial position in a constituent has a special status in production. This special status is also apparent when one considers the *linguistic behavior* of segments in this position. Initial segments are particularly impervious to diachronic (see section III-B and synchronic (section III-C) variations across languages. In addition, initial segments play a particularly important part in speech processing (see section III-D). We will see that these properties can be interpreted as a phonologization of the articulatory (phonetic) characteristics of this position.

III.A. Word-initial position: a "strong" position

Traditionally, final and unstressed positions in a Word are considered as *weak* positions. Conversely, Word-initial position and stressed position are described as *strong* positions. According to Vaissière (1986a, 1988), these latter positions can be specified by a prosodic feature [+strong] which would account for the articulatory variations observed in these two positions (velum raising, glottal constriction). In French, the initial syllable of a Word can bear initial stress. Although initial stress is not always realized (see section II.A), the initial position in French can thus be considered to be doubly strong because it is initial and it can carry stress.

Examination of the kind of diachronic and synchronic variation that segments can be subject to as a function of their position indicates that such a distinction between positions is not merely binary (strong vs. weak) but also scalar: *some positions are stronger than others*. For example, Straka (1964) proposes a strength scale between positions which governs the diachronic evolution of segments. The following table summarizes the positions on the scale for consonants and vowels, from strongest to weakest:

Consonants	**Vowels**
1. Word-initial [#C]	1. Syllable-initial before C [.VC]
2. Syllable-initial after C [C.**C**]	Syllable-medial between C [.CVC.]
3. Word-final [**C**#]	2. Syllable-medial (without C) [.**V**.]
4. Syllable-initial after V [V.**C**V]	3. Syllable-final after C [CV.][CV#]
5. Syllable-final before C [**C**.C]	

(Note: positions 3 and 4 for consonants can be switched)

III.B. Initial position in diachronic evolution. Evidence from historical linguistics

B.1. An example from the diachronic evolution from Latin to French

A segment's resistance to diachronic change from Latin to French mainly depends on its position within a Syllable or a Word. Segments in strong positions were preserved, whereas segments in weak positions were reduced or disappeared [Brunot & Bruneau 1937, Martinet 1955, Bourciez & Bourciez 1967, Zink 1986].

The strength hierarchy between positions is illustrated clearly by the evolution of the Latin word *factum* into the French word *fait* as shown in Haden (1938). The word *factum* exhibits a strength difference between the stressed syllable (*fac*) and the unstressed syllable (*tum*), which is weaker. There is also a difference between the initial consonants (/f/ and /t/), which occur in strong positions, and the final consonants (/k/ and /m/), which occur in a weak position within the syllable. The final consonant in the unstressed syllable is in the weakest position, and so it is lost first: ['faktum] becomes ['factu]. The final consonant in the stressed syllable is now in the weakest position. It is altered, then absorbed into the vowel: ['faktu] becomes [façt'] then [fajt] then [fɛt]. /t/ thus becomes final and is dropped, since it is in a weak position, resulting in the French word [fɛ]. The initial consonant in the stressed syllable (which is also Word-initial) remains unchanged.

The general trend of consonant change can be summarized as follows (from Bourciez & Bourciez 1967):

- Consonants may remain intact. This is the case for word-initial consonants (except velars /k/ and /g/, which are altered before front vowels) and medial consonants following another consonant.
- Consonants may change. This is the case for medial consonants, which undergo various changes depending on the consonant's class (lenition or deletion). A medial consonant occurring before another consonant is preserved only if the cluster is easy to pronounce, otherwise it weakens or assimilates.

Final consonants must be differentiated:

- Before the end of an utterance or a pause, the consonant is in a strong position and it is preserved (e.g. *J'en ai six* [sis]). The deletion of pre-pausal final consonants occurred later, in Middle French. This is one of the few cases in which a segment's position within a constituent higher than the Word is mentioned (here the final position in a sentence or an Intonational Phrase, which is also described as [+strong] in Vaissière 1988).
- Before a vowel-initial word, the consonant is in a weak position. It adjoins to the following word, thus finds itself in intervocalic position, and is thus weakened (it becomes voiced) (e.g. S*ix hommes* [sizɔm]).
- Before a consonant-initial word, the consonant is in a weak position (both syllable-final and before a consonant) and it is deleted (e.g. *Six femmes* [sifam]).

Brunot & Bruneau (1937) note that, in the evolution of the language, only the most frequent of these three allomorphs of consonant-final words was generally retained[12]:

> *Of these three forms, si, si-z and sis, which all consonant-final French words did have, only the most frequent one was generally preserved. [Brunot & Bruneau 1937:53, our translation]*

For vowels, the strongest position is the stressed position, but the initial position is also considered to be strong. However, in the literature, the initial position for a vowel generally refers to the vowel *in the initial syllable.* I have not found a distinction between a vowel in the initial

[12] Martinet (1955:295) provides a similar account of the preservation of weak initial consonants in Celtic. While the language contrasted weak and strong consonants in initial position, weak consonants were preserved because they were the most common, in forms generally following a copula or preverbal elements.

syllable and a truly initial vowel (VC) in historical linguistics literature. The preservation of initial vowels has been explained by the presence of secondary stress on the initial syllable of words in Latin. This hypothesis is not accepted by all, and Bourciez & Bourciez ascribe the preservation of initial vowels to their articulatory characteristics during the classical period and even into the French period:

> *The vowel in the initial syllable of words in Latin was pronounced with special clarity, and this is why this vowel has regularly been preserved in French. [Bourciez & Bourciez 1967:42, our translation]*

Vaissière (1996) also distinguishes the evolution of stressed vowels from that of vowels in initial syllables: the former were subject to diphthongization/monophthongization variations, whereas the latter always resisted diphthongization.

<u>B.2. What do these changes consist of?</u>

Language change is characterized by constant pressure in favor of lenition. For example, Brunot & Bruneau (1937) explain consonant lenition by a "gradual disarticulation" of the consonant: articulatory gestures become less and less strong and eventually disappear completely. Lenition mainly targets weaker positions, where acoustic or perceptual cues are the least salient [Steriade p.c., Ohala & Kawasaki 1984, Kohler 1992], or those segments that exhibit the most variability:

> *The consonants which proved to be less resistant with time are the ones which are the more likely to disappear in spontaneous speech. [Vaissière 1996:70]*

To account for this weakening phenomenon, Martinet (1955) invokes functional economy within a given structure, i.e. a tendency to minimize effort that would explain the nature, causes and origin of language change. However, although this weakening tendency dominates the speech signal, we have seen that it does not affect all positions equally: *the initial position appears to resist this weakening tendency.*

According to Venneman (1993), language change follows a principle of syntagmatic emphasis of contrasts within a syllable or a word. Following this principle, language change favors the formation of an optimal word comprising a very strong initial consonant, a strong consonant at the onset of the stressed syllable, and weak consonants everywhere else. Martinet (1955) ascribes this tendency towards the optimal word to processes of "medial obstruent weakening" and "initial sonorant strengthening," which will result in a strong consonant in initial position and weak consonants in medial positions. Hock (1992) notes that initial strengthening does not affect any particular segment or feature but this particular position within the word.

The special status of the initial position in language change can be seen in the particular resistance of segments in initial position to diachronic change. The (strong) initial position is thus distinguished from the following (weak) non-initial positions, in which segments undergo lenition.

III.C. Initial position in synchronic change. Evidence from phonology

The phonological behavior of the sounds of language also exhibits asymmetry between initial and non-initial segments. Here again, initial segments are characterized by their resistance to change. We now turn to synchronic change.

C.1. Less frequent phonological variations in initial position

Segments exhibit different phonological behaviors depending on their position within the Syllable or the Word.

The initial position is much less apt to undergo assimilation, lenition or deletion than non-initial positions are [see among others Bell & Hooper 1978, Ohala & Kawasaki 1984, Goldsmith 1990, Harris 1990, Kohler 1992]. To name but a few examples: the neutralization of voicing contrast observed word-finally is much less frequent word-initially [Kohler 1990, Gow et al. 1996]; non-geminated Spanish /r/ is produced with only one tap in intervocalic medial position, but it retains its multiple taps in initial position [Martinet 1955, Delattre 1965]; certain English consonants can be realized as flaps in medial position, but they are realized as stops word-initially; in several languages, place of articulation contrasts in nasal consonants are lost syllable-finally but retained in initial position [Fujimura & Erickson 1997].

The initial/final position asymmetry also appears in the structure of linguistic units. At the syllable level, languages tend to favor the presence of a strong element in the onset. For example, although this is not the only constraint driving syllable formation, languages tend to follow the sonority scale. A syllable begins with a sequence of increasingly sonorous segments, with the least sonorous (i.e. strongest) element occurring first. The "onset first" principle in Clements and Keyser (1983) also favors heavy onsets over heavy syllable codas.

Finally, the special status of the initial position also appears when one considers the distribution of allophones or the diversity of phone inventories in languages according to their position. Keating et al. (1983) showed that the existence of various allophones for voiceless and voiced consonants in languages depends on their position. Consonant inventories tend to be richer in initial position. In French, for example, this tendency appears in the distributions of the allophones of /R/: in initial position, three allophones are possible, whereas only one (more or less devoiced) appears word-finally [Chafcouloff 1984]. This asymmetry also appears in the consonant inventory of Spanish, which has 19 consonants in initial position but only 7 in final position: one stop out of six, one nasal out of three, and two liquids out of five can appear word-finally [Delattre 1965].

C.2. Are the articulatory characteristics of the initial position phonologized?

Compared to non-initial positions, the initial position thus seems to be characterized by greater resistance to phonological change and by more variety in the nature of the segments it licenses. This positional asymmetry can be due partly to the articulatory and acoustic characteristics of initial segments.

The phonological stability of initial segments has been explained by a "stabler" and "more precise" articulation of segments in this position [Ohala & Kawasaki 1984, Kohler 1990, Browman & Goldstein 1995]. We have seen that the phasing of articulatory gestures of initial segments is less variable and more precise than that of non-initial segments (cf. Section II.B.3).

Conversely, the articulatory gestures of final segments are more variable and thus more apt to overlap [Kohler & Hardcastle 1974, Byrd 1994].

The differences in the behavior of initial and final positions have also been explained by the predominance of anticipatory over perseverative coarticulation. Together with the fact that coarticulation from consonant to vowel is greater than coarticulation from vowel to consonant, the weakening of final consonants is understandable. If a vowel inherits the characteristics of the following final consonant, then the consonant only has a lesser, redundant function, and can thus be weakened or even disappear. This is what happened to final nasal consonants in French: they disappeared after their nasal featured was assimilated by the preceding vowel [Ohala & Kawasaki 1984]. The predominance of anticipatory coarticulation can also account for the changes undergone by final consonants under the influence of a following initial consonant (e.g. "bat cave" → "/bæk/ cave" but never "back tap" → back /kæp/") [Gow et al. 1996]. However, as we have already noted, the anticipatory/perseverative coarticulation asymmetry reflects not only a positional difference within a constituent but mostly a difference in the combination of sounds.

Lastly, the resistance of initial segments to phonological change has also been explained by the fact that these segments are considered to be more salient acoustically and perceptually [Ohala & Kawasaki 1984, Manuel 1991, Kohler 1992, Gow et al. 1996]. Again, this account has been proposed mostly to explain the reduction of final (coda) consonants compared to initial consonants. Initial consonants would be preserved because their informational content is more salient and robust perceptually than that of final consonants. For instance, the perceptual cues carried by CV transitions are richer than those carried by VC transitions (acoustic modulations, amplitude, spectral shape, F0) [Malmberg 1950, Malécot 1960, Fujimura et al. 1978]. Consonant place of articulation information carried by the release burst and consonant voicing information carried by VOT are apparent in onset consonants but not (or to a very small degree) by final consonants (especially in languages such as English in which final consonants are often unreleased):

> *There are thus stronger and more numerous cues as to the identity of place of articulation, particularly for stops, which are differentiated by spectral characteristics of the burst and (in the case of voiceless ones) of the aspiration, over and above the difference in formant transitions. [Kohler 1990:88].*

It follows that contrasts are more apt to be preserved in a highly salient CV syllable, but will be reduced in final position:

> *The syllable- or word-initial position has a higher signaling value for a listener and must therefore be given a more precise articulation by a speaker. Thus the final position has a higher reduction coefficient than the initial one [...] What is not distinctive for a listener anyway may be reduced by a speaker more easily to yield to the principle of economy of effort. [Kohler 1992:209].*

The importance of the salience of information carried by a linguistic unit has also been invoked to explain why words with less informational saliency tend to lose their phonological substance in language change to become function words or affixes [Givon 1975].

III.D. INITIAL POSITION IN SPEECH PRODUCTION AND PERCEPTION. PSYCHOLINGUISTIC CUES

Being more salient and less prone to reduction, the beginning of a word is considered to be an *informational island or "island of recoverability"* (Gow et al. 1996). It may thus have a special status in the perceptual processing of speech, notably in segmentation and lexical access.

Because it is phonologically more stable, the beginning of a word is the most robust (invariant) source of information contained in the underlying phonological form that is supposed to be stored in the mental lexicon [Gow et al. 1996]. Less reduced, the beginning of a Word can facilitate lexical access by providing the hearer with surface forms that are very similar to the underlying forms.

Similarly, position-dependent allophonic variations can be considered as cues facilitating lexical segmentation: when the hearer encounters an allophone in the speech stream which only occurs in initial position, s/he can deduce the presence of a word boundary. Allophonic variations are thus no longer considered as a source of noise for speech processing but as distributional information [Church 1987]. Studies of lexical segmentation of ambiguous strings show that hearers can use the presence of initial allophones to segment the string correctly [e.g. Nakatani & Dukes 1977]. However, hearers perform poorly if word boundaries are not marked in the signal by particular allophones or strong cues (e.g. aspiration or glottalization). This suggests that the articulatory variations observed Word- or Syllable-initially cannot necessarily all be perceived or used in segmenting by the hearer. Those articulatory variations that are greater at the beginning of higher-level constituents may be more robust cues to segmentation. These cues would then combine with other cues to prosodic boundaries (variations in F0, duration, etc.). Indeed, hearers' performance in segmentation tasks shows that they achieve better results when the sequences to be segmented are based on a boundary greater than a Word boundary [e.g. O'Connor & Tooley 1964, Nakatani & Dukes 1977, Rietveld 1980, Quené 1992].

The special status of the initial position is also apparent when one considers the consequences on speech processing of distortions in this position. Several studies have shown that degraded information Word-initially has a much stronger effect on lexical access than distortion occurring later in the Word [e.g. Bagley 1900, Cole & Jakimik 1980].

These results have been used by Marslen-Wilson (1980) in his theory on lexical access. He claims that a word is divided into two parts, one carrying more information than the other. The informative part (the beginning of the word) triggers the activation of all the lexical entries sharing the same initial string of phonemes (the initial cohort). As they are heard, the following phonemes will make a selection between all the entries in the initial cohort, until the word reaches its disambiguation point and is recognized. This model, said to be overly restrictive, is controversial [e.g. Hawkins & Cutler 1988]. However, the informational importance of the beginning of a word in lexical access is commonly accepted (it may also be explained simply by the fact that the beginning of a word is what reaches the ear first).

The importance of the beginning of a word is also apparent in word retrieval tasks, in which the hearer has to guess a word after hearing only a small part of that word. These studies show that subjects retrieve words faster and more easily when hearing an initial sequence than a final sequence, and perform worst when hearing a medial sequence [Horowitz et al 1968, Nooteboom 1981]. Likewise, in tip-of-the-tongue phenomena, speakers generally have a better idea of the beginning of the word they're looking for than of its end [Browman 1978, Brown & McNeill 1966].

Data on the distribution and characteristics of production errors according to position within a word also highlight the special status of the initial position. Speech error research has shown the importance of the structuring of speech units during encoding. Inverted segments are always pairs occurring in the same position in the original words. For example, consonants /f/ and /p/, word-initial in the sequence "**p**arade **f**ad **f**oot **p**arole" are more likely to be exchanged ("farade pad foot farole") than if they are not in the same position in the words of the sequence (e.g. "re**p**eat **f**ad **f**oot re**p**air") [Shattuck-Hufnagel 1985]. Additionally, Shattuck-Hufnagel (1986) shows that these errors are most frequent in initial position, the next most frequent position being under stress (especially in the case of vowels). Then,

> *The implication of these findings is that both structure and lexical stress are part of the representation that is in force when consonant errors occur, and that similarity in word-onset position is the more powerful of the two factors in determining which pairs of consonants will interact in errors. [Shattuck-Hufnagel 1986:134]*

However, it is not known why these segments (in initial and stressed positions) are more likely to be exchanged than others. Klich et al. (1979) claims that the greater occurrence of exchanges in initial position is due to greater complexity in encoding. Van Lieshout et al. (1995) propose a similar account for the greater occurrence of stuttering in initial position:

> *We would predict that whenever in such situations more articulatory effort is required, as in sentence initial position and with longer words or sentences, disfluency will increase. [van Lieshout et al. 1995:371]*

III.E. THE SPECIAL STATUS OF THE INITIAL POSITION: CONCLUSION

This section has shown that the initial position has a special status on the basis of the linguistic behavior of initial segments. The initial position is considered as a strong position in which segments are more resistant to reduction. It is interesting to note that certain characteristics of the initial position are similar to those of the position under stress, whose special status is uncontroversial.

The special status of the initial position (and of the position under stress) is thus apparent at several levels in the representation of speech:

- Psycholinguistic data show that the assignment of various positions in the skeleton of a word occurs at the *level of encoding* (e.g. Levelt 1992). The initial position appears as marked.
- The articulatory data in section II also show that these positions exhibit particular characteristics at the *production level*.
- Some of these phonetic/articulatory characteristics have been phonologized at the *linguistic level*. Segments in initial position are more resistant to a language's reducing tendencies, in both diachronic change and synchrony.
- Lastly, at the *level of perceptual processing*, the initial position is distinguished by its informational content for segmentation and lexical access.

Figure 1.10: Comparison of articulatory variations observed under stress (gray background) and in initial position in a prosodic constituent (white background). The variations are illustrated for each articulator with one segment representing its category (e.g. /t/ for oral obstruents, /n/ for nasals, /u/ for rounded vowels, /p/ for labials).

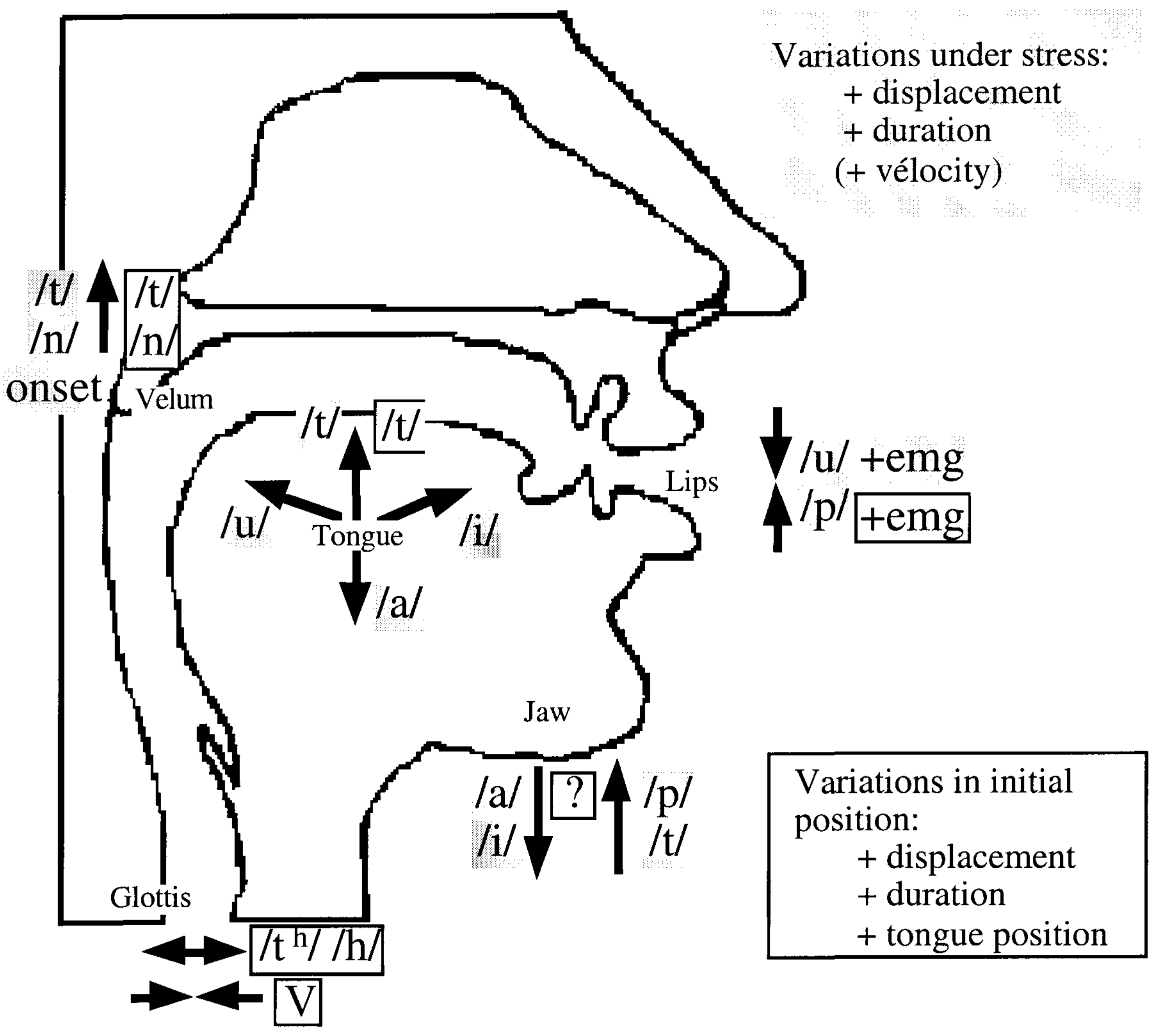
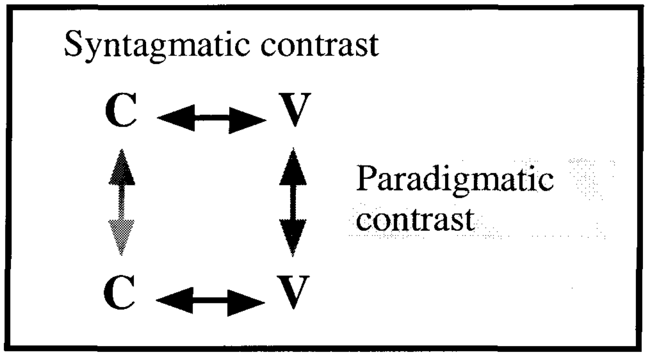


Figure 1.11: In black, the sonority expansion hypothesis highlights the syntagmatic contrast between a consonant (C) and a vowel (V). In gray, the contrast enhancement by local hyperarticulation hypothesis highlights the paradigmatic contrast between segments.



IV. HYPOTHESES ABOUT THE MECHANISM GOVERNING ARTICULATORY VARIATIONS IN INITIAL POSITION

IV.A. A COMPARISON BETWEEN ARTICULATORY VARIATIONS UNDER STRESS AND IN INITIAL POSITION

Figure 1.10 summarizes the variations observed under stress (gray arrows) and those affecting segments in initial position in a constituent (black arrows). The comparison of these two variation factors is not always easy because the segments and articulators that have been studied under stress and in initial position are not always the same. However, the following observations stand out:

- Both under stress and in initial position, the amplitude and duration of articulatory gestures tend to increase.
- Labial articulation: stress and initial position have a similar effect, in that they increase muscular activity in the lips. The effect of the initial position on the movement of the lips (rounding, protrusion) has not been studied enough to draw any conclusions.
- Nasal articulation: stress and initial position appear to interact (see Vaissière 1988 for English). Word-initially, the velum is always higher than in coda position, for both oral and nasal consonants. Stress seems to reinforce this initial/final distinction: stress contributes to raising the velum in Word-initial position and to lowering it in coda position.
- Lingual articulation: stress and initial position reinforce the consonantal aspect of consonants (constriction) by raising the tongue. For vowels, stress most frequently reinforces the inherent quality of vowels by making them more peripheral (see section II.A.3). In initial position, only a few acoustic data suggest that, if vowels are affected (in Finnish but not in Czech, Lehiste 1964), variation also contributes to making the vowels more peripheral. There may be differences in jaw raising in initial position as has been shown under stress, but this has not been studied enough to draw any conclusions.
- Glottis: initial position and stress both favor the glottalization of vowels.

The variations observed in initial position thus generally resemble those observed under stress. Since stress and boundaries define important positions in prosodic organization, one might expect that the variations in these two kinds of position are similar in nature and governed by the same physiological mechanism.

IV.B. HYPOTHESES PROPOSED TO ACCOUNT FOR ARTICULATORY VARIATIONS

B.1. The sonority expansion hypothesis and the distinctive feature enhancement hypothesis

Two opposing hypotheses have been proposed to describe the underlying mechanism governing variations under stress. These hypotheses are based on old notions but they have recently been reformulated and discussed in the articulatory prosody literature. They can be summarized as follows.

(1) The sonority expansion hypothesis

Edwards & Beckman (1988) and Beckman et al. (1992) propose that stress contributes to increasing the sonority distinction within the syllable by reinforcing the segments' intrinsic sonority. In a stressed position, a consonant (C) would be less open and so less sonorous, whereas a vowel (V) would be more open and remain so for more time, and as such more sonorous. This hypothesis is comparable to the one proposed by Straka (1963) to account for the articulatory variation observed in accented and initial positions and in reinforced speech. Straka does not account for the increase in consonant-vowel contrast with sonority but with aperture: consonants exhibit a stricter closure and vowels are more open in reinforced speech.

(2) The distinctive feature enhancement or local hyperarticulation hypothesis

De Jong (1995) rejects the idea that stress implies a sonority expansion mechanism. Stress affects not only the jaw but every articulator. Furthermore, variations under stress can conflict with an increase in a segment's sonority: for example, closing the lips for /U/ or raising the tongue for non-low vowels both diminish these segments' sonority. De Jong refers to Lindblom's (1990) notion of hyperarticulation, explaining that segments are hyperarticulated under stress. This notion no longer applies to a difference in communication situation (Lindblom 1990) but applies locally to stressed syllables in the speech chain. Consequently, articulatory gestures under stress are more extreme in the direction of their assumed target. Thus stress would have a paradigmatic effect, reinforcing the distinctive features of segments.

In the light of the articulatory variations observed under stress (section II.A), it appears that both mechanisms coexist:

- The sonority expansion hypothesis is confirmed by the behavior of the jaw, which is higher in consonants and lower in vowels [Farnetani & Vayra, de Jong, Giot, Beckman, Macchi, etc.]
- The distinctive feature enhancement by local hyperarticulation hypothesis is confirmed by observations of tongue and lip gestures [Kent & Netsell, de Jong, Macchi, Farnetani & Vayra, Giot, Vaissière, etc.]. However, the distinctive feature enhancement hypothesis is not helped by the fact that various authors do not observe the same variations in vowels (opening for Houde, stability for Macchi, peripherality for others) and that the velum's variations in initial

position induce a decrease in the realization of the nasal feature in stressed nasal consonants [Vaissière 1988].

The articulatory realization of stress could correspond to a combination of these two mechanisms aiming at increasing segment distinctiveness, both syntagmatically and paradigmatically. The effect of either one of these two functions would be articulator- or speaker-specific (as speakers sometimes exhibit different strategies in the realization of stress) [Beckman et al. 1992, 1994, Krakow 1993, de Jong 1995, Harrington et al. 1996].

B.2. Can these hypotheses account for the variations observed in initial position?

Under stress, the two hypotheses (sonority and distinctive features) can be differentiated by the predictions they make with respect to the lingual articulation of vowels and by their effect on the jaw. However, as we have seen, initial-position articulatory variations have hardly been studied for initial vowels, and with respect to the jaw, the two available studies give contradictory results.

Let us still attempt to test these two hypotheses based on the articulatory variations known to occur in initial position (section II.B):

- The CV-contrast-increase hypothesis predicts the variations observed in initial consonants: obstruents exhibit more lingual closure, nasals (with a higher velum) and /h/ (with weaker energy) are more consonantal [Manuel 1991, Pierrehumbert & Talkin 1992, Fujimura & Erickson 1997]. The reinforcement of syntagmatic contrast within a CV syllable is also confirmed by the results in Farnetani & Vayra (1996), who observe more consonantal closure and more vocalic opening in initial position. However, this hypothesis does not predict the glottalization of initial vowels, which does not reinforce the syntagmatic contrast between the glottalized vowel and the adjacent consonants. Besides, if one generalizes Lehiste's (1964) acoustic observations, it seems that close vowels are closer in initial position, which does not contribute to enhancing a sonority contrast.
- The distinctive feature enhancement hypothesis also predicts an increase in constriction in stops in initial position: obstruents show more constriction in initial position. However, this hypothesis is strongly refuted by the behavior of the velum: raising the velum in initial nasals goes against an increase in the realization of the nasality feature. The glottalization of initial vowels cannot be explained by a reinforcement of the vocalic feature either.

In conclusion, neither of these hypotheses is confirmed by the articulatory variations observed in initial position. As in the case of variations under stress, the behavior of certain articulators or segments confirms either one or the other hypothesis. I will thus present another hypothesis, which explains articulatory variations in initial position as a consequence of "articulatory strengthening" in this position.

IV.C. THE INITIAL POSITION ARTICULATORY STRENGTHENING HYPOTHESIS

C.1. History and definitions of "strength"

C1.a. "lenition" or "weakening" vs. "fortition" or "strengthening"

Articulatory changes in segments in medial or final position are generally described as "lenition" or "weakening." Conversely, variations in initial position are often called "fortition" or "strengthening." These terms, which directly or indirectly refer to the notion of "strength," have been commonly used in historical linguistics and in phonology to explain certain diachronic or synchronic segmental changes [e.g. Lass 1984, Straka 1964, Hock 1992].

Lenition can be defined as follows:

> *Lenition (also weakening): Any phonological process in which a segment becomes either less strongly occluded or more sonorous, such as [k] → [x], [x] → [h] or [k] → [g]. Often the term is extended to various other processes, such as loss of aspiration, shortening of long segments and monophthongization of diphthongs, which represent "weakening" in some intuitive sense. [Trask 1996].*

> *Lenition is the term used to describe a mutation in consonants, which normally originated in a decrease in the energy used in their articulation. [Thurneysen 1946, cited in Martinet 1955: 257].*

Duez considers the weakening observed in spontaneous speech as an

> *obscuration process in which a consonant is modified in the direction of lesser constriction or weaker articulation, such as a stop becoming an affricate or fricative, or a fricative becoming a sonorant. [Duez 1995: 409].*

Fortition (or strengthening) is considered to be the reverse mechanism. Trask defines it as follows:

> *Fortition (also strengthening): Any phonological process in which some segment becomes "stronger" (more consonant-like). An example is the development of the glide [j] into some kind of fricative, affricate or plosive in most varieties of Basque. [Trask 1996].*

Other authors do not really define fortition. It is only suggested in their descriptions under such terms as "greater articulatory force," "greater energy," "greater tension," "sharper and more precise articulations," "more extreme articulations," "more consonant-like articulations" (Ohala & Kawasaki 1984, Vaissière 1986a, 1988, Fujimura 1990, Keating 1995, Dilley et al. 1996, Fujimura & Erickson 1997, Fougeron & Keating 1997, etc.).

C1.b. What is "articulatory force" or "articulatory effort"?

The notion of "articulatory force" or "articulatory effort" was largely used in the literature in the early 20th century (no doubt on the basis of the first available articulatory data) but rarely defined in physiological terms. However, as Slis notes, descriptions of quite precise articulatory variations are made possible by this idea:

articulatory effort... is commonly used in the literature, [but] it is not particularly well defined, and seems to be largely based on intuition. Nevertheless, it may be shown that in a number of linguistic oppositions, allegedly differing in articulatory effort, there are consistent behavioural correlates... [Slis 1971: 398].

Delattre (1940a) notes that Rousselot referred to "la force de l'articulation." Then with his students (Genévrier, for instance) the term became "force d'articulation." But while this term originated with Rousselot's school, neither Rousselot himself nor his followers defined it, and they all use it rather vaguely. In his paper, Delattre (1940a) suggests the following definition:

We believe that consonantal articulatory force refers to the amount of energy necessary to the realization of all the muscular effort involved in the production of a consonant. [Delattre 1940a: 111, our translation].

Malécot, who extensively studied articulatory force, proposes a very similar definition:

Force of articulation is a physiological attribute of consonants; the degree of force of articulation of a given consonant is defined as the relative amount of muscular energy required to utter it. [Malécot 1955: 35].

Straka (1963) defines it more precisely by comparing it to other aspects of articulatory intensity (p.102ff). He takes as synonymous "articulatory energy," "articulatory force" and "articulatory effort," which he defines as follows:

Articulatory energy is simply the contraction force of the muscles involved in a given articulation. [Straka 1963: 91, our translation].

This articulatory energy (or force) is that of the *articulatory gesture* (or articulatory movement). For example, it is the force with which the tongue and maxilla muscles contract in order to set up these organs for a particular articulation.

"Articulatory force" as force of the "articulatory gesture" is distinguished from "overall muscular effort," which is the sum total of the muscular effort required by a given articulation. This "overall muscular effort" includes "articulatory force" as well as "laryngeal or phonatory force" and "expiratory force[13]." This distinction between the different components of the overall effort appears to Straka to be necessary, because the variations in muscle contraction do not necessarily affect all of the "overall muscular effort" but only some of its components (its "forces"). In addition, variations in different forces can compensate for each other (e.g. a weakening of supraglottal articulations can be compensated with a strengthening of expiratory force). Thus phonetic variations are not due to changes in the overall effort but to variations in specific forces.

Straka also distinguishes between "articulatory force" and "articulatory tension" (p. 108). Articulatory force corresponds to the first phase of the articulation: *the muscle contraction phase*

[13] Straka also distinguishes between "articulatory force," which applies to lip, jaw and tongue gestures, and "velo-pharyngeal force," which is involved in the gestures of the velum and of the pharyngeal wall. However, when discussing variations in nasal articulation in reinforced speech, he includes variations in the contraction of the velum muscles in "articulatory force," together with the contraction of the tongue or lip muscles.

necessary to set up the appropriate organs to perform an articulation (Grammont's *catastasis*). This is a dynamic notion related to movement. On the other hand, articulatory tension (not to be confused with muscular tension) is a static notion. It serves to "keep the muscles contracted" in order to "keep the organs in position for a very short time." During the course of a given articulation, the task of maintaining muscle contraction (articulatory tension) at the level resulting from the initial gesture (articulatory force) can also vary. But this variation is different from that which affects articulatory force during the production phase of the gesture.

In summary, Straka considers variations in articulatory force to be variations in muscle contraction/relaxation. The affected muscles are those of the organs involved in a particular articulation and not those of the overall production mechanism. These muscle contractions can be more or less strong. As they set the relevant organs in position for a given articulation, the stronger they are, the more extreme the organ's position. If the relevant muscle is a raising muscle, the organ will be higher, and if it is a lowering muscle, the organ will be lower[14].

C1.c. Contrasts described with reference to "articulatory force"

α. Classifying sounds according to their "articulatory force"

The notion of "articulatory force" has long been used to characterize various classes of sounds (see e.g. Rousselot 1901, Roudet 1910, Martinet 1955, etc.). Sounds are ordered along one or more "force" scales which can be summarized as follows:

stops > affricates > continuants > vowels
voiceless > voiced > nasals and liquids
palatal > non-palatal

(From Straka 1964. ">" stands for "stronger than")

β. The fortis/lenis contrast

The terms "fortis" and "lenis" also refer to articulatory force. They are generally used to describe a phonological contrast between two classes of consonants which cannot be characterized as a voicing contrast, or to describe secondary phonetic characteristics associated with a voicing contrast between consonants. For French consonants, Malmberg (1943) gives the following contrast:

fortis: p, t, k, f, s, ʃ
lenis: b, d, g, v, z, ʒ, m, ɲ, n, l, r, j, ɥ, w

The existence of a fortis/lenis feature as a primary contrastive feature has long been highly controversial (see Malécot 1970, Jeagger 1983, Kohler 1985, etc.). First, because the articulatory

[14] This distinction between raising and lowering muscles, and their often subjective definition, is what makes Straka's theory most debatable. We address this later.

or physiological basis of this alleged contrast is very vague. Second, because the phonetic characteristics attributed to this contrast are highly variable and very often circular (see e.g. Jeagger 1983). This study does not attempt to justify the existence of a fortis/lenis contrast in the phonological representation of a system. I only mention this notion because it has been associated with the idea of "articulatory force." "Fortis" consonants are supposedly articulated with greater articulatory force, as Pike notes:

> *Fortis articulation entails strong, tense movements [...] relative to a norm assumed for all sounds...weak articulation is lenis. [Pike 1943: 128].*

The differences in articulatory force used in the literature to distinguish these segments often refer to impressionistic accounts of the muscular force required for the articulation of a segment (the "muscular sense" in Delattre 1940a). They can also correspond to an auditory impression of force (e.g. Malécot 1955). They have also been illustrated by articulatory differences. However, the great number and diversity of articulatory (or acoustic) variations presented in the literature as correlates of articulatory force have led to heated debates about the usefulness of this notion (see e.g. the Lebrun-Malécot exchange cited in Debrock 1977[15]). Quite frequently, when reading the literature, one does not know what exactly these articulatory variations consist of, whether they really reflect articulatory force, or whether it is even possible to measure articulatory force. Nevertheless, as we shall see, the articulatory characteristics that have been associated with articulatory strengthening exhibit striking similarities with the articulatory variations observed in initial position.

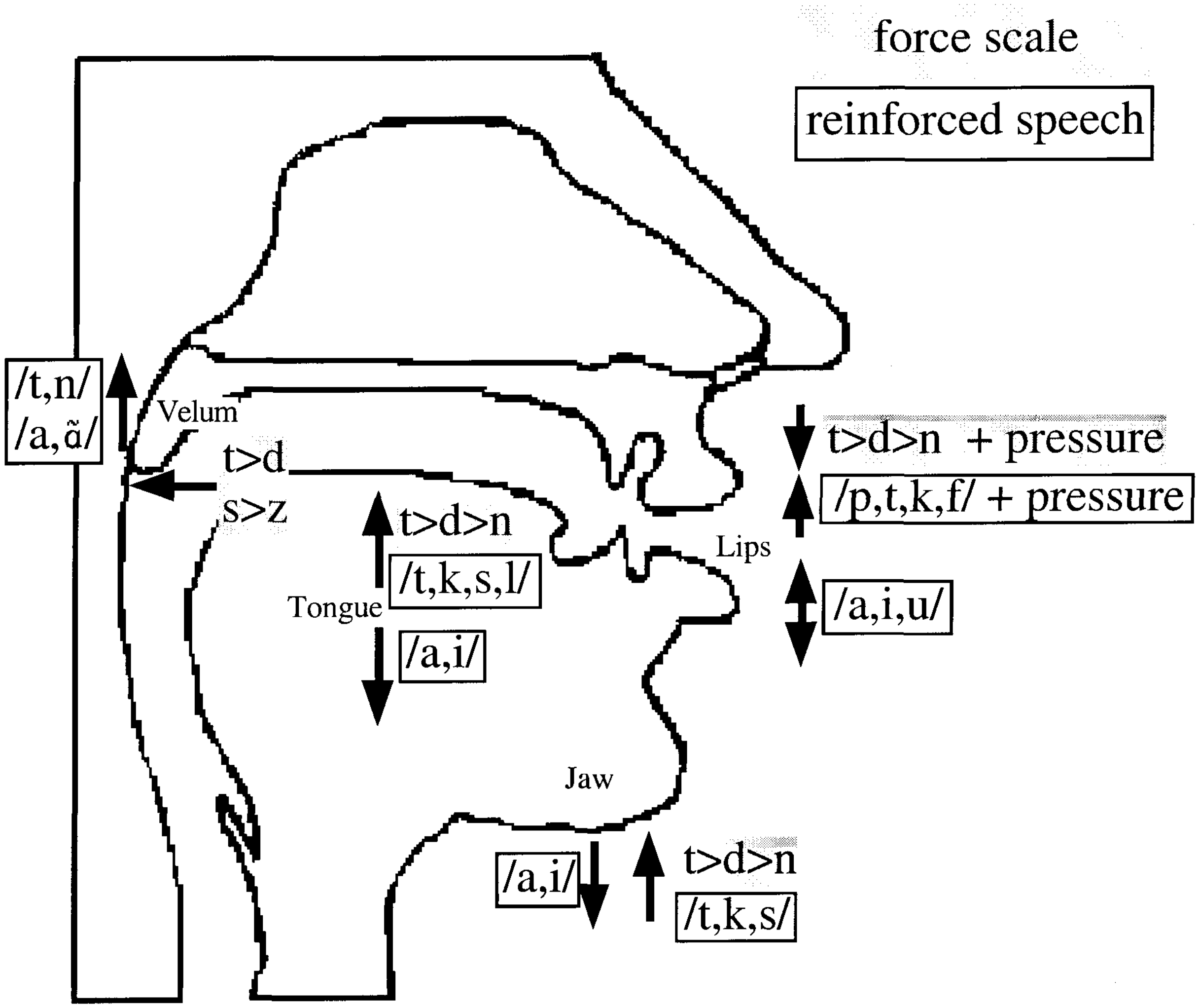
C.2. Articulatory characteristics of "strong", "fortis" segments and of reinforced speech

α. "Strong" segments

So-called "strong" segments are fortis consonants and the segments at the top of the force scale shown above. Their articulatory characteristics are shown in figure 1.12 (gray background). ">" stands for "greater than" for each articulatory variable.

[15] For example, fortis consonants have been described as more resistant to airflow (Trubetzkoy 1949), as involving greater intra-oral pressure (Rousselot 1897, Stetson 1951, Malécot 1955) or greater pulmonic force, as longer (Jakobson, Fant & Halle 1963, Straka 1963), as exhibiting more abrupt transitions with the neighboring vowels (Debrock 1977, Jaeger 1983), etc.

Figure 1.12: Articulatory characteristics of sounds along the force scale (voiceless>voiced>nasals) (gray background) and in reinforced speech (white background).



They are characterized by:

- Higher tongue and jaw, narrower labial opening (voiceless>voiced>nasals, stops and fricatives) [Haden 1938, Chlumsky 1938, Simon 1967, Giot 1977]. Consequently, the antagonism between consonantal and vocalic articulatory gestures (CV) gestures with respect to the jaw, the lips and the tongue position is greater (voiceless>voiced>nasals) [Giot 1977].
- Wider linguo-palatal contact (voiceless>voiced>nasals) [Rousselot 1901, Simon 1967, Marchal 1979, 1984].
- Greater stiffness of the tongue body before release of closure, thus making the tongue less likely to be affected by coarticulation with the following vowel (voiceless>voiced>nasals) [Giot 1977].
- Firmer approximation of the velum against the pharyngeal wall for non-nasal consonants (voiceless>voiced, stops and fricatives) [Simon 1967].
- Greater (mechanical) lip pressure for labial consonants and greater linguopalatal pressure (voiceless>voiced>nasals) [Rousselot 1901, Malécot 1966].

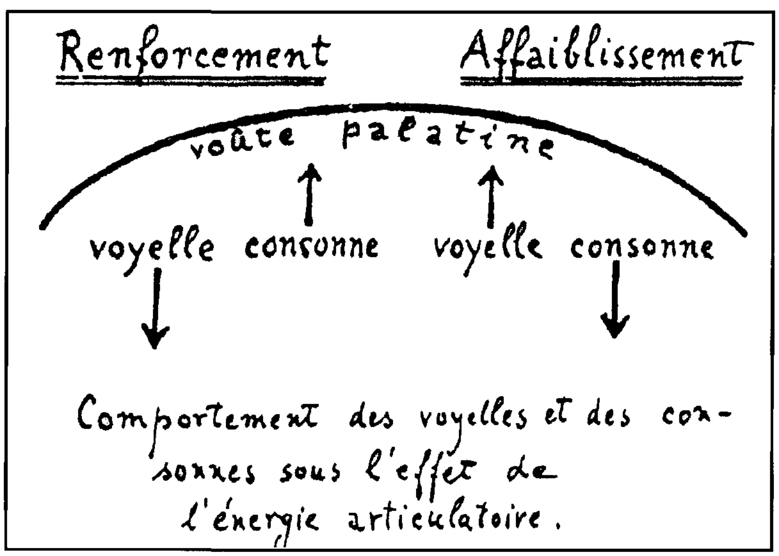
β. "Reinforced" speech

The effects of an increase in articulatory force on several types of segments were studied by Straka (1963) in a style of speech he terms "reinforced speech[16]" ("la parole renforcée"). He does not clearly define what this speech style consists of or what instructions the speakers were given. He describes it as "more energetic pronunciation," "reinforced pronunciation" or "more energetic flow." These energetic pronunciations are compared (1) to "normal" pronunciations," "produced with usual force" and (2) to "weakened," "produced with less, diminished energy." The reinforced (more energetic) pronunciations are probably produced with greater vocal intensity or hyperarticulated. Even though one does not really know whether the speech style he studies is a good indicator of articulatory force, the results he shows are interesting for our study because the variations observed are very similar to those observed as a function of prosodic position.

The author uses a detailed examination of the effects of variations in articulatory force to account for some historical changes. This study also allows him to distinguish consonants and vowels physiologically as two distinct segment classes in a language. Variations in articulatory force have completely opposite effects in these two classes of sounds, which are shown in figure 1.13:

> *As a result of an increase in articulatory force, a consonant's closure and a vowel's opening increase. Conversely, as a result of articulatory weakening, a consonant's closure and a vowel's opening decrease. [Straka 1963: 77, our translation].*

Figure 1.13: Behavior of vowels and consonants under the influence of strengthened or weakened articulatory force (energy). From Straka (1963: 79). With strengthened energy, consonant closure and vowel opening increase.



[16] The articulatory data collected by Straka include palatography, X-rays, cineradiography, oscillograms, kimograms, from his or other authors' work. He presents data from French, Spanish, Catalan, Italian, English, German, Polish, Serbo-croatian, Albanian, Alsatian, Estonian, etc.

Consequently, an increase in consonant closure and vowel opening result in a greater oral aperture contrast in a CV syllable, as shown by Straka's illustration in figure 1.14.

Figure 1.14: Opening ratio of consonants and vowels as a function of a syllable's articulatory energy. From Straka (1963: 83). Greater articulatory force enhances the aperture contrast in a CV syllable.

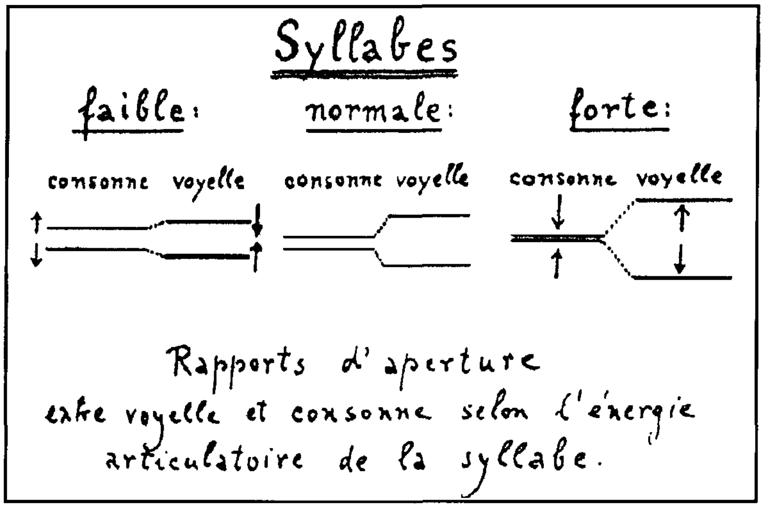


Table I.1: Straka's observations of reinforced speech, by segment type and articulator. ↑ indicates a raising of the articulator (and a decrease in labial opening for the lips)and ↓ indicates a lowering of the articulator (and an increase in labial opening for the lips):

	Jaw	Tongue	Lips	Velum
Stops				
oral	↑	↑	↑	↑
nasal	↑	↑	↑	↑
lateral	↑	↑	↑	
Fricatives	↑	↑	↑	
Vowels				
open	↓	↓	↓	↑
close	↓	↓	↓	↑
nasal				↑

The articulatory characteristics of segments in reinforced speech are summarized in table I.1 and shown in figure 1.12 (white background). Organized by segment and articulator, these observations are:

- Lingual stop consonants are characterized by a decrease in oral aperture caused by a sharper jaw angle and a higher tongue position. This higher tongue position results in a greater linguo-palatal contact area. The contact area widens towards the hard palate (towards the back for anterior constrictions and the front for posterior constrictions) along the median axis and the sides of the palate. These changes appear on figures 1.15 and 1.16, which show palatographic data for /t/ and /k/ and X-ray tracings of /t/, pronounced with various degrees of strength. Lip aperture for consonants is also reduced in reinforced speech. For labial stops, Straka observes in frontal-view photographs that the lip closure is stronger when consonants are pronounced with energy. For nasal and oral stops, the velum is higher in reinforced pronunciation. This appears in figure 1.15 for /t/. Unfortunately, Straka does not provide any illustration for nasals.

- Continuants undergo similar changes in reinforced speech. The oral opening is reduced by a raising of the jaw and tongue, and is performed articulatorily by a widening of linguopalatal contact on the edges of the palate towards the median axis. Under extreme strengthening, lateral contact areas can meet and form a stop closure, and fricatives can become stops. For labial and labiodental fricatives, the labial opening is narrower in reinforced speech.
- The lateral /l/ in reinforced speech is produced with a widening of the median contact of the tongue tip against the palate. The contact area then tends to spread towards the back.
- In vowels, articulatory strengthening causes an increase in oral opening via a lowering of the jaw and tongue. This is observed for all vowels, whether open or close. Figure 1.17 shows X-ray and palatographic tracings of various vowels pronounced with great and normal force. The palatographic data suggest that a lowering of the tongue causes a decrease in linguopalatal contact. For rounded vowels, Straka notes that there is no "intensifying of labiality" (rounding? protrusion?) but an increase of the labial opening, as in other vowels. A greater lip aperture can be the consequence of the lowering of the jaw. The velum is higher in both nasal and oral vowels. This appears on the X-ray tracings of /i/ and /o/ in figure 1.17. Again, Straka does not provide any illustration of the position of the velum for nasal vowels.

Figure 1.15: Articulation of /t/ pronounced with great and normal force (figure 10 in Straka 1963). The vocalic context is not the same. Note the greater lingual closure, raised jaw, tongue and velum in reinforced pronunciation.

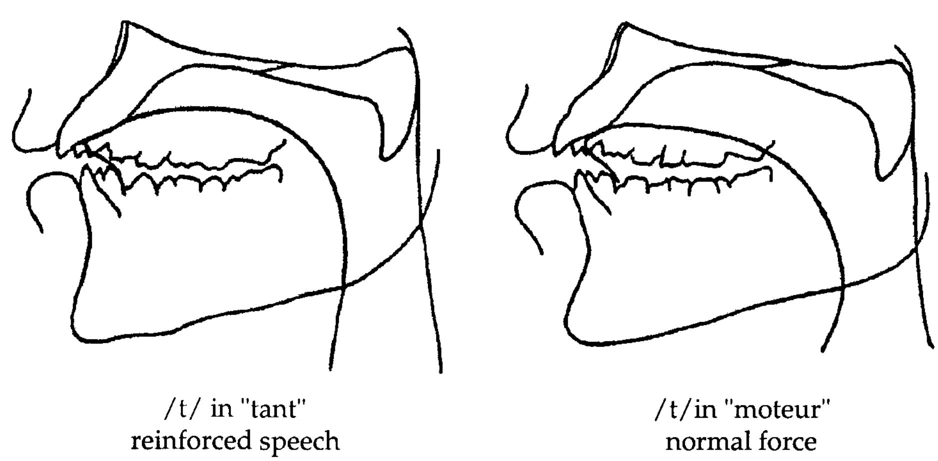


Figure 1.16: Palatographic tracings showing linguo-palatal contact area (hatched) for /t/ and /k/ pronounced with different degrees of force: weak, normal and strong (figure 6 in Straka 1963). Note the widening of the contact area towards the back for /t/ and the front for /k/.

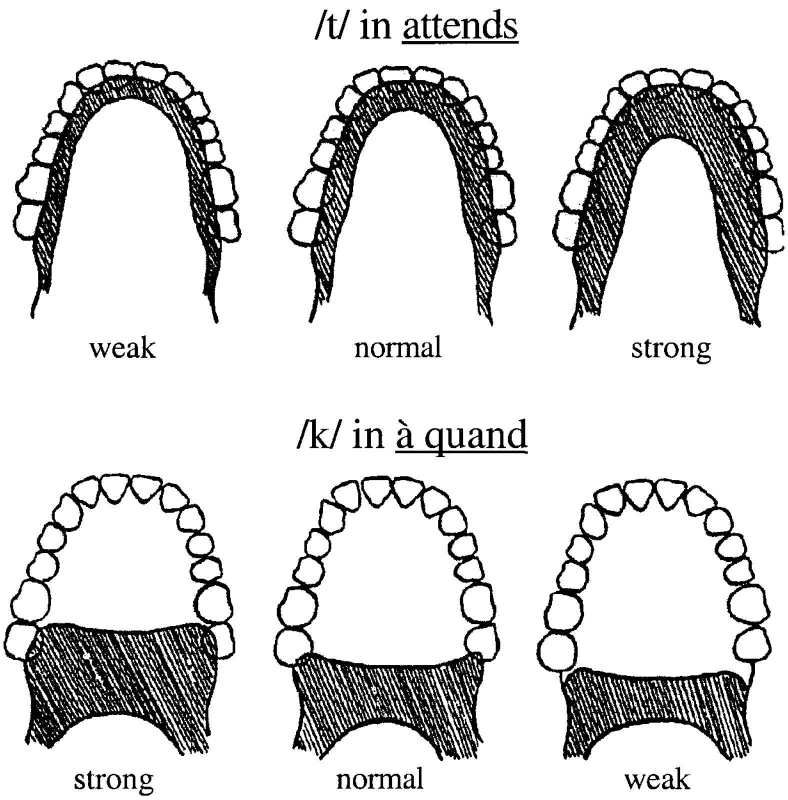
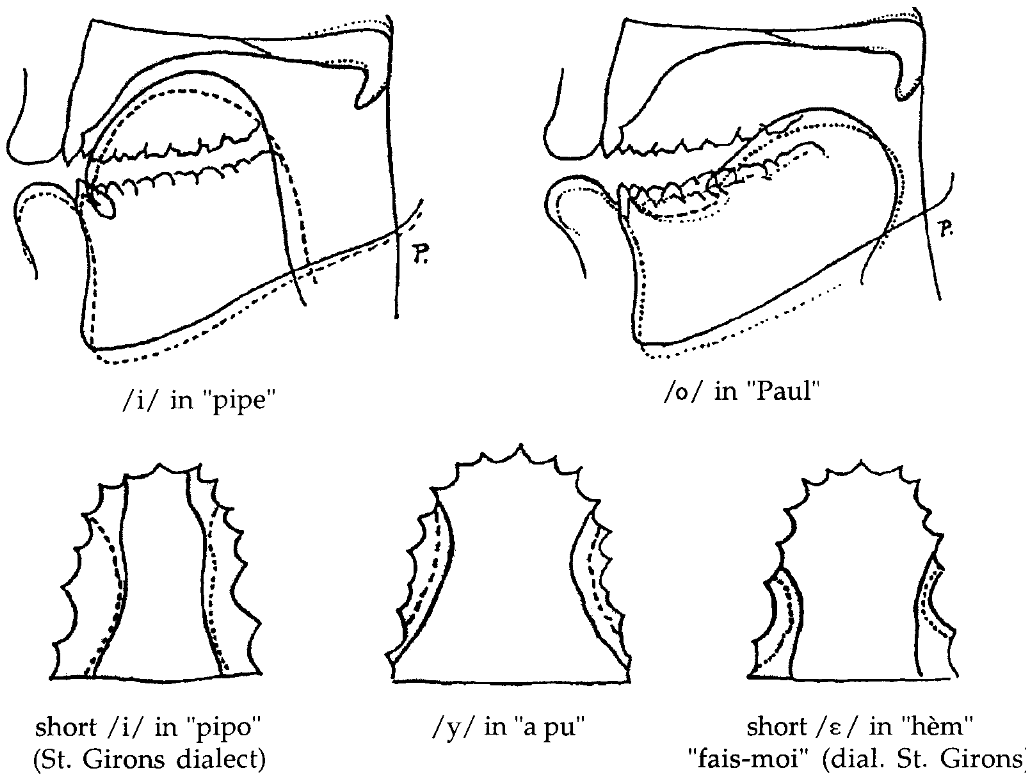


Figure 1.17: Articulation of several vowels pronounced with strong articulatory force (dotted line) and normal force (solid line), based on figures 7a and 11 in Straka (1963). Note the greater lingual closure area, lowering of the tongue and jaw, lip aperture and raising of the velum in the strong condition.



Solid lind: normal articulatory force
Dotted line: great articulatory force

In conclusion, comparing the articulatory characteristics of the segments described as "strong" or "strengthened" (see figure 1.12) to those of initial segments within a prosodic constituent (see figure 1.10), the similarities are striking. Initial segments share several of the characteristics associated with "strong" or "strengthened" segments, notably with respect to variations in lingual and nasal articulation.

C.3. How greater articulatory force can account for variations in initial position

The notion of articulatory strengthening has thus been applied to the articulation of particular segments (strong-fortis) and to a specific speech style (reinforced speech). This notion has also been used to identify *particular positions*: position under stress (Straka 1963, Vaissière 1988) and the initial position in a Word (Vaissière 1986a, 1988). Vaissière does not use the term "strengthening" but the same idea is found in her work. She proposes that (muscular) tension increases in the whole production mechanism at the beginning of a Word and under stress[17].

We have observed that initial consonants, like "strong" or "strengthened" consonants, exhibit greater tongue height and pressure against the palate. The velum is also higher in both oral and nasal consonants. If one accepts Straka's definition, articulatory strengthening consists in *increasing the contraction force of the muscles involved in a given articulation* (Straka 1963: 91). We will see how this increase in muscle contraction can predict these variations:

- The raising of the tongue in consonants is due to the contraction of the tongue's raising muscles (genioglossus, palatoglossus, styloglossus) [Straka 1963, Lieberman & Blumstein 1988, UCLA 1990]. Articulatory strengthening in initial position would cause these muscles to contract more, thus raising the tongue. For stops, a higher tongue position would result in greater pressure against the palate, and so a widening of the contact area.
- A raising of the velum in initial position can also be explained by greater muscle contraction. Variations in velum height are strongly correlated with the activity of the levator veli palatini muscle [Bell-Berti & Hirose 1975, see Bell-Berti 1993 for a review]. During the closing phase of the velopharyngeal port in the production of non-nasal segments, the contraction of the levator palatini pulls the velum upwards and towards the back. This action may be accompanied by the contraction of the superior pharyngeal constrictor, but not all authors agree on this point. The mechanism governing the opening of the velopharyngeal port in nasal segments is more controversial. For some, this opening is an active mechanism caused by the action of several lowering muscles [Lubker et al. 1970], but for most the lowering of the velum is a passive mechanism [see Bell-Berti 1993]. For these authors, the velum is lowered by the relaxing of the muscles causing its raising, i.e. the levator palatini. Together with this relaxing of the levator palatini, an increase in the activity of the palatoglossus has also been observed by some researchers [Fritzell 1969, Lubker et al. 1970] but not by others [Bell-Berti 1973, 1976, Bell-Berti & Hirose 1973]. If one accepts the idea that raising and lowering the velum is mainly achieved by contracting and relaxing the levator palatini, respectively, greater articulatory force predicts the observed variations:
    - In non-nasal segments, which always exhibit a contraction of the levator palatini [Lieberman & Blumstein 1988], a greater contraction would cause the velum to be raised higher
    - In nasal segments, greater articulatory force would decrease the relaxing of the levator palatini, and would result in a lesser lowering of the velum[18]. This position is shared by Fujimura (1990) who notes that, even though the physiological mechanism of velum control in speech is not perfectly understood (p. 201),

[17] The increase in tension initially would be associated to the idea of "beginning." It appears at the beginning of a Word, but also at the beginning of a speech act, and so it could partly account for the "speech ready gesture" (Vaissière 1986b). This initial tension would be followed by a gradual articulatory relaxing along the utterance until a state of maximal relaxation, which would be associated with the idea of "end" (see also Kohler 1992).

[18] Straka (1963): "The lowering of the velum in nasal articulations is the result of the relaxation of the peristaphylines, or even the palatostaphylines (uvular azygos). Their contraction raises the velum and causes a rhinopharyngeal closure in oral articulations." (p. 95, our translation). 'Palatostaphylines' (or 'uvular azygos') for Straka (and Rousselot) appear to correspond to the levator veli palatini; the internal and external 'peristaphylines' appear to correspond to the tensor palatini ('staphyline' means uvular). However, even if the muscles considered as essential differ, the mechanism remains the same: the raising of the velum is achieved by contracting some muscle and its lowering by relaxing some muscle.

> *muscles surrounding the soft palate may be more tense in initial position, resulting in a somewhat higher velum position, in spite of the fact that the nasalization feature is characterized by a lowered velum position and should manifest maximally by a more (deeply) lowered position. [Fujimura 1990: 233].*

- For vowels, greater articulatory force predicts different variations, depending on which muscles are considered to be dominant in vowel articulation. To Straka, the dominant muscles are the lowering muscles of the tongue:

> *...vowels' articulatory gestures are essentially controlled by the contraction of the hyoglossus, the lowering muscles in the tongue body, and by the contraction of the lowering muscles in the tongue tip. [Straka 1963: 90, our translation].*

An increase in the contraction of these muscles would thus cause a greater lowering of the tongue. Thus, even close vowels would be more open as a result of articulatory strengthening (which Straka observes in reinforced speech). However, as Straka notes, vowel articulation is also driven by the activation of raising muscles, such as the palatoglossus and the genioglossus for front vowels (UCLA 1990). To Straka, the activity of these raising muscles is secondary and thus not affected by articulatory strengthening:

> *...articulatory strengthening only affects those muscles that perform the main task: raising muscles for consonants and lowering muscles for vowels. [Straka 1963: 92, our translation].*

On the other hand, in the production of close vowels, if the activity of raising muscles is considered to be at least as important as that of lowering muscles, then articulatory strengthening will have different effects on close and open vowels. In close vowels, the tongue will be higher if strengthening mainly affects raising muscles. However, if strengthening equally affects the two types of antagonist muscles, the tongue cannot change position, since the strengthening will act equally in both directions.

IV.D. SUMMARY OF THE ARTICULATORY VARIATIONS PREDICTED BY THESE THREE HYPOTHESES

The sonority expansion hypothesis and the distinctive feature enhancement hypothesis emphasize the *linguistic function* of articulatory variations: they emphasize either a syntagmatic contrast in a syllable (sonority), or a paradigmatic contrast between segments (distinctive features). Articulatory strengthening, on the other hand, is not initially considered as a functional mechanism linguistically, but as a *motor mechanism*. It is a physiological mechanism that would act at the encoding level for a particular position: the initial position. This does not exclude the possibility that the articulatory consequences of initial strengthening are used linguistically (e.g. resistance to reduction as seen in section III).

Table I.2 lays out the predictions on articulatory variation in initial position resulting from these three hypotheses. We only present the predictions made for lingual and nasal articulation, which will be the main point of this study. The last column in the table shows the segments that I concentrate on in Fougeron (1998).

- The variations predicted by the sonority expansion hypothesis (H1) and the articulatory strengthening hypothesis as formulated by Straka (H3) are the same. As seen in figure 1.14, Straka has shown that greater articulatory force enhances the aperture contrast between consonants and vowels since consonants become more constricted and vowels become more open. This corresponds to a greater sonority contrast. However, if one considers the role of raising muscles in the production of close vowels (H4), these two hypotheses make different predictions with respect to tongue gestures. In Fougeron (1998), I examine the behavior of the close vowel /i/ in initial position to test these two hypotheses. Lehiste (1964) shows acoustic results which would tend to confirm H4 but not Straka's view: initial vowels are more peripheral in Finnish.
- The predictions made by the distinctive feature enhancement hypothesis (H2) differ from those made by the articulatory strengthening hypothesis (H3) essentially with respect to velum gestures. For nasal segments (vowels and consonants), an enhancement of the nasal feature predicts a lowering of the velum, whereas articulatory strengthening (H3) or increased initial tension (H5) predict a raising of the velum. In Fougeron (1998) I study the behavior of the velum with aerodynamic data for a nasal vowel and a nasal consonant. For tongue gestures in close vowels, the distinctive feature enhancement hypothesis (H2) predicts more raising of the tongue and thus opposes the predictions made by the sonority expansion hypothesis (H1) or by the articulatory strengthening hypothesis as proposed by Straka (H3).

Table 1.2: Articulatory variations predicted by the different hypotheses for the tongue and the velum as a function of segment type. (1) shows the sonority (and CV contrast) expansion hypothesis [Beckman et al. 1994]. (2) shows the distinctive feature enhancement hypothesis [de Jong 1995]. (3) shows the articulatory strengthening hypothesis as formulated by Straka (1963). (4) shows the articulatory strengthening hypothesis, taking into account the action of raising muscles in vowel production. (5) shows the initial tension hypothesis for the velum (Vaissière 1986a, 1988). The symbols
↑ and ↓ stand for the raising and lowering of an articulator, respectively. Shaded cells indicate contradictory predictions. The last column shows the segments examined in Fourteron (1998).

	tongue				velum				
	1	2	3	4	1	2	3	5	
Stops									
oral	↑	↑	↑	↑		↑	↑	↑	/t k/
nasal	↑	↑	↑	↑	↑	↓	↑	↑	/n/
lateral	↑	↑	↑	↑				↑	/l/
Fricatives	↑	↑	↑	↑				↑	/s/
Vowels									
open	↓	↓	↓	↓			↑	↑	
close	↓	↑	↓	↑ =			↑	↑	/i/
nasal						↓	↑	↑	/ɑ̃/

Acknowledgements

I would like to deeply thank Roger Billerey for his great translation of this chapter, Moto Ueyama for invaluable help with ms. preparation, and Pat Keating for making this publication possible. I am also grateful to my committee members, Jean-Marie Hombert, Shinji Maeda, Alain Marchal, and particularly Jacqueline Vaissière and Pat Keating, for their comments on my thesis.

Bibliography

Abbreviations:

ASA: Acoustical Scoety of America meeting
C.U.P.: Cambridge University Press
ICPhS: International Congress of Phonetic Sciences
ICSLP: International Conference on Spoken Language Processing
JASA: Journal of the Acoustical Society of America
JEP: Journés d'Études sur la Parole
JSHR: Journal of Speech and Hearing Research
LSA: Lingusitic Society of America meeting
MIT QPR: MITQuarterly Progress Report
STL-QPSR: Speech Trans. Lab. Quarterly Progress and Status Report, Stockholm

Beckman, M.E. & Edwards, J. (1990). Lengthening and shortening and the nature of prosodic constituency. In J. Kingston & M.E. Beckman (eds.) *Papers in Laboratory Phonology I: Between Grammar and Physics of Speech.* Cambridge: C.U.P..

Beckman, M.E. & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P.A. Keating (ed.) *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form,* 7-33. Cambridge: C.U.P..

Beckman, M.E. & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook,* **3,** 255-309.

Beckman, M.E. & Edwards, J., and Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. Docherty & D.R. Ladd (eds.) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody,* 68-86. Cambridge: C.U.P..

Bell, A. & Hooper, J. (1978). Issues and evidence in syllabic phonology. In A. Bell & J. Hooper (eds.) *Syllables and Segments,* 3-24. Amsterdam: North-Holland Publishing Co..

Bell-Berti, F. (1993). Understanding velic motor control: studies of segmental context. In M.K. Huffman & R.A. Krakow (eds.) *Nasals, Nasalization and the Velum,* 63-86. San Diego: Academic Press.

Bell-Berti, F. & Hirose, H. (1975). Palatal activity in voicing distinctions: a simultaneous fiberoptic and EMG study. *Journal of Phonetics,* **3**, 69-74.

Benguerel, A.-P. (1973). Corrélats physiologiques de l'accent en français. *Phonetica* **27**, 21-35.

Benguerel, A.-P. (1977). Velar coarticulation in French: a fiberscopic study. *Journal of Phonetics,* **5**, 145-158.

Bolinger, D. (1958). A theory of pitch accent in English. *Word,* **14**, 109-149.

Bonnot, J.F.P., Chevrie-Muller, C., Arabia-Guidet, C., Maton, B., & Grenier, G.F. (1986). Coarticulation and motor encoding of labiality and nasality in CVCVCV nonsense words. *Speech Communication,* **5**, 83-95.

Bourciez, E. & Bourciez, J. (1967). *Phonétique française, étude historique.* Paris: Klincksieck.

Brichler-Labaeye, C. (1970). *Les voyelles françaises: Mouvements et positions articulatoires à la lumière de la radiocinématographie.* Paris: Klincksieck.

Browman, C. (1978). Tip of the tongue and slip of the ear: implications for language processing. *UCLA Working Papers in Phonetics*, **42**.

Browman, C. & Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica,* **49**, 155-180.

Browman, C. & Goldstein, L. (1995). Gestural syllable position effects in American English. In F. Bell-Berti & L.J. Raphael (eds.) *Producing speech: A Festschrift for Katherine Safford Harris*, 19-34. Woodbury, N.Y.: American Institute of Physics Press.

Brown, R. & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior,* **5**, 325-337.

Brunot, F. & Bruneau, C. (1937). *Précis de grammarie historique de la langue français.* Paris: Masson.

Byrd, D. (1994). *Articulatory timing in English consonant sequences.* PhD. diss., UCLA (distributed as *UCLA Working Papers in Phonetics* **86**).

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics,* **24 (2)**, 209-244.

Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (1996). Phrasal signatures in articulations. Paper presented at *the 5th Conference on Laboratory Phonology*, Evanston, USA (to appear in *Papers in laboratory phonology V*).

Byrd, D. & Saltzman, E. (1996). Distinguishing multiple prosodic boundaries in articulatory kinematics and dynamics. *JASA,* **100 (4 pt.2)**, 2823.

Byrd, D. & Saltzman, E. (1998). Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics, 26,* 173-199.

Carlson, R. & Grandström, B. (1995). Word accent, emphatic stress, and syntax in a synthesis by rules scheme for Swedish. *STL-QPRS, 2/3, 31-35.*

Caspers, J. & van Heuven V.J. (1991). Phonetic and linguistic aspects of pitch movements in fast speech in Dutch. *Proceedings of the 12th ICPhS,* **5**, 174-77. Aix en Provence, France.

Caspers, J. & van Heuven V.J. (1995). Effects of time pressure on the choice of accent-lending and boundary-making pitch configuration in Dutch. *Proceedings of Eurospeech,* **2**, 1001-1004. Madrid, Spain.

Chlumsky, G. (1938). *Radiografie Francouzskych Samohlasek a Polosamohlasek.* Prague: Czeck Academy.

Chomsky, N. & Halle, M. (1968). *The Sound Pattern of English.* New York: Harper & Row.

Church, K.W. (1987). Phonological parsing and lexical retrieval. In U. Frauenfelder & L. Komisarjewsky-Tyler (eds.) Spoken word recognition. *Cognition,* **25**, 53-69.

Clements, G.N. & Keyser, S.J. (1983). *CV Phonology: A Generative Theory of the Syllable.* Cambridge, Mass.: MIT Press.

Cooper, A. (1991). Laryngeal and oral gestures in English /p, t, k/. *Proceedings of the 12th ICPhS,* **2**, 50-53. Aix en Procence, France..

Cooper, W.E. (1976). *Syntactic Control of Timing in Speech Production.* PhD diss., MIT.

Crompton, A. (1980). Timing patterns in French. *Phonetica,* **37**, 205-234.

Crystal, T. & House, A.S. (1990). Articulation rate and the duration of syllables and stress groups in connected speech. *JASA*, **88 (1)**, 101-112.

Debrock, M. (1977). An acoustic correlate of the force of articulation. *Journal of Phonetics,* **5**, 61-80.

De Jong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *JASA,* **97 (1)**, 491-504.

De Jong, K., Beckman, M.E., & Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and Speech,* **36 (2, 3)**, 197-212.

Delais-Roussarie, E. (1995). *Pour une approche parallèle de la structure prosodique.* Thèse de doctorat, U.

Toulouse-le Mirail.

Delattre, P. (1938). L'accent final en français: accent d'intensité, accent de hauteur, accent de durée. *The French Review,* **12 (2)**, 141-145.

Delattre, P. (1939). Accent de mot et accent de groupe. *The French Review,* **13 (2)**, 1-6.

Delattre, P. (1940). La force d'articulation consonantique en français. *The French Review,* **14**, 220-232.

Delattre, P. (1965). *Comparing the Phonetic Features of English, French, German and Spanish: An Interim Report.* New York: Chilton Books.

Delattre, P. (1966). Les dix intonations de base du français. *The French Review,* **40 (1)**, 1-14.

Dell, F. (1984). L'accentuation dans les phrases en français. In Dell, F., Hirst, D. & Vergnaud, J.R. (eds.) *La forme sonore du langage: la nature des représentations en phonologie.* Paris:Herman.

Di Cristo, A. (1978). *De la microprosodie à l'intonosyntaxe.* Thèse de doctorat, U. de Provence. (published in 1985).

Di Cristo, A. & Hirst, D. (1993). Rhythme syllabique, rhythme mélodique et représentation hiérarchique de la prosodie du français. *Travaux de l'Inst. de Phon. d'Aix,* **15**, 9-24.

Di Cristo, A. & Hirst, D. (1996). Vers une typologie des unités intonatives du français. *Actes des XIe Journées d'Études sur la Parole,* 219-222. Avignon, France.

Di Cristo, A. & Hirst, D. (1997). L'accentuation non-emphatique en français: stratégies et paramètres. In J. Perrot (ed.) *Polyphonie pour Ivan Fonagy,* 71-101. Paris: L 'Harmattan.

Dilley, L. & Shattuck-Hufnagel, S. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics,* **24**, 423-444.

Duez, D. (1987). *Contribution à l'etude de la duré vocalique, le timbre et la structure syllabique en français.* Thèse de doctorat U. Aix Marseille.

Duez, D. (1995). On spontaneous French speech: aspects of the reduction and contextual assimilation of voiced stops. *Journal of Phonetics,* **223**, 407-427.

Edwards, J. & Beckman, M.E. (1988). Articulatory timing and the prosodic interpretation of syllable duration. *Phonetica,* **45**, 156-174.

Erickson, D. & Fujimura, O. (1992). Acoustic and articulatory correlates of contrastive emphasis in repeated corrections. *Proceedings of ICSLP,* 835-837. Banff, Canada.

Erickson, D. & Fujimura, O. (1996). Maximum jaw displacement in contrastive emphasis. *Proceedings of ICSLP (CD-ROM).* Philadelphia, USA.

Fant, G. & Kruckenberg, A. (1991). Language specific patterns of prosodic and segmental structures in Swedish, French, and English. *Proceedings of the 12th ICPhS,* **4**, 118-121. Aix en Provence, France.

Farnetani, E. (1986). A pilot study of the articulation of /n/ in Italian using electropalatography and airflow measurements. *Actes des XVe J.E.P.,* 23-26. Aix en Procence, France.

Farnetani, E. & Vayra, M. (1996). The role of prosody in the shaping of articulation in Italian CV syllables. *Proceedings of the 1st ESCA Workshop on Speech Production Modeling, 4th Speech Production Seminar,* 9-12. Autrans, France.

Faure, G. (1962). L'intonation et l'identification des mots dans la chaîne parlée (exemples empruntés à la langue française). *Proceedings of the 4th ICPhS,* 598-609. Helsinki, Finland.

Fletcher, J. (1991). Rhythm and final lengthening in French. *J. of Phonetics,* **19**, 193-212.

Fletcher, J. & Vatikiotis-Bateson, E. (1994). Prosody and intrasyllabic timing in French. *Ohio State U. Working Papers,* **43**, 41-46.

Fougeron, C. & Jun, S.-A. (1995). Properties of French intonation at fast speech rate. *Proceedings of the 13th ICPhS,* **3**, 488-491. Stockholm, Sweden.

Fougeron, C. & Jun, S.-A. (1998). Rate effects on French intonation: prosodic organization and phonetic realization. *Journal of Phonetics,* **26** (1), 45-70. (preliminary version in UCLA Working Papers in Phonetics, 94, 61-87 (1996))

Fougeron, C. & Keating, P.A. (1995). Demarcating prosodic groups with articulation. *JASA,* **97 (5 pt2)**, 3384.

Fougeron, C. & Keating, P.A. (1996). Articulatory strengthening in prosodic domain-initial position. *UCLA Working Papers in Phonetics,* **92**, 61-87.

Fougeron, C. & Keating, P.A. (1997). Articulatory strengthening at edges of prosodic domains. *JASA,* **106 (6)**, 3728-3740.

Fritzel, B. (1969). The velopharyngeal muscles in speech. *Acta Oto-Laryngology,* suppl. **250**.

Fromkin, V. (1965). Some phonetic specifications of linguistic units: an electromyographic investigation. *UCLA Working Papers in Phonetics,* **3**, 1-184.

Fry, D. B. (1965). Experiments in the perception of stress. *Language and Speech,* **1**, 126-152.

Fujimura, O. (1990). Methods and goals of speech production research. *Languages and Speech,* **33 (3)**, 195-258.

Fujimura, O. & Erickson, D. (1997). Acoustic phonetics. In W.J. Hardcastle & J. Laver (eds.) *The Handbook of Phonetic Sciences,* 65-115. Oxford: Blackwell.

Fujimura, O. & Lovins, J.B. (1978). Syllables as concatenative phonetic units. In Bell, A. & Hooper, J.B. (eds.) *Syllables and Segments.* 107-120. Amsterdam: North-Holland Publishing Co..

Garde, P. (1968). *L'accent.* Paris: Presses Universitaires de France.

Gårding, E. (1967). Internal juncture in Swedish. *Lund Working Papers.,* **VI**.

Gay, T. (1978). Physiological and acoustic correlates of perceived stress. *Languages and Speech,* **21 (4)**, 347-353.

Giot, J. (1977). Etude comparative de syllabes accentuées et prétoniques du français sur les plans articulatoire et acoustique. *Travaux de l'Inst. de Phon. de Strasbourg,* **9**, 89-169.

Goldsmith, J. (1990). *Autosegmental and Metrical Phonology.* Oxford: Blackwell.

Goldstein, L. (1992). Comments on chapters 3 and 4. In G. Docherty and D.r. Ladd (eds.) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody,* 120-124. Cambridge: C.P.S.

Gordon, M. (1966). The effects of stress and prosodic phrasing on duration, acoustic amplitude and air flow of nasal consonants in Estonian. *UCLA Working Papers in Phonetics,* **92**, 151-159.

Gow, D.W., Melvold, J., Manuel, S. (1996). How word onsets drive lexical access and segmentation: evidence from acoustics, phonology and processing. *Proceedings of ICSLP* (CD-ROM). Philadelphia, USA.

Grammont, M. (1933). *Traité de Phonétique.* Paris: Delagrave.

Grosjean F. & Deschamps, A. (1972). Analyse des variables temporelles du français spontanée. *Phonetica,* **26**, 129-156.

Grosjean, F., Grosjean, L. & Lane, H. (1979). The pattern of silences: performance structures in sentence production. *Cognitive Psychology,* **11**, 58-81.

Haden, E.F. (1938). The physiology of French consonant changes. *Language dissertation, Language (suppl.),* **26**, vol. **14 (4)**, 3-117.

Hardcastle, W. (1985). Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication,* **4**, 1-40.

Harrington, J., Fletcher, J., Beckman, M.E. (1996). Manner and place conflicts in the articulation of accent in Australian English. Paper presented in *the 5th conference on Laboratory Phonology,* Evanston, USA (to appear in *Papers in Laboratory Phonology V*).

Harris, J. (1990). Segmental complexity and phonological government. *Phonology,* 7, 255-300.

Harris, K.S., Gay, T., Sholes, G.N., & Liberman, P. (1968). Some stress effects on electromyographic measures of consonant articulations. *Haskins Laboratory Status Report,* **13/14**, 137-151.

Hayes, B. (1989). The prosodic hierarchy in meter. In P. Kiparsky & G. Youmans (eds.) *Phonetics and phonology, vol.I: Rhythm and meter,* 47-96. New York: Academic Press.

Hirst, D. & Di Cristo, A. (1984). French intonation: a parametric approach. *Die Neuren Sprachen,* **83 (5)**, 554-569.

Hirst, D. & Di Cristo, A. (1996). Y-a-t-il des unités tonales en français? *Actes des XIe JEP,* 223-226. Avignon, France.

Hjemslev, L. (1936). Accent, intonation, quantité. *Studi Baltici,* **66**, 1-57.

Hock, H.H. (1992). Initial strengthening. In W.U. Dressler et al. (eds.) *Phonologia 1988.* Cambridge, Mass.: C.U.P.

Holst, T. & Nolan, F. (1995). The influence of syntactic structure on [s] to [ʃ] assimilation. In B. Connell & A. Arvaniti (eds.) *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidences,* 315-333. Cambridge: C.U.P..

Horowitz, L., Chilian, P.C. & Dunnigan, K.P. (1969). Word fragments and their reintegrative powers. *Journal of Experimental Psychology,* **80**, 392-404.

Houde, R.A. (1967). *A Study of Tongue Body Movement During Selected Speech Sounds.* Ph.D. Diss., U. of Connecticut.

Hsu, C.-S.& Jun, S-A. (1997). Prosodic strengthening in Taiwanese: syntagmatic or paradigmatic? Unpublished manuscript, UCLA.

Ito, J. & Mester, A. (1983). Weak layering and word binarity. UC Santa Cruz, ms.

Jaegger, J.J. (1983). The fortis/lenis question: evidence from Zapotec and Jawoñ. *Journal of Phonetics,* 11, 177-189.

Jakobson, R., Fant, G. & Halle, M. (1969). *Preliminaries to speech analysis. The Distinctice Features and Their Correlates.* Cambridge, Mass.: MIT Press.

Jun, S.-A. (1993). *The Phonetics and Phonology of Korean Prosody.* PhD diss. Ohio State University.

Jun, S.-A. & Fougeron, C. (1995). The accentual phrase and the prosodic structure of French. *Proceedings of the 13th ICPhS,* vol. **2**, 722-725. Stockholm, Sweden.

Jun, S.-A. & Fougeron, C. (1997). A phonological model of French intonation. Paper presented in the *ESCA Workshop on Intonation.* Athens, Greece. (to appear in A. Botinis (ed.) *Intonation research and applications* (temporary title), Cambridge UK:C.U.P.)

Keating, P.A. (1995). Effects of prosodic position on /t, d/ tongue/palate contact. Poster presented at the *13th ICPhS* (unpublished version). Stockholm, Sweden.

Keating, P.A. (1997). Word-initial versus word-final consonant articulation. *JASA,* **102 (5 pt2)**, 3094.

Keating, P.A., Linker, W. & Huffman, M. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics,* **11**, 277-290.

Keating, P.A. & Wright, R. (1994). Effects of position-in-syllable on consonant articulation and acoustics. *JASA,* **95 (5 pt2)**, 2819.

Kent, R.D. & Netsell, R. (1971). Effects of stress contrast on certain articulatory parameters. *Phonetica,* **24**, 23-44.

Klatt, D.H. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics,* **3**, 129-140.

Klich, R.J., Ireland, J.V. & Weidner, W.E. (1979). Articulatory and phonological aspects of consonant substitutions in appraxia of speech. *Cortex,* **15**, 451-470.

Kohler, K.J. (1985). The perception of lenis and fortis plosives in French, a critical re-evaluation. *Phonetica,* **42**, 116-123.

Kohler, K.J. (1990). Segmental reduction in connected speech in German: phonological facts and phonetic explanations. In W. Hardcastle & A. Marchal (eds.) *Speech Production and Speech Modeling,* 69-992. Dordrecht: Kluwer.

Kohler, K.J. (1992). Gestural reorganization in connected speech: a functional viewpoint on "articulatory phonology". *Phonetica,* **49**, 205-21..

Kohler, K.J. (1994). Glottal stops and glottalization in German. *Phonetica,* **51**, 38-51.

Kohler, K.J. & Hardcastle, W. (1974). The instability of final alveolars in English and German: proposal for an instrumental investigation. *Speech Communication Seminar,* **2**, 95-98. Stockholm, Sweden.

Kozhevnikov, V.A. & Chistovich, L.A. (1965). *Speech: Articulation and Perception.* (US Dept. of Commerce, Washington D.C.) JPRS, **30**, 524.

Krakow, R.A.. (1989). *The articulatory organization of syllables; A Kinematic Analysis of Labial and Velic Gestures.* PhD diss., Yale U.

Krakow, R.A., Bell-Berti, F. & Wang, Q. (1991). Supralaryngeal patterns of declination: Labial and velar kinematics. *JASA,* 90 (4 pt.2), 2343.

Krakow, R.A., Bell-Berti, F. & Wang, Q. (1994). Supralaryngeal declination: evidence from the velum. In F. Bell-Berti & Raphael L.J. (eds.) *Producing speech : A Festschrift for Katherine Safford Harris*, 333-353. Woodbury, N.Y.: American Institute of Physics Press.

Ladd, D.R. (1984). Declination: a review and some hypotheses. *Phonology Yearbook,* **1**, 53-74.

Ladd, D.R. (1986). Intonational phrasing: the case of recursive prosodic structure. *Phonology Yearbook,* 3, 311-340.

Ladd, D.R. (1988). Declination 'reset' and the hierarchical organization of utterances. ASA, 84 (2), 530-544.

Ladd, D.R. (1996). *Intonational Phonology.* Cambridge UK: C.U.P..

Lehiste, I. (1961). Acoustic studies of boundary signals. *Proceedings of the 4th ICPhS,* 178-187. Helsinki, Finland.

Lehiste, I. (1964). Juncture. *Proceedings of the 5th ICPhS,* 172-200. Münister, Germany.

Lehiste, I. (1970). Temporal organization of higher level linguistics units. Paper presented at ASA, Atlantic City, USA. (unpublished).

Lehiste, I. (1972). The timing of utterance and linguistic boundaries. *JASA,* **51**, 2018-2024.

Lewis, J., Daniloff, R. & Hammarberg, R. (1975). Apical coarticulation at juncture boundaries. *Journal of Phonetics,* 3, 1-7.

Liberman, M. (19785). *The intonational system of English.* PhD diss., MIT.

Liberman, M. & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry,* **8**, 29-336.

Liberman, P. (1960). Some acoustic of word stress in American English. *JASA*, **32**, 451-454.

Liberman, P. & Blumstein, S. (1988). *Speech Physiology, Speech Perception, and Acoustic Phonetics.* Cambridge UK: C.U.P..

Lindblom, B. (1968). Temporal organization of syllable production. STL-QPSR, **2**, *1-5.*

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W.J. Hardcastle & A. Marschal (eds.) *Speech Production and Speech Modeling*, 403-439. Dordrecht: Kluwer.

Lindblom, B. & Rapp, K. (1973). Some temporal regularties of spoken Swedish. *Papers in Lingusitics, U. of Stockholm,* **21**, 1-59.

Lisker, L. & Baer, T. (1984). Laryngeal management at utterance-internal word boundary in American English. *Language and Speech,* **27**, 163-171.

Lubker, J.F., Fritzell, B. & Lindqvist, J. (1970). Velopharyngeal function: an electromyographic study. *STL-QPSR,* **4**, 9-20.

Lucci, V. (1983). *Etude phonétique du français contemporain à travers la variation situationnelle.* Grenoble: Publ. U. de Grenoble.

Macchi, M. (1988). Labial coarticulation pattern associated with segmental features and syllable structure in English. *Phonetica,* **45**, 109-121.

MacClean, M. (1973). Forward coarticulation of velar movement at marked junctural boundaries. *Journal of Speech and Hearing Research*, **16**, 286-296.

Malécot, A. (1955). An experimental study of force of articulation. *Studia Linguistica,* **9**, 35-44.

Malécot, A. (1960). Vowel nasality as a distinctive feature in American English. *Language,* **36**, 222-229.

Malécot, A. (1970). The lenis-fortis opposition: its physiological parameters. *JASA,* **47**, 1588-1593.

Malmberg, B. (1950). The phonetic basis for syllable division. *Studia Linguistica,* **IX**, 80-87.

Malmberg, B. (1971). Voyelle, consonne, syllabe, mot. In *Phonétique Générale et Romane,* 131-140. The Hague-Paris: Mouton.

Manuel, S. (1991). Some phonetic bases for the relative malleability of syllable-final vs. syllable-initial consonants. *Proceedings of the 12th ICPhS,* **5**, 118-121. Aix en Provence, France.

Marchal, A. (1979). Etude des occlusives t/d du français par l'electropalatographie. *Proceedings of the 9th ICPhS,* **1**, 203. Copenhagen, Denmark.

Marslen-Wilson, W.D. (1980). Speech understanding as a psychological process. In J.C. Simon (ed.) *Spoken Language Generation and Understanding.* Dordrecht: Reidel.

Martin, P. (1987). Prosodic and rhythmic structures in French. *Linguistics,* **25**, 925-949.

Martinet, A. (1955). *Economie des changements phonétiques: traité de phonologie diachronique.* Berne: Francke SA.

McGlone, R. & Proffit, W.R. (1967). Lingual pressure associated with speaker consistency and syllable variations. *Phonetica,* **17**, 178-183.

McGlone, R., Proffit, W.R. & Christiansen, R.L. (1967). Lingual pressure associated alveolar consonants. *JSHR,* **10 (3)**, 606-615.

Mertens, P. (1987). *L'intonation du français, de la description linguistique à la reconnaissance automatique.* Thèse de doctorat, U. Leuven.

Meynadier, Y., Marchal, A. & Piterman, M. (1998). Frontière syllabique, accent focal constrastif et coarticulation linguale dans le groupe consonantique /kskl en position m´diane de mot en français. *Actes des XXXII2 JEP,* 303-306. Martigny, Switzerland.

Nakatani, L.H. & Dukes, K.D. (1977). Locus of segmental cues for word juncture. *JASA,* **62 (3)**, 714-719.

Nakatani, L.H. & Schaffer, J.A. (1978). Hearing "words" without words: prosodic cues for word perception. *JASA,* **63 (1)**, 234-245.

Nespor, M. (1977). Segment and word durational correlates of syntactic boundaries in Italian. In R. Rodman (ed.) *Explorations in Language and Linguistics.* Chapel Hill: U. of North Carolina.

Nespor, M. & Vogel, I. (1986). *Prosodic Phonology.* Dordrecht: Foris Publications.

Nooteboom, S.G. (1972). *Production and Perception of Vowel Duration: A Study of Durational Properties of Vowels in Dutch.* PhD diss., Utrecht.

Nooteboom, S.G. (1981). Lexical retrieval from fragments of spoken words: beginnings vs. endings. *Journal of Phonetics,* **9**, 407-424.

Ohala, J. & Kawasaki, H. (1984). Prosodic phonology and phonetics. *Phonology Yearbook,* **1**, 113-127.

Oller, D.K. (1973). The effect of position in utterance on speech segment duration in English. *JASA,* **54 (5)**, 1235-1247.

Pasdeloup, V. (1990). *Modèle de règles rhythmiques du français appliquées à la synthèse de la parole.* Thèse de doctorat, U. de Provence.

Pierrehumbert, J. & Beckman, M.E. (1988). *Japanese Tone Structure..* Cambridge, Mass.: MIT Press.

Pierrehumbert, J. & Talkin, D. (1992). Lenition of /h/ and glottal stop. IN G. Docherty and D.R. Ladd (eds.) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody,* 90-117.

Pike, K.L. (1943). *Phonetics: A Critical Analysis of Phonetic Theory and a Technique for the Practical Description of Sounds.* Ann Arbor: U. of Michigan Press.

Post, B. (1993). *A Phonological Analysis of French Intonation.* MA thesis, U. Nijmegen.

Quené, H. (1992). Durational cues for word segmentation in Dutch. *Journal of Phonetics,* **20**, 331-350.

Rietveld, A.C.M. (1980). Word boundaries in the French language. *Language and Speech,* **23 (3)**, 289-296.

Rigault, A. (1962). Rôle de la fréquence, de l'intensité et de la durée vocalique dans la perception de l'accent en français. *Proceedings of the 4th ICPhS,* 735-748. Helsinki, Finland.

Rossi, M. (1980). Le français, langue sans accent? In Fonagy, I. & Léon, L. (eds.) *L'accent en français contemporain. Studio Phonetica,* **15**, 13-51.

Rossi, M. (1985). L'intonation et l'organisation de l'énoncé. *Phonetica,* **42** , 135-153.

Rossi, M., Di Cristo, A., Hirst, D., Martin, P.H. & Nishinuma, Y. (1981). *L'intonation de l'acoustique à la sémantique.* Paris: Klincksieck.

Rousselot, P.J. (1901). *Princepes de phonétique expérimentale.* Paris: Welter.

Sabio, F. (1996). *Description prosodique et syntaxique du discours en français. Données et hypothèses.* Thèse de doctorat, U. de Provence.

Saussure, F. (1915) *Cours de linguistique générale.* Paris: Payot.

Selkirk, E. (1980). *Phonology and Syntax: The Relation between Sound and Structure.* Cambridge, Mass.: MIT Press.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook,* **3**, 371-405.

Shattuck-Hufnagel, S. (1986). The representation of phonological information during speech production planning: evidence from vowel errors in spontaneous speech. *Phonology Yearbook,* **3**, 117-149.

Shattuck-Hufnagel, S. & Turk, A.E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research,* **25 (2)**, 193-247.

Simon, P. (1967). *Les consonnes françaises: mouvements et positions articulatoires à la lumière de la radiocinématographie.* Paris: Klincksieck. (et These de doctorat, U. de Strasbourg).

Slis, I.H. (1971). Articulatory effort and its durational and electromyographic correlates. *Phonetica,* **23 (3)**, 171-188.

Sluijter, M.C. (1995). *Phonetics Correlates of Stress and Accent.* PhD diss., U. of Leyden.

Stone, M. (1981). Evidence for a rhythm pattern in speech production : observations of jaw movement. *Journal of Phonetics,* **9**, 109-120.

Straka, G. (1963). La divison de sons du language en voyelles et consonnes peut-elle être justifiée? *Travaux de Ling. et de Littérature, U. de Strasbourg,* **1**, 17-99.

Straka, G. (1964). L'évolution phonétique du latin au français sous l'effect de l'énergie et de la faiblesse articulatoires. *Travaux de Ling. et de Littérature, U. de Strasbourg,* **2**, 17-98.

Thurneysen, R. (1946). *A Grammar of Old Irish.* Dublin: Inst. for Advanced Studies.

Touati, P. (1987). Structures prosodiques du suédois et du français. *Lund U. Working Papers,* **21**.

Trask, R.L. (1996). *A Dictionary of Phonetics and Phonology.* New York: Routledge.

Troubetskoy, N,S, (1949). *Grundzüge der phonologie.* Travaux du Cercle Linguistique de Prague VII.

UCLA Phonetics Laboratory. (1990). Dissection of the speech production mechanism. *UCLA Working Papers in Phonetics,* 77.

Umeda, N. (1977). Consonant duration in American English. *JASA,* **61**, 846-858.

Vassière, J. (1975). Further note on French prosody. *MIT QPR,* **115**, 251-261.

Vassière, J. (1977). Premiers essais d'utilisation de la durée pour la segmentation en mots dans un système de reconnaissance. *Actes des VIIIe JEP,* 345-352. Aix en Provence, France.

Vassière, J. (1983). Language-independent prosodic features. In Cutler, A. & Ladd, D.R. (eds.) *Prosody: Models and Measurements.* Berlin: Springer-Verlag, 53-66.

Vassière, J. (1986a). Variance and invariance at the word level. Invited contribution, in Perkell, J.S. & Klatt, D.H. (eds.) *Invariance and Variability in Speech Processes,* 534-553. Hillsdale: Erlbaum Ass.

Vassière, J. (1986b). Invariance and variability in speech production: a distinction between linguistic intent and its neuromotor implementation, Comment on Abbs' paper. In Perkell, J.S. & Klatt, D.H. (eds.) *Invariance and Variability in Speech Processes,* 534-553. Hillsdale: Erlbaum Ass.

Vassière, J. (1988). Prediciton of velum movement from phonological specifications. *Phonetica,* **45**, 122-139.

Vassière, J. (1989). *Contribution à l'analyse des phénonmène de parole continue lue.* Habilitation à diriger des recherches, U. de Strasbourg (unpublished).

Vassière, J. (1992). Rhythm, accentuation and final lengthening in French. In J. Sundberg, Nord, L. & Carlson (eds.) *Music, Language, Speech and Brain,* 108-120. Stockholm: Wenner-Gren.

Vassière, J. (1996). From Latin to modern French: on diachronic changes and synchronic variations.

Arbeitsberichte, U. Kiel, **31**, 61-74.

Vassière, J. (1997). Langage, prosodies et syntaxe. *Revue Traitement Automatique des Langues, ATALA,* **38 (1)**, 53-82.

van Lieshout, P., Starweaker, C.W., Hulsijn, W. & Peters, H.F.M. (1995). Effects of linguistic correlates of stuttering on EMG activities in nonstuttering speakers. *JSHR,* **38**, 360-372.

Vatikiotis-Bateson, E. (1988). *Linguistic Structure and Articulatory Dynamics.* Bloomington: Indiana U. Linguistic Club.

Vatikiotis-Bateson, E. & Fowler, C. (1988). Kinematic analysis of articulatory dynamics in English, French and Japanese. *Journal of Phonetics,* **21**, 231-265.

Vayra, M. & Fowler, C. (1992). Declination of supralaryngeal gestures in spoken Italian. *Phonetics,* **49**, 48-60.

Vennemann, T. (1993). Language changes as language improvement. In Jones, C. (ed.) *Historical Linguistics: Problems and Perspectives.* London-New York: Longman.

Verluyten, P. (1983). La structure de l'énoncé en groupes intonationnels. *ITL-Review of Applied Linguistics,* **60-61**, 77-104.

Wightman, C.W., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P.J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *JASA,* **91 (3)**, 1707-1717.

Willbur, R. & Zelanik. (1997). Kinematic correlates of stress and phrase position in ASL. Paper presented at the *71st Annual Meeting of the LSA.* Chicago, USA.

Wright, R. (1994). Coda lenition in American English consonants: an EPG study. JASA, **95 (5 pt.2)** , 2819.

Zink, G. (1986). *Phonétique historique du français.* Paris: Presses Universitaires de France.

Articulatory properties of initial segments in several prosodic constituents in French.

Cécile Fougeron

Lab. de Phonétique, ILPGA, Univ. Paris III
&
Lab. de Psycholinguistique Expérimentale, Univ. de Genève

ABSTRACT

This study reports on how what is usually called "segmental" articulation may be affected by prosodic structure. Articulatory properties of segments of various types are compared in initial position of four prosodic constituents in French: an Intonational Phrase, an Accentual Phrase, a Word and a Syllable. Modification of linguopalatal contact, nasal flow, acoustic measures of glottal articulation, as well as durational properties of segments are examined depending on these prosodic positions. Results show that the articulation of initial segments varies depending on the prosodic level of the constituent. Initial segments in higher prosodic domains tend to have more linguopalatal contact, somewhat less nasal flow for nasals, and vowels tend to be more frequently glottalized than in initial position in lower domains. The magnitude and the realization of the articulatory modification observed in these positions is found to depend on speakers, articulators, and segment types. However, the regular trend observed suggests that these articulatory properties can reflect the prosodic encoding of constituents of various levels within utterances. Correlation of articulatory properties with initial lengthening showed that the temporal properties of initial segments can not account for all of the variance found in their articulatory properties. Moreover, comparison with articulatory variations described in the literature for accented position, for a strengthened style of speech, or for strong segments suggests that the properties of initial segments result from a local articulatory strengthening in initial position that progressively varies depending on the prosodic level of the constituent.

1. Introduction

As early as 1901, the Abbé Rousselot noted in his *Principes de phonétique expérimentale* that the nature of morpho-syntactic phrases and their boundaries influences a segment's articulation. He concluded that one must consider "les nombreuses variétés auxquelles une même consonne est soumise en raison de sa place dans le mot ou dans la phrase" (the many forms a consonant can have because of its position in a word or a sentence, p.601).

Since then, several studies have shown that the production of a sound can be influenced by the structure in which it is placed. Sometimes known as "articulatory prosody", this trend of studies is concerned with the relationship between the segmental and suprasegmental (or prosodic) aspects of speech, and more particularly with the influence of the latter on the former. Most of the prosodically-conditioned segmental variations that have been reported so far are the ones concerned either with the difference between accented and unaccented positions, or with the distinction between initial and final position in a syllable or a word (see Fougeron, this volume, for a review). It is more recently that studies have looked at whether traces of the higher level hierarchical prosodic organization of an utterance could be found in the articulation of the segments it contains. The study presented here shares this objective. It aims at determining whether articulatory variations reflect the prosodic encoding of utterances into constituents (domains) of various levels in French.

French is considered to be a "boundary language" rather than a "stress language", following Vaissière's expression (1983, 1992), since *most* of its prosodic cues are associated with constituent boundaries. As a consequence, one can expect that if there are articulatory variations associated with the marking of prosodic structure, these could occur at constituent boundaries in such a language. Two positions around a boundary can be explored: the pre-boundary position which corresponds to the final position in a constituent, and the post-boundary position which corresponds to the initial position in a constituent. Pre-boundary, or domain-final, position has been shown in several studies and in many languages to be the site of realization of prosodic markers, such as final lengthening and intonational boundary contour or tone (among other, Delattre 1965, Klatt 1975, Vaissière 1983, Wightman et al. 1992). In this study, we will focus on the mirror position, that is the domain-initial or post-boundary position.

Evidence for a particular status of domain-initial position has been widely documented at the syllable and word levels. Initial segments in these constituents show a particular behavior regarding synchronic and diachronic variations (e.g. Bell & Hooper 1978, Ohala & Kawasaki 1984). They are particularly resistant to reduction or lenition processes. For example, in the sound changes from Latin to French, most of the syllable- and word-initial consonants have been maintained while medial or final ones have lenited or disappeared (Brunot & Bruneau 1937, Bourciez & Bourciez 1967). Initial segments have also been shown to have specific articulatory properties. Within a syllable or a word, the glottal and supraglottal articulations of a segment may vary depending on its position. Along with some well-known positional allophones, like the aspirated stops or the light form of /l/ occurring in initial position in English (e.g. Kahn 1976, Lehiste 1960), more subtle articulatory variations have been described in word- or syllable-initial position in many languages. Among others, initial consonants can have, compared to final ones, a larger glottal opening (Cooper 1991), a greater linguopalatal constriction (Rousselot 1901, Byrd 1994, Keating, Wright & Zhang 1999, this volume) or a higher velum position for both nasal (Fujimura 1977, Krakow 1989) and oral segments (Benguerel 1977, Vaissière 1988).

This study is aimed at extending these results by looking at constituents higher than the syllable and the word. My hypothesis is that initial position is, along with final position, a site governed by a specific prosodic encoding which is linked to phrasal organization. In order to test this hypothesis, a comparison between initial and final segments is not sufficient. Initial positions in different prosodic domains have to be compared. A series of recent studies have indeed shown that articulatory variations such as those described for initial consonants at the syllable or word level, do reflect the hierarchical level of the constituent studied. The width of glottal opening as approximated by VOT (Pierrehumbert & Talkin 1992 in English; Jun 1993, Cho & Keating 1999, this volume, in Korean, Hsu & Jun 1997 in Taiwanese), or the extent of linguopalatal contact (Fougeron & Keating 1997 in English, Cho & Keating 1999, this volume, in Korean), increases for segments initial in higher domains compared to segments initial in lower domains. Hence, as shown for final lengthening (e.g. Wightman, Shattuck-Hufnagel, Ostendorf & Price 1992), the progressive articulatory changes in initial position seem to reflect the hierarchical organization of prosodic constituents in distinguishing several levels of boundaries.

However, in most of these studies, it is noteworthy that the realization of these prosodically-dependent articulatory properties is quite variable. Table I summarizes some aspects of the studies that have reported articulatory variations in several domain-initial positions so far. It can be seen in the fourth column that not all the prosodic constituents studied are found to be distinguished by a variation in articulation. For example, Fougeron and Keating (1997) found in English that out of the five prosodic constituent considered, three to four were distinguished by the amount of linguopalatal contact of their initial segment. Furthermore, the three subjects studied did not distinguish the same constituents this way. For one subject, the amount of linguopalatal contact in /n/ was smaller in Syllable- and Word-initial positions, greater in Phonological Phrase-initial position, and even greater in Intonational Phrase-initial position. For another subject, there was a distinction between Syllable-initial and Word-initial /n/ but no distinction between Phonological Phrase- and Intonational Phrase-initial positions. A similar subject-dependent and constituent-dependent variability was found by Cho and Keating (1999, this volume) in Korean. Thus, it appears that the observed phenomenon needs to be examined more deeply. If the articulatory variations observed in domain-initial position do reflect the prosodic organization of an utterance, one needs to understand why these variations are not always present and whether they are just a side effect of some other prosodic markers.

Table I: Summary of the studies showing articulatory variation of segments in initial position of constituents of different prosodic levels. The language and the segments examined in these studies are given in column 2 and 3. In the last column is indicated the number of constituents observed to be distinguished by the articulatory measure out of all the constituents studied. The variability in this number indicates that not all speakers or all segments made the same distinctions.

Studies	**Language**	**Segments**	**# distinctions**
Lingual articulation			
Byrd et al. '96	Tamil	/n/	2 out of 3
Hsu & Jun '96	Taiwanese	/t/	2 to 3 out of 3
Fougeron & Keating '97	English	/n/	3 to 4 out of 5
Cho & Keating '99	Korean	/t/, /tʰ/, /t*/, /n/	3 to 5 out of 5
Nasal articulation			
Gordon '96	Estonian	/n/	2 to 3 out of 4
Glottal articulation: VOT			
Pierrehumbert & Talkin '92	English	/tʰ/	2 out of 2
Jun '93	Korean	/pʰ/	3 out of 3
Hsu & Jun '97	Taiwanese	/kʰ/, /b/ /t/	3 out of 3 0 out of 5
Cho & Keating '99	Korean	/t/, /tʰ/ /t*/	3 to 5 out of 5 0 out of 5
Glottal articulation: vowel glottalization			
Pierrehumbert & Talkin '92	English	/ɔ/ + divers	2 out of 2
Dilley et al. '96	English	divers	2 to 3 out of 3

The work presented here aims at providing a better understanding of this effect by examining in a more comprehensive fashion its realization in French. The main objective is to uncover the nature of the articulatory variations undergone by the segments in initial position. From the third column of table I, it appears that the influence of prosodic position on articulation has been studied mostly for stop consonants. It is only for glottal articulation that this effect has been extended to vowels (Pierrehumbert & Talkin 1992, Dilley & Shattuck-Hufnagel 1996). Hence, one must examine first if this effect generalizes to other types of segments. Second, the physiological mechanism governing these articulatory variations needs to be understood. In their study of lingual articulation, Fougeron & Keating (1997) have proposed to describe the increase in linguopalatal contact observed in domain-initial position as an *articulatory strengthening*. This term meant to reflect that the articulation of a consonant is more extreme (with a greater constriction for stops) in initial position compared to medial, and more extreme at the beginning of higher level constituents than at the beginning of lower ones. Indeed this "strengthening" appears as a mirror modification of the "weakening" known for final segments. In order to test this hypothesis, the nature (i.e. the directionality) of the articulatory variations in domain-initial position has to be examined in detail. In this study, I will look in a single language, French, at the nature of these variations both for several segments that differ in place and manner of articulation, and for three different articulatory subsystems (lingual, nasal, glottal).

2. Method

2.1. Test segments and prosodic positions

Five consonants varying in place and manner of articulation, /t, n, k, l, s/, and two vowels /i, ɑ̃/ were chosen for study. In order to examine their articulatory properties at different prosodic position, these test segments were placed in initial position in four different prosodic constituents: a Syllable (S), a lexical Word (W), an Accentual Phrase (AP), an Intonational Phrase (IP). Table II gives an example of the sentences used for the test segment /n/. Each prosodic position will be henceforth named with an abbreviation, where "i" stands for "initial" and the capital letters before stand for the name of the constituent (e.g. "APi" = "Accentual-Phrase initial"). In all sentences used in experiment 1 (lingual articulation), the serial position of the test segment within the sentence was controlled so that the segments appeared at the onset of the fifth syllable in every prosodic position (see example in table IIa). This was done in order to avoid a possible articulatory declination that would modify the articulation of the segment because of its early-to-late position in the sentence (see Krakow, Bell-Berti & Wang 1994, but see Fougeron & Keating 1997 for counter-evidence against articulatory declination). In the corpus used in experiment 2 (nasal flow), the serial position of the test segment was not as well controlled, and the test segments could appear between the 5th and 8th syllable of the sentence (see table IIb for /n/). For all test segments, the surrounding segmental context was kept constant: /a_a/ for /n, l, s, k/, /ɔ̃_ɔ̃/ for /t/, /k_d/ for /ɑ̃/ and /p_p/ for /i/.

Table II a, b: Example of sentences used for the test segment /n/. The test segment, in bold, is placed at four different prosodic positions: Intonational Phrase-intial (IPi), Accentual Phrase-initial (APi), Word-initial (Wi), and Syllable-initial (Si). In (a) are the sentences used in experiment 1 (EPG) where the serial position of the test segment is constant, in (b) are the sentence used for experiment 2 (nasal flow) where the serial position of /n/ varies.

	(a)	(b)
IPi	La pauvre Tata, **N**adia et Paul n'arriveront que demain. (Poor Auntie, Nadia and Paul will arrive only tomorrow.)	Pauvre Tata, **N**adia et Paul n'arriveront que demain. (Poor Auntie, Nadia and Paul will arrive only tomorrow.)
APi	Tonton, Tata, **N**adia et Paul arriveront demain. (Uncle, Auntie, Nadia and Paul will arrive tomorrow.)	Tonton, Tata, **N**adia et Paul arriveront demain ... (Uncle, Auntie, Nadia and Paul will arrive tomorrow...)
Wi	Paul et Tata **N**adia arriveront demain matin. (Paul and Auntie Nadia will arrive tomorrow morning)	Tonton Paul et Tata **N**adia arriveront demain par le train. (Uncle Paul and Auntie Nadia will arrive tomorrow by the train)
Si	Tonton et A**nn**abelle arriveront demain matin. (Uncle and Annabelle will arrive tomorrow morning)	Tonton Paul et Tata A**nn**abelle arriveront demain ... (Uncle Paul and Auntie Annabelle will arrive tomorrow...)

In several studies on the distinction between onsets and codas, the stimuli used were monosyllabic words (e.g. Byrd 1994). Therefore, the difference observed between the positions can not be attributed with certainty to an effect of position in the syllable or position in the word. Hence, it appears that in order to study the articulatory properties of segments at prosodic boundaries, one needs to control the hierarchical level of this boundary, i.e. the level of the constituents on both sides of the boundary. In this work, I follow the Strict Layer Hypothesis (Selkirk 1986), which states that constituents of a same hierarchical level can not be nested. It follows that a constituent of a level X^p can only be preceded by a constituent (if any) of the same level. Thus, a segment initial in a constituent is at the boundary between two constituents of the same hierarchical level.

In order to verify that the test segments were produced at the intended prosodic position, a transcription of the sentences produced by the speakers was done according to a set of prosodic criteria. These criteria define the prosodic level of the constituent preceding the test segment and, according to the Strict Layer Hypothesis, the level of the constituent in which the test segment is initial. Table III summarizes these criteria for the four prosodic positions studied. The criteria are: presence (+) or absence (Ø) of a pause before the test segment, degree of final lengthening of the preceding vowel (large: ++, medium: +, no lengthening: 0), demarcative function of the tone preceding the test segment (major boundary (%%), minor (%), not demarcative (Ø)). A handful of sentences that did not correspond to these criteria were eliminated – these were mostly sentences constructed for the APi position that had been produced with a pause before the test segment. The height of the tone preceding the test segment is also given in the table. In order to control for an effect of accent on the articulation of the test segments, it was verified that all test segments were not accented (all were realized with a low pitch contour).

Table III: Criteria used for the prosodic transcription of the prosodic position. See text for explanation.

	pause	lengthening V1	boundary tone
IPi	+	++	%% (L or H)
APi	Ø	+	% (H)
Wi	Ø	0	Ø (L or H)
Si	Ø	0	Ø (L)

2.2. Articulatory measures, subjects

Two types of articulatory data are examined in this study. In experiment 1, an index of lingual articulation was obtained with electropalatography (EPG), which measures linguopalatal contact, that is, the contact of the tongue against the palate. Data were collected with the Kay Elemetrics Palatometer 6300 in the Phonetics Laboratory at UCLA. Figure 1 shows pictures of the custom-made pseudo-palates. The white circles indicate the placement of the 96 electrodes that cover the hard palate and the inner surface of the upper molars. In order to capture the dental articulation of French /t/ and /n/, two electrodes have been placed in the middle of the inner surface of the incisors. Each sweep of the 96 electrodes takes 1.7 ms. and the sampling rate is 100 Hz. The audio signal was acquired in parallel at 12800 Hz.

Variation in the lingual articulation of the consonants and the close vowel /i/ depending on their prosodic position was measured in terms of amount of linguopalatal contact. This was done by computing the percentage of electrodes contacted in the frame showing the largest amount of contact, that is at the point of maximum constriction in the articulation of the segment. Most of the analyses describe the amount of contact over the whole palate (96 electrodes). Thus, one percent corresponds approximately to one electrode. For /s/ and /l/ some sub-regions of the palate have been defined to cover specific zones of articulation. These will be described in the appropriate sections.

Spk 1M **Spk 2F**

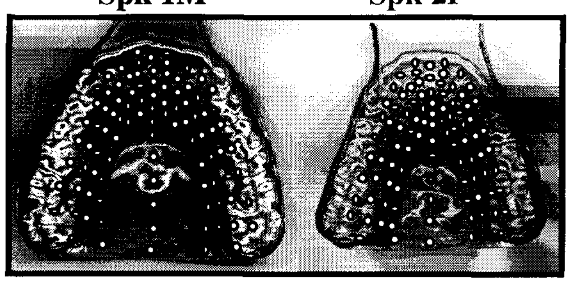


Figure 1: Picture of the pseudo-palates made for the two speakers. White circles represent the location of the 96 electrodes.

In experiment 2, an indirect measure of articulation at the velum was obtained with aerodynamic data. The amount of nasal flow during the production of a nasal segment (/n/ and /ɑ̃/) has been considered to be an approximation of velopharyngeal opening and of velum height. A rough control for the contribution of overall airflow to nasal flow was done by looking at the acoustic energy of the oral segments (vowels) surrounding the test segment. Even though the relationship between acoustic energy and flow is not straightforward, it was hypothesized that an increase in overall flow at particular prosodic positions could induce an increase in acoustic energy.

In that case, a concomitant increase in nasal airflow would have been due not only to a change in velopharyngeal aperture but also to a change in overall airflow. Results of this testing showed that acoustic energy of the vowels was similar in every prosodic position. Any change in nasal flow was thus interpreted as resulting from a change in velopharyngeal aperture. Nasal flow, collected through a Rothenberg split mask was filtered at 30 Hz and directly digitized, along with the audio signal, at a 1000 Hz sampling rate by a multi channel Kay Elemetrics CSL (model 4300B), and was then calibrated.

For both nasal segments, the amount of nasal flow in each prosodic position was taken at the point where nasal flow is maximum in the segment.

2.2 Other measures of articulatory variation

In order to approximate any change in glottal articulation depending on prosodic position, VOT was measured for the voiceless stops /t, k/. VOT was taken to be the time between the beginning of the burst and the onset of glottal vibrations on the acoustic waveform.

For the vowel /i/, variation in glottal articulation was measured in terms of the occurrence of glottalization in each prosodic position. Presence of glottalization at the onset or during the vowel was determined visually from the acoustic signal either by the presence of isolated glottal pulses before the onset of regular voicing or by irregular glottal vibration in the initial part of the vowel (Dilley & Shattuck-Hufnagel 1996).

Variation in the temporal properties of the initial segments was also measured. First, duration of the lingual closure of stop consonants was measured from the EPG data, as the time between the first frame showing a full lingual closure and the last frame before the lingual seal was broken. Second, acoustic duration of all segments was measured from the audio signal. In experiment 2 (nasal flow), even though the signal was filtered by the Rothenberg mask, its quality was sufficient for durational measurements. Segment boundaries were determined from both the waveform and a spectrogram. Consonant offset was determined by the onset of voicing and the appearance of F2 of the following vowel. Consonant onset was determined by the onset of frication noise on the spectrogram for /s/ and by the appearance of mid-frequency resonances and/or the onset of voicing for /l/ and /n/. For the stops, the closure duration (between onset of silence and burst) was measured only for positions where the stop was not preceded by a pause. For vowels, segment boundaries were determined both by the onset/offset of voicing in the waveform and the presence of higher formants on the spectrogram.

2.3 Speakers and procedure

Four Parisian French subjects participated in this study. Speaker 1M (male) and speaker 2F (female, the author), were recorded for experiment 1 (EPG) and experiment 2 (nasal flow). Two additional female speakers (3F and 4F) participated in experiment 2 (nasal flow). All were between 20 and 30 years old and had spent from one to five years in the US at the time of the study.

Data acquisition was made in the Phonetics Laboratory at UCLA in several recording sessions. The test sentences were presented to the subjects in blocks containing all four prosodic positions for each test segment (like in Table II). Subjects had to read each sentence 5-6 times and proceed to the next prosodic position. At the end of the block, they had to repeat the same sentences but in the reverse order (e.g. IPi , APi , Wi, Si then Si, Wi, APi, IPi) in order to avoid a list effect.

Speakers were not told how to phrase the test sentences. In a few cases, the subjects were just reminded to read them in a more lively style.

In general, 20 repetitions of each prosodic position were recorded except for /l/ (10 rep.) and /k, i/ (15 rep.). In the nasal flow experiment, some recordings out of the 20 had to be eliminated from the data set because of an unusual nasal flow or because the subject had breathed during the pause (in IPi position) and the nasal flow had not come back to normal at the start of the test segment. These correspond to about 15% of the data recorded.

For all the measures considered, the comparison between the prosodic positions was tested with a one factor (position) ANOVA and post-hoc Fisher PLSD tests with a 95% significance level. A prosodic position was considered significantly different if it was different from *all* the other positions. For more details on the method and the measures, the reader is referred to Fougeron (1998).

3. Results

3.1 Lingual articulation

3.1.1 Linguopalatal contact of stop consonants

Statistical analysis shows that the lingual articulation of the three stops /t, n, k/ varies depending on prosodic position. The amount of linguopalatal contact of initial stops tends to increase progressively from the lowest constituent (S) to the highest constituent (IP). Figure 2 illustrates this variation in the three stops for the two speakers. Table IV gives the statistical results of the comparison for each segment and speaker.

Table IV: Lingual variation: Statistical comparison done with a one factor (position) ANOVA and Fisher PLSD post-hoc test both with a 95% significance level. Direction of the significant differences in the amount of palatal contact is indicated by ">" for more contact and "<" for less contact. When a prosodic position does not appear in the comparison, that means that it is not significantly different from the surrounding positions (e.g. Wi for speaker 2F /t/ is not distinct from Si and APi). For /k/ and /i/ the position Si is not studied. In the 3rd and 5th columns is given the number of positions distinguished by linguopalatal contact out of the number of positions studied.

Segment	**Speaker 1M**		**Speaker 2F**	
/n/	F(3,73)=105.9, p=0.0001 Si < Wi < APi < IPi	4/4	F(3,76)=74.2, p=0.0001 Si < Wi < APi < IPi	4/4
/t/	F(3,77)=101.0, p=0.0001 Si, Wi < APi < IPi	3/4	F(3,78)=53.3, p=0.0001 Si < APi < IPi	3/4
/k/	F(2,42)=49.2, p=0.0001 Wi, APi < IPi	2/3	F(2,44)=25.9, p=0.0001 Wi < APi < IPi	3/3
/l/ front region	F(3,35)=23.4, p=0.0001 Si < Wi < APi, IPi	3/4	F(3,36)=23.4, p=0.0001 Si < Wi < APi, IPi	3/4
Asymmetry Index /l/	F(3,35)=37.7, p=0.0001 Si, Wi > APi > IPi	3/4	F(3,36)=7.7, p=0.0004 Si, Wi > APi, IPi	2/4
/s/ front region	F(3,75)=4.4, p=0.006 Wi < APi, IPi ; Si < APi	2-3/4	F(3,74)=9.4, p=0.0001 Wi, APi < IPi, Si	2/4
/i/	F(2,43)=7.4, p=0.002 Wi < APi, IPi	2/3	F(2,45)=22.5, p=0.0001 Wi < APi, IPi	2/3

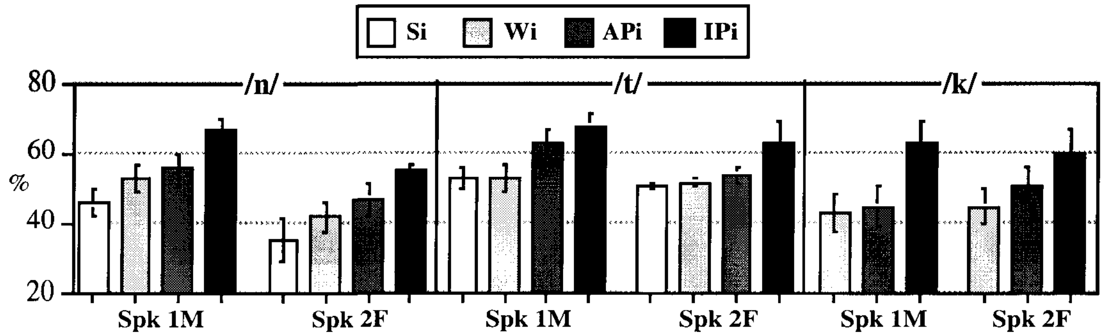


Figure 2: Amount of linguopalatal contact for the stops /n/, /t/, and /k/ in the different prosodic positions examined, Si: Syllable-initial, Wi: Word-initial, APi: Accentual Phrase-initial, IPi: Intonational Phrase-initial.

For the nasal stop /n/, the increase in linguopalatal contact significantly distinguishes the 4 prosodic positions considered for the 2 speakers, with a progressive increase in contact from the lowest position (Si) to the highest (IPi). For /t/, only 3 positions are distinguished by the progressive increase in contact. Speaker 1M does not show a distinction between the Syllable and Word levels, and for speaker 2F the Word level is not distinct from either the Syllable or AP level. For /k/, only Wi, APi and IPi positions were examined. While for speaker 2F, these 3 positions are distinguished by the amount of linguopalatal contact, speaker 1M produces only a two-way distinction between Wi/APi and IPi. The difference between Wi and APi follows the same trend with a greater amount of contact in APi but the difference is only marginally significant (p=.05).

A closer look at the distribution of the linguopalatal contact shows that the increase in contact observed in higher constituents results from a widening of the lingual constriction against the palate. For the two front consonants, /n/ and /t/, the surface of occlusion extends more toward the center of the palate when the consonant is initial in a high-level constituent. This is illustrated in the examples of contact profiles given in figure 3a. In these examples, the contact for /n/ in Si is maximal in the two front-most rows of electrodes and some electrodes are contacted on the third row. In IPi, the contact is maximal in the three front rows and extends to the fourth row. There is also an increase in electrodes contacted on the sides of the palate. This probably results from a better anchoring of the tongue against the palate in higher positions. For the back consonant /k/, the increase in contact in higher prosodic position appears as a widening of the back contact toward the middle of the palate, as illustrated in figure 3b.

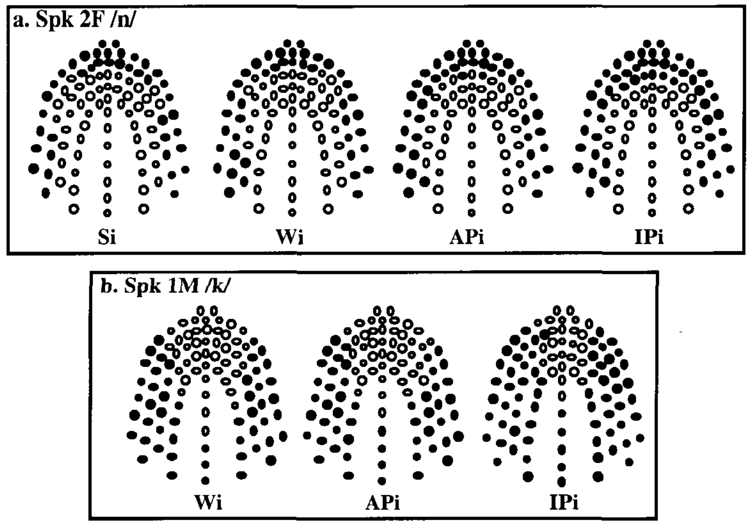


Figure 3 a, b: example of contact profile for /n/ and /k/ produced by speaker 2F. Black circles represent the electrodes contacted by the tongue. Each profile is taken at the frame of maximum contact in the rendition.

3.1.2. Linguopalatal contact of /l/

For the lateral consonant, three regions have been defined on the palate. These are illustrated in figure 4. The front-central region contains 46 electrodes. It extends from the front-most part of the palate toward the pre-palatal region and ends on the side after the pre-molars. The two lateral regions each contain 15 electrodes and cover the three most external lateral rows of electrodes.

Observation of the amount of contact over the front region shows that the lateral consonant follows a similar trend to that shown by the stops. The contact in the front region increases for initial /l/s in higher constituents. However, this increase in contact does not differentiate the same prosodic constituents as it does with the stops. For the two speakers, there is no distinction between the two highest levels: APi and IPi. Thus, as can be seen in table IV and in figure 4a, the two speakers make a three-way distinction between a sub-lexical level (Si), a lexical level (Wi) and a supra-lexical level (APi/IPi).

In the lateral regions there is also a change in the amount of contact depending on prosodic position. This change is dependent on the location of the lateral channel, which varies for the two speakers (on the right for 1M and on the left for 2F). In the region of the lateral channel, the contact tends to increase when the /l/ is initial in higher constituents. This means that the opening of the lateral channel becomes smaller in this position. On the opposite side, i.e. the side of tongue anchoring, the contact tends to decrease slightly. As a consequence, there is a reduction of the asymmetry of the contact between the two lateral regions in initial position in higher domains. This is shown by a reduction of the asymmetry index computed as [(left contact – right contact) / total right & left contact], in figure 4b and table IV.

In sum, for the 2 speakers, /l/ has a smaller asymmetry of tongue contact in initial position of higher constituents along with a widening of front central contact.

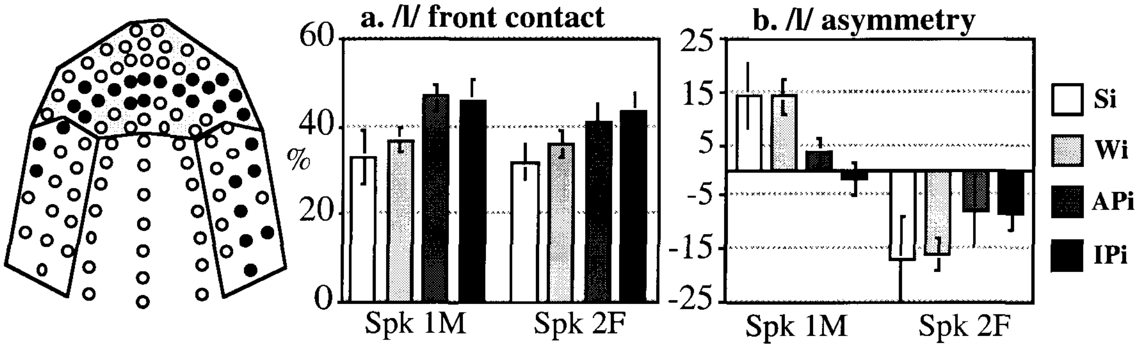


Figure 4a, b: Amount of linguopalatal contact for /l/ in the front region depending on prosodic position (a). Asymmetry in the lateral contact between the left and right regions depending on prosodic position (b). Asymmetry to the left appears as a positive index and asymmetry to the right as a negative index.

3.1.3. Linguopalatal contact of /s/

The articulation of the fricative /s/ is examined in the front region of the palate where the fricative groove is located. This region was defined as in figure 5. It includes 33 electrodes, extends towards the alveolar region, and ends on the side at the pre-molars. Variation in the amount of contact in this region is illustrated in figure 5.

It can be seen that /s/ is less systematically affected by prosodic position compared to the other consonants studied. For the 2 speakers, some variations can be seen but the tendency is less clear than in the case of other consonants. Excluding the Si position, there is an increase in contact from Wi to APi/IPi for speaker 1M, and from Wi/APi to IPi for speaker 2F. Thus, even if the variation observed follows the same tendency as that of other consonants, i.e. an increase in contact in higher position, very few positions are distinguished this way for this consonant. Moreover, the lowest position Si does not follow this trend since a large amount of contact is observed there.

Closer examination of the characteristics of the fricative groove did not reveal any more effect of prosodic position on /s/. The width of the fricative groove, its area (length*width) and the groove center location were measured following Fletcher (1989)'s indices, but no clear trend was found.

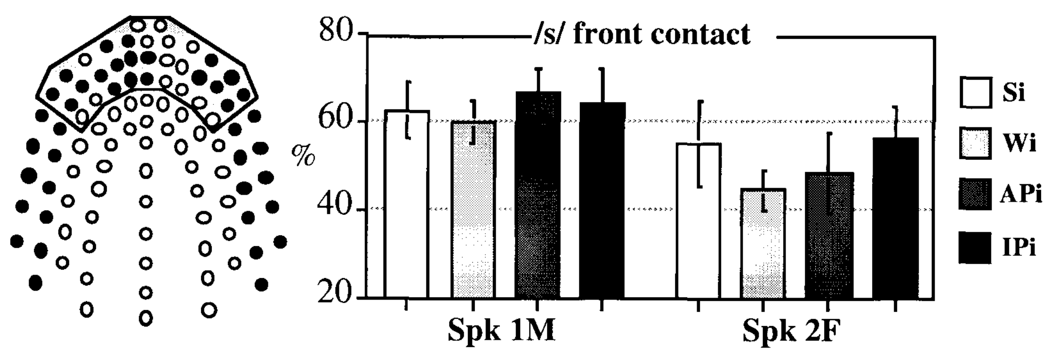


Figure 5: Amount of linguopalatal contact for /s/ in the anterior region depending on prosodic position.

3.1.4. Linguopalatal contact of the vowel /i/

The amount of linguopalatal contact during the articulation of the close vowel /i/ was measured at the point of maximal contact, as done for the consonants. Only three positions are compared: Wi, APi, IPi.

The left panel of figure 6 shows the amount of linguopalatal contact over the whole palate for the vowel. As observed in the consonants, there is an increase in contact in the vowel that follows the hierarchy of the positions. However, significant differences are only found between the lexical level (Wi) and the two phrasal levels (APi and IPi) for the two speakers (see table IV).

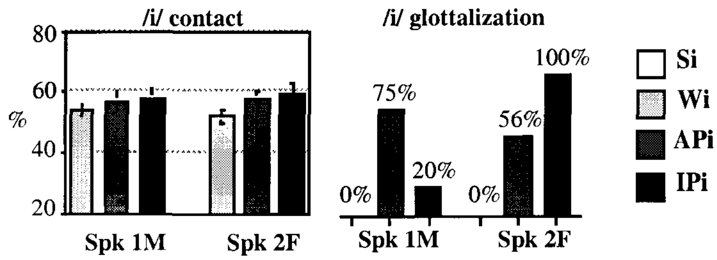


Figure 6: On the left, amount of linguopalatal contact in /i/ and on the right, percent occurrence of glottalization in the vowel depending on prosodic position

3.2. *Nasal flow for nasals*

The amount of nasal flow depending on prosodic position was measured for the nasal consonant /n/ and the nasal vowel /ɑ̃/. For the nasal vowel, only three prosodic positions are studied: Wi, APi, IPi.

For both nasal segments, there is a tendency for nasal flow to decrease in initial position in higher constituents. However, the progressive nature of this trend is less striking than the increase observed for linguopalatal contact. As can be seen in table V and figure 7a, the common pattern shown for three out of the four subjects is a diminution of nasal flow for initial /n/ in IP compared to initial /n/ in lower constituents. But for speaker 4F, the trend is the opposite. This speaker, who was the least lively in her productions, shows an increase in nasal flow in IPi position.

For /ɑ̃/, there is variation in nasal flow depending on prosodic position for two (1M, 3F) out of the four speakers and no variation for the others (see table V and figure 7b). For these two speakers, the trend is similar to the one shown for /n/: a diminution of nasal flow in higher positions. Speaker 3F presents a gradual decrease of nasal flow between the three positions studied (Wi > APi > IPi). Speaker 1M makes only a two-way distinction between the lexical level (Wi) and the phrasal levels (APi/IPi). Therefore, only the diminution from Wi to IPi is shared by these speakers.

Table V: Statistical comparison done with a one factor (position) ANOVA and Fisher PLSD post-hoc test both with a 95% significance level. Direction of the difference in the amount of nasal flow is indicated by ">" for more flow and "<" for less flow. For /ɑ̃/, the position Si was not studied. In columns 3, 5, 7, 9 is given the number of positions distinguished by linguopalatal contact out of the number of positions studied.

	Speaker 1M		Speaker 2F		Speaker 3F		Speaker 4F	
/n/	F(3,54)=5.8, p=0.002 Si, Wi, APi > IPi	2/4	F(3,70)=5.2, p=0.003 Si, Wi, APi > IPi	2/4	F(3,68)=8.3, p<0.001 Wi, APi > IPi ; Si < Wi	2-3/4	F(3,59)=6.6, p<0.001 Si, Wi, APi > IPi	2/4
/ɑ̃/	F(2,24)=9.4, p=0.001 Wi > APi, IPi	2/3	F(2,27)=0.3, p=0.8 ns.	0/3	F(2,25)=21.0, p<0.001 Wi > APi > IPi	3/3	F(2,27)=0.9, p=0.4 ns.	0/4

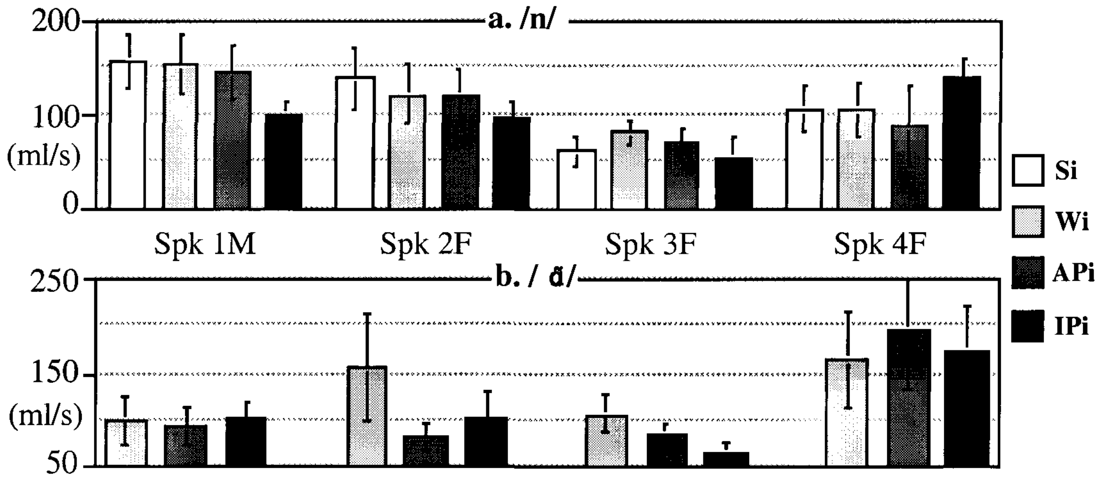


Figure 7a, b. Amount of nasal flow at the point of maximum flow in /n/ (a) and in /ɑ̃/

3.3. Glottal articulation

3.3.1. VOT of /t/ and /k/

The VOT of /t/ shows little variation depending on the prosodic position of the consonant and the only common variation shown by the two speakers is that VOT is shorter in Syllable-initial position. For /k/ no difference is found in VOT duration depending on position. The comparisons of VOT are given in the first two rows of table VI.

3.3.2. Occurrence of glottalization in /i/

The percent occurrence of glottalization observed for the vowel /i/ in the different prosodic positions is illustrated in the right panel of figure 6. Results show that, in this corpus, /i/ is never glottalized in Word-initial position, while it is more frequently glottalized in initial position of higher phrasal constituents. Depending on the speakers, the frequency of glottalization increases progressively from Wi to APi to IPi, as for speaker 2F, or is more frequent in AP-initial position, as for speaker 1M.

3.4. Durational variation

3.4.1. Duration of articulatory events: lingual closure duration for stops

Comparison between prosodic positions (table VI) shows that the duration of the lingual closure (as captured in EPG data) varies in a similar way as the amount of linguopalatal contact: consonants initial in higher constituents have a larger amount of contact and also a longer lingual closure. However, fewer distinctions are made by the temporal measure (closure duration) than by the spatial measure (linguopalatal contact): speaker 2F loses the distinction between Si and Wi for /n/ and between Wi and APi for /k/; speaker 1M loses the distinction between Wi and both Si and APi for /n/.

In order to examine the relationship between the temporal and spatial properties of the lingual articulation, correlations were calculated between these two measures. These coefficients are given in table VII. In (a) the correlations between the 2 factors are computed for all positions. It can be seen that there is a relationship between the spatial and the temporal measures, with the duration measure accounting for 20% to 76% of the variance of the linguopalatal measure. The duration of the occlusion is usually very long in IP-initial position. In /n/, for example, the occlusion is initiated, and even achieved, well before the voicing starts, that is during the preceding pause. In order to avoid exaggerating the relationship between linguopalatal contact and closure duration by the characteristics of IPi position (large amount of contact and extra-long closure), the correlation was also computed only for the three positions not preceded by a pause: Si, Wi and APi. These results are shown in (b) in table VII. As expected, the exclusion of IP-initial position somewhat the correlation between the two variables reduces (especially for speaker 1M), so that the variance of closure duration can not account for more than 66% of the variance of linguopalatal contact (and vice-versa).

Table VI: Durational variation: significant distinction between prosodic position determined by a one factor (position) ANOVA and Fisher PLSD post-hoc test both with a 95% significance level. Direction of the duration difference is indicated by ">" for longer and "<" for shorter.

Segment	Speaker 1M	Speaker 2F	Speaker 3F	Speaker 4F
VOT duration				
/t/	Si, Wi < APi, IPi	Si < Wi, IPi		
/k/	ns	ns		
Lingual seal duration (correlation only for positions not preceded by a pause)				
/n/	Si < APi < IPi	Si, Wi < APi < IPi		
/t/	Si, Wi < APi < IPi	Si < Wi, APi < IPi		
/k/	Wi, APi < IPi	Wi, APi < IPi		
Acoustic duration (comparison and correlation only for positions not preceded by a pause)				
/n/ exp.1	Si < Wi < APi	Si < Wi < APi		
/n/ exp.2	Si < Wi < APi	Si, Wi < APi	Si < Wi < APi	Si < Wi < APi
/t/ closure	Si < Wi < APi	Si < Wi < APi		
/k/ closure	Wi < APi	Wi < APi		
/l/	Si, Wi < APi	Si, Wi < APi		
/s/	Si, Wi < APi	Si < Wi < APi		
/i/	Wi < APi	ns.		
/ã/	ns.	Wi > APi	ns.	ns.

Table VII: Correlation coefficients between segment durational properties and amount of linguopalatal contact. In (a) are given the results of the computation including all of the prosodic positions (i.e. including IPi), in (b) are given the results for all of the positions except IPi (i.e. all the positions not preceded by a pause).

		Closure duration			Acoustic duration					
		/n/ seal	/t/ seal	/k/ seal	/n/	/t/ closure	/k/ closure	/l/	/s/	/i/
(a)	Spk 1M	0.76	0.60	0.20	0.02			0.20	0.002	0.09
	Spk 2F	0.63	0.60	0.60	0.11			0.18	0.11	0.03
(b)	Spk 1M	0.63	0.47	0.0003	0.68	0.51	0	0.70	0.04	0.15
	Spk 2F	0.66	0.61	0.0001	0.69	0.35	0.09	0.67	0.13	0.09

3.4.2. Acoustic duration

For all test segments, Intonational Phrase-initial position is characterized by a large variability in acoustic duration, as shown by the large standard deviation in figure 8. While the segments can be quite long in this position, they are most often shorter than in other positions. In IP-initial position, segments are preceded by a pause. It is known that the setting of appropriate aerodynamic conditions required for voicing to take place may take some time after a pause, so that the acoustic duration of the segment may be shortened (e.g. Lisker & Abramson 1967, Flege 1982). Because of this large variability in IPi position, the discussion about segment duration will focus on the positions not preceded by a pause: Si, Wi, APi.

The last rows in Table VI summarize the differences observed in segment duration in these three positions. It can be seen in this table that the pattern shown by all consonants is a lengthening in APi position compared to initial position in lower constituents. In several cases there is also a lengthening in Wi position compared to Si, thus a progressive lengthening from Si to Wi to APi.

The vowels exhibit variation in acoustic duration depending on prosodic position only for one speaker. Speaker 1M shows a significant lengthening in APi compared to Wi (and also IPi) for /i/ and a lengthening of /ɑ̃/ in Wi compared to APi (and IPi). However, it must be noted that initial vowels have been shown to be glottalized in higher constituents, and some of the glottalized portion may not have been included in the duration measurement.

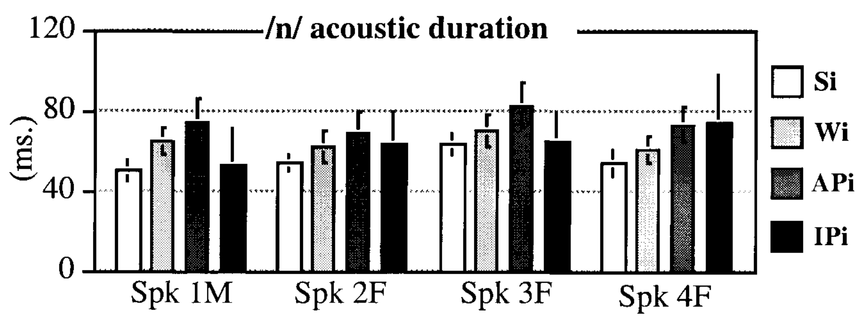


Figure 8: Acoustic duration of /n/ measured in experiment 2 (nasal flow) depending on prosodic position.

Results of the correlation between the acoustic duration and the amount of linguopalatal contact of segments in experiment 1 (EPG) are given in table VII. When the computation includes all of the positions (in (a)), it can be seen that the temporal and articulatory variables are quite independent from each other. On the contrary, when we look at the positions not preceded by a pause (in (b)), a positive correlation between the two variables appears clearly for /n/ and /l/, and to a lesser extent for /t/. For these segments, the difference between the two types of computation (with or without IPi) results from the fact that in IP-initial position, these segments have a large amount of linguopalatal contact but a short acoustic duration. In the other positions, the increase in linguopalatal contact from lower to higher position is accompanied by a lengthening of these segments.

4. Discussion

As previously shown for the word and syllable levels, this study demonstrates that segments placed in initial position in higher level constituents are characterized by specific articulatory properties in French. These properties are found to differentiate segments initial in higher domains from segments initial in lower domains as found in other languages. At the same time, the comparison between several domains showed that initial segments in the constituents studied are different from medial segments, since a segment initial in a lower constituent (e.g. a word) is also medial in a higher constituent (e.g. an AP).

Thus, these results confirmed for French that the articulation of speech segments can be affected by the prosodic organization of the utterances in which they are produced. For each of the articulatory properties considered, the magnitude of the variation observed between segments in initial position tends to follow the prosodic level of the constituents. For lingual articulation, for example, the magnitude of the increase in linguopalatal contact tends to be progressive from the lowest to the highest constituent studied. Hence, these articulatory properties in initial position do reflect the hierarchical organization of the prosodic constituents (i.e. of boundary strength).

4.1. Variability in the realization of prosodically-driven articulatory properties

This study also showed that there is a large amount of variability in the phenomenon observed. The realization of these prosodically-driven articulatory properties varies depending on the segment type (a), the articulator (b), the speaker (c) and the constituent (d) considered.

(a) As far as segment type is concerned, it is found that prosodic position clearly affects the articulation of the stops and the lateral, and to a smaller extent that of the vowels. Conversely, the lingual articulation of the fricative seems to be less affected by prosodic position. This finding is not surprising. In Byrd's comparison between onset and coda in English (1994), for example, /s/ also showed fewer positional differences. In fact, this consonant is known to be less subject to articulatory variation in general because of its small articulatory and acoustic degree of freedom (e.g. Shadle & Scully 1995). In our study, it could also be the case that the variations taking place in initial position in higher domains consist of an increase in the upward force with which the tongue contacts the palate. This would affect the cross-sectional shape of the tongue (Stone, Faber, Raphael & Shawker 1992:266), but this modification could not be inferred from EPG alone since linguopalatal contact does not reflect the exact shape of the tongue.

(b) In experiment 2, it is found that the amount of nasal flow distinguishes at most 3 prosodic positions and in most cases, only a two-way distinction is made between IPi position and lower positions. Hence, variations in nasal flow show fewer distinctions than linguopalatal contact.

This observation is hard to interpret since it could be due either to the type of articulator or to the type of articulatory data. Indeed, it is possible that velum height, especially for nasals, has a smaller degree of freedom than tongue position. It is also possible that aerodynamic data, which are an indirect measure of velopharyngeal aperture, do not reflect subtle differences in aperture between prosodic positions with enough precision.

(c) Another factor of variation is the speaker. Although all the speakers studied show some articulatory variations depending on prosodic position, they do not always realize the same distinctions. In the EPG experiment, it can be seen in table IV that there are only three cases where the two speakers share exactly the same distinctions (for /n, l, i/). In the nasal flow experiment, there is even more variability: three out of the four speakers show a reduction of nasal flow in /n/ in IPi position while the other speaker shows an increase of nasal flow in that position. This difference remains to be explained. For /ã/, while two of the speakers show some variation in nasal flow depending on prosodic position, the other two speakers do not. As was noted in the introduction, comparable speaker variability was also observed in other related studies (see table I). In their study of vowel glottalization in English, Dilley and Shattuck-Hufnagel (1996:442) suggested that "[...] glottalization is one of a collection of strategies that speakers use to mark prosodic events, which together are rule-governed but which may individually vary with importance across speakers". A similar conclusion can be applied to the variations in supra-glottal articulation reported here.

(d) As noted above, all the speakers or all the segments do not show the same distinctions between prosodic positions. In other words, it is not the case that each constituent studied is consistently marked by variation in the articulation of its initial segment. As indicated in table IV and V, articulatory variations observed in initial positions differentiate at least 2 levels of constituents (for /s, i/ linguopalatal contact and for /n/ nasal flow) and at the most 4 (out of 4) constituents (for /n/ linguopalatal contact). In general, the most robust distinctions observed are the ones made between the most extreme constituents of the hierarchy: the highest constituent IP and the lowest constituents S or W. Similar variability in the number and the nature of the constituents differentiated by articulatory variations was observed in other languages (e.g. Keating, Cho, Fougeron & Hsu 1998). Lehiste (1964) also noted that positional allophones are not always produced in words and that their realization is more frequent in a "maximally differentiated style" of speech. This phenomenon need therefore to be studied in a more casual style of speech and in spontaneous speech.

Interestingly, the Word level does not appear to be consistently different from either the Syllable or the Accentual Phrase level in our data. This fact has to be interpreted with caution, since the definition of "word" in our corpus is lexical. It is possible that what I considered as a Wi position in a sequence like "Tata Nadia" does not correspond to an initial position in a prosodic word (which remains to be defined for French). Also the interaction with the presence or absence of an initial accent on the word preceding the test word (on the first /ta/ of "Tata Nadia") was not controlled. This needs to be explored further. Moreover, words in French are often preceded by a determiner. In my corpus, the use of first names was meant to avoid having a determiner so that the word would be placed at the left edge of the constituents. Future studies will have to consider whether articulatory variation in initial position is purely an edge-effect. In that case, the articulation of the determiners should vary depending on the prosodic level of the constituents in which they are initial.

In sum, it appears that the realization of the prosodically-conditioned articulatory properties observed is *optional*. However, this fact does not reduce the interest of the phenomenon. As noted by Danes in his comments of Lehiste's results (1964), what is interesting is the "potentiality" of these articulatory properties to occur: in some positions, but not in others, segments can be articulated in a different way. Then, it is the possibility vs. the impossibility of occurrence of these properties that makes the difference between the prosodic positions. This suggests that the initial position in a constituent is a particular site where these properties can be realized. Nonetheless, the optionality of the phenomenon leads to the fundamental question of whether these articulatory properties are intentionally produced by the speakers or whether they are side effects of other prosodic variations.

4.2. Relationship with other prosodic variations

The signaling of prosodic boundaries is multi-parametric. As noted earlier, particular intonational contour, the occurrence of pauses, or presence of final lengthening contribute to setting off constituents' boundaries. These markers were used here as criteria for defining the prosodic positions studied. It is therefore possible that the articulatory properties found in initial position of these constituents are only secondary effects of these other prosodic variations. The following arguments will show that it is not the case.

Articulatory variations in initial position do not seem to be conditioned by the occurrence of a pause, since they also occur in constituents not preceded by a pause (Si, Wi, APi). For nasal flow, the distinction between IPi position and the other positions may be influenced by the presence of a pause before the test segment that could induce a change in aerodynamic conditions. However, pausing can not explain the few variations observed in positions that do not involve a pause (e.g. Wi vs. APi for /A)/).

The articulatory properties of initial segments do not appear to be a secondary effect of the type of intonation contour either. Recall that all the test segments are unaccented and produced with a low tone. The nature and the height of the tone preceding the test segment are variable since they are one of the criteria used to define prosodic positions. However, it is not possible to associate any articulatory variation to a specific intonational property: in APi and IPi segments show articulatory distinctions although they are both preceded by a final demarcative accent (realized usually as a high tone and sometimes as a low boundary tone); in Wi and Si, segments also show articulatory distinctions although they are both preceded by a low f0 contour without a demarcative accent.

Final lengthening is also a criterion used for determining prosodic position. Consequently, articulatory variations do follow the pattern shown by final lengthening. For example, there is more linguopalatal contact in IPi where the test segment is preceded by a large final lengthening. However, in the cases of Si and Wi, it is not always the case that the vowel preceding a Word-initial segment is more lengthened that the vowel preceding a Syllable-initial segment. Nevertheless, we observed some articulatory variation between Si and Wi.

Lengthening of initial segments is also a factor found to vary with prosodic position in French. Initial lengthening was also reported at the word level (among other Oller 1973, Vaissière 1983) and for higher constituents (see Keating et al. 1998) in several languages. Then, it could be the case that the spatial modification of the articulation of initial segments is a consequence of their lengthening. As suggested in Fougeron and Keating (1997), following Lindblom's (1963) reasoning, it is possible that lengthened initial segments have more linguopalatal contact in higher constituents because the raising movement of the tongue has enough time to reach its target position, and even to overshoot it. However, the correlation between the two variables showed that the change in articulation in initial position is only partly linked to lengthening in French: no more than 70% of the variation in linguopalatal contact can be explained by duration. Furthermore, for most of the segments, variations in linguopalatal contact distinguish a larger number of prosodic positions than the variation in initial lengthening does. In fact, initial lengthening appears to be mostly a characteristic of Accentual Phrase initial position in French. In English, Fougeron and Keating (1997) also found only a weak correlation between the amount of initial lengthening and the amount of linguopalatal contact for /n/. In Tamil, Byrd and colleagues (Byrd, Kaun, Narayanan & Saltzman 1996), found a lengthening of the lingual occlusion of initial /n/ in higher constituents, but this lengthening was not associated with any change in the magnitude of the lingual gesture. Conversely, Cho and Keating (1999, this volume) found in Korean that the variations in linguopalatal contact show a larger positive correlation with initial lengthening. The relationship between the two factors seems then to be language-specific.

To conclude, the articulatory properties observed in initial position in different constituents in French seem to be governed by a mechanism independent from the prosodic markers reviewed above. The cumulative nature of the articulatory variations observed suggests that these variations are associated with the prosodic phrasing of utterances: the articulatory properties found in initial segments reflect the hierarchical organization of the constituents. In that sense, *prosodic phrasing is realized in the articulation of the segments*. However, it is difficult at this point to conclude that these articulatory properties are perceptible prosodic markers or that they can be used by listeners for prosodic segmentation. Acoustic correlates of these articulatory properties first have to be found. In this study, we saw that initial lengthening often but not always accompanied these properties. An analysis of the acoustic energy of /n/ and /l/ depending on prosodic position (not reported here) has shown a trend for the energy distinction between the initial consonant and the following vowel to increase in IPi position (see Fougeron 1998). However, this characteristic is not shared by all the speakers or by all the consonants, and it cannot distinguish other prosodic positions. Another interesting finding in Fougeron (1998) was that the spectral characteristics of initial /i/ are affected by the height of the constituents. From Wi to APi to IPi there is a significant progressive increase in F3. Since /i/ is surrounded by labial consonant in the corpus, this increase in F3 could be interpreted as a reduction of the contextual labialization of the vowel in initial position in a higher constituent. This would suggest that initial segments are more resistant to coarticulation in higher constituents, but this hypothesis has to be investigated further.

Whatever the linguistic function of these articulatory properties, what seems clear from this study and others is that the variations in the articulation of initial segments appear as traces of the prosodic encoding of an utterance. These variations are not produced randomly; they may be optional but are rule-governed: (1) they are cumulative and follow the prosodic hierarchy, and (2) they consist of subtle modifications of the articulation of the segment, a trend which is shared by all the segments studied. Hence, the puzzling question is now to understand the nature of the physiological mechanism responsible for the occurrence of these articulatory variations.

4.3. On the nature of prosodically-conditioned articulatory variations

4.2. Comparison with articulatory variations found in other languages

Lehiste (1964:196) noted that "the manner in which boundaries are realized in a language constitutes an integral part of its structure, and has to be included in its phonological description". From the data reported so far on articulation at constituent boundaries, it is difficult to draw a comparison between languages since most of the studies have compared only a few types of segments or articulations, and the prosodic constituents studied are specific to each language. Nonetheless, what appears clearly is that the production of specific articulatory properties at initial boundaries is shared by several languages and that the nature of the variations is similar in these languages.

As shown in figure 9, modifications in the articulation of initial segments found in French are comparable to the ones described in other studies, both in comparisons between initial and non-initial segments at the word or syllable levels, and in comparisons between initial positions in several prosodic constituents:

(1) For lingual articulation, initial segments in French show an increase in linguopalatal contact and a widening of the constriction area in higher constituents. These properties were also found to characterize syllable/word-initial consonants in French by Rousselot (1901) and Straka (1963), and by others in several languages. This result also replicates the properties observed for initial stops in higher domains in English, in Korean and in Taiwanese (see table I in introduction). The present study adds to these results by showing that the initial close vowel /i/ presents the same type of variation as consonants.

(2) For nasal articulation, initial nasal segments in French show some decrease of nasal flow. A similar observation was made by Gordon (1996) for Estonian /n/ in initial position of higher prosodic domains. Fujimura (1977) and Krakow (1989) observed that velum position is higher in word initial position compared to medial or final position for nasal consonants in English and Japanese. If we extend these observations made at the word level to our results, it is possible to interpret the observed decrease in nasal flow as a raising of the velum, which induces a reduction of velopharyngeal aperture in higher position. Again, this study shows that nasal consonants and nasal vowels are affected by prosodic position in the same way (when they show variation).

(3) For glottal articulation, initial /i/ is more frequently glottalized in higher constituents than in Word-initial position in French. In other languages, vowel glottalization has been shown to characterize initial as opposed to medial or final vowels (e.g. Lehiste 1964, Kohler 1994). In French, processes like liaison or enchaînement reduces the likelihood of a Word-initial vowel's being glottalized. In the corpus used here, the test vowel was produced in the sequence "Philippe Ippine". It is probable that there was enchaînement between these two words, although this was not tested empirically. That initial vowels are more frequently glottalized in higher prosodic constituents in French corresponds to findings for English by Pierrehumbert and Talkin (1992) and Dilley and Shattuck-Huffnagel (1996): initial vowels are more frequently glottalized in IP-initial position compared to Word-initial position, and compared to Intermediate Phrase-initial position if the vowels are unaccented (Dilley & Shattuck-Huffnagel 1996). As far as VOT is concerned, in Korean (Jun 1993, Cho & Keating 1999, this volume), in Taiwanese (Hsu & Jun 1997) and in English (Pierrehumbert & Talkin 1992) it has been found that the VOT of an aspirated voiceless consonant varies depending on its prosodic position. In Korean for example, Jun (1993) showed a progressive lengthening of VOT of aspirated stop from Si to Wi to APi. In French, whose stops are unaspirated, this effect does not appear so clearly.

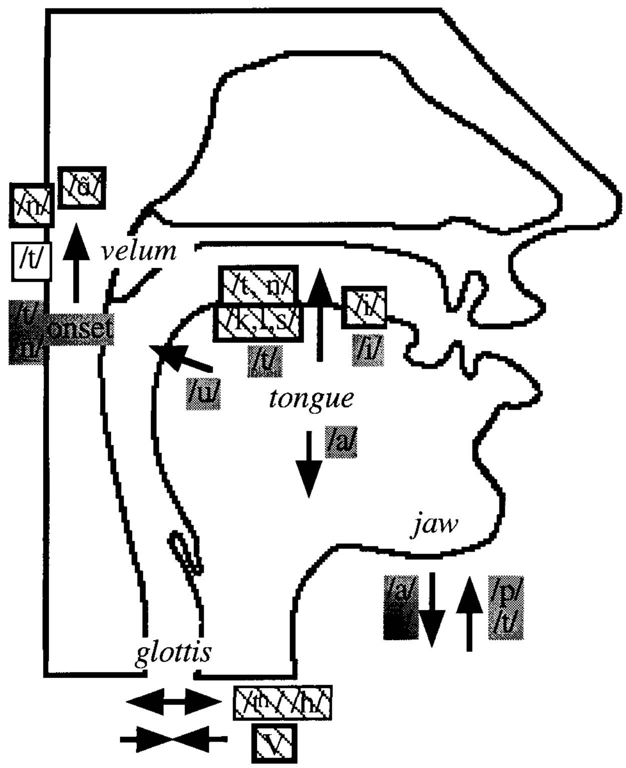


Figure 9: Schematic representation of the articulatory properties found to characterize initial segments. The direction of the arrows indicates the nature of the articulatory property for each articulator concerned. For example, the rising arrow for the tongue means that initial segments are characterized by a higher tongue position. White boxes indicate the properties found to characterize segments in initial position at the word or syllable level (compared to medial or final positions). Hatched boxes indicate the segments for which these articulatory properties have been found to increase in magnitude for segments initial in higher prosodic constituents. Among these, the ones in bold boxes have been confirmed for French in this study. For comparison, gray boxes show the articulatory properties reported to characterize segments in accented syllables.

4.2. Comparison with articulatory properties found in accented position

As noted by Beckman and Edwards (1994:8) the phrasing organization of an utterance "can be indicated by marking the edges or by marking the heads". It is therefore possible that articulatory variations in these two prosodically important positions (boundary position and accented position) are governed by the same mechanism. Indeed, the nature of the variations found in initial position of prosodic constituents is comparable to that observed in accented position as illustrated in figure 9:

- For most of the articulators, modification in the articulation of segments under accent have been reported for vowels. For lingual articulation, several studies have shown that tongue position is more peripheral for vowels in accented compared to unaccented position. For example, while /a/ is found to be more open in accented position, the tongue position for /i/ is higher and sometimes more front in this position (Rousselot 1901, Giot 1977 in French, Kent & Netsell 1971 in English, Farnetani & Vayra 1996 in Italian). This modification in the articulation of /i/ is thus similar to that found in initial position in higher constituents: /i/ has more linguopalatal contact, reflecting a higher and/or more advanced tongue position.

- For consonants, Farnetani and Vayra (1996) showed in Italian that /t/ has a greater amount of linguopalatal contact in accented syllables. In French, lingual closure is also longer under accent (Simon 1967). Vaissière (1988) showed that the velum is higher for nasal and oral consonants in accented syllables, but only if those consonants are in onset position. In coda position, accented consonants have a lower velum than unaccented consonants. Nasal and lingual articulation thus show somewhat similar modifications for consonants in accented position as in initial position.

Two main hypotheses have been advanced to explain articulatory variations conditioned by accent. Based on the observation of the jaw, Beckman and colleagues (Edwards and Beckman 1988, Beckman, Edwards and Fletcher 1992) proposed that the articulatory variations in accented position contribute to increasing the sonority distinction within a syllable: in accented syllables, consonants with a higher jaw position are less sonorous while vowels with a lower jaw position are more sonorous. Straka (1964) made a similar hypothesis stating that it is the aperture distinction within the syllable that is increased under accent as well as in initial position. In these two hypotheses, it is then the syntagmatic contrast in the syllable that is enhanced. However, de Jong (1995) has argued that the articulation of accent does not imply a sonority expansion mechanism. He showed that in English emphatic stress modifies not only the position of the jaw but also that of other articulators. Moreover, for some segments, the articulatory modifications go against a sonority expansion: for non-open vowels, for example, the elevation of the tongue in accented position reduced their sonority. Consequently, de Jong proposed that the articulatory variations found in accented syllables result from a local hyper-articulation. Hence, the distinction between accented and unaccented position is made on the paradigmatic scale by an enhancement of the distinctive features of accented segments.

Considering the articulatory variation observed in domain-initial position, it appears that neither of these two hypotheses can explain all the results. The increase in linguopalatal contact for initial /i/ in higher constituents does not result in an increase in the vowel's sonority or aperture. However, measurement of the acoustic energy of sonorants (not presented here) showed that the articulatory modifications observed do not involve a systematic increase in acoustic distinction between the consonant and the following tautosyllabic vowel (see Fougeron 1998). The hypothesis of an enhancement of distinctive features is also challenged by the variation observed for nasal segments. Initial nasal consonants and vowels show a decrease in nasal air flow in initial position of higher constituents. Neither their nasal feature nor their sonorant feature is therefore enhanced.

In fact, it appears that articulatory variations observed in initial position may contribute to enhance syntagmatic as well as paradigmatic contrast (see for example the lingual articulation of the consonants), but it seems that the underlying process governing these variations has to be found in a more physiological mechanism.

4.3. A local and progressive articulatory strengthening in domain initial position

Initial and accented positions have been traditionally described as "strong positions" in order to reflect the particular resistance of the segments in these positions to diachronic evolution. The notion of strength has also been used to describe the articulatory properties found in the segments placed in these two positions. For example, Vaissière (1986, 1988) proposed to associate a [+strong] feature to the Word-initial and accented positions. She suggested that this feature is realized as an increased tension of the overall production mechanism. For Fujimura (1990:233), initial positions in Syllables, Words and Phrases are characterized by more "forceful" articulatory gestures along with a reinforced source intensity. For Fougeron and Keating (1997), articulatory variations in initial positions are described as an "initial strengthening", which mirrors the "final weakening" often described in the literature.

The notions of "strength", "strengthening", "articulatory force" have also been widely used in the literature to characterize the articulation of particular segments: strong consonants along a strength scale and fortis (vs. lenis) consonants (e.g. Delattre 1940, Malmberg 1943, Sraka 1964 for French). Straka (1963) also used this notion to characterize a particular style of speech, "la parole renforcée" (strengthened speech), which he described as involving a more "energetic or more forceful" articulation. Interestingly, the segmental articulatory properties, allegedly explained by an increase in articulatory force, present striking similarities with the properties found to vary in initial position in prosodic constituents. For example, an increase in linguopalatal contact and a widening of the occlusion surface was found to distinguish strong (voiceless) consonants compared to weak (voiced and nasal) consonants (e.g. Rousselot 1901, Simon 1967, Marchal 1979 for French). Similar properties were found for /t, k, s, l/ in "parole renforcée" compared to normal pronunciation by Straka (1963). Strong voiceless stops have also been shown to have a closer tightening of the velum against the velopharyngeal wall compared to voiced stops (Simon 1967). Moreover, Straka found that in "parole renforcée" velum position is higher for both oral and nasal segments.

However, the notion of "articulatory force" (or "articulatory strengthening") has been strongly debated in the literature, mostly because of its intuitive use, its lack of physiological definition and the difficulties one has to quantify it (see e.g. Debrock 1977). Nonetheless, if one considers the definition given by Straka (1963), it seems that articulatory strengthening could be the mechanism involved in initial position. For him, articulatory force is "rien d'autre que la force de contraction des muscles entrant en action pour l'articulation donnée" (nothing but the contraction strength of the muscles involved in a given articulation, p. 91). What must be underlined in this definition is that it is only the muscles involved in the positioning of the articulators for the given articulation that are affected by this increased contraction. For Straka, articulatory strengthening is thus an increase in the "force of the articulatory movements" but not an increase in the overall "muscular effort" (including the phonatory and expiratory forces) involved in articulation.

With this definition, articulatory strengthening seems appropriate for explaining the modification observed in the lingual, nasal and glottal articulation of segments in initial positions:

(1) the larger amount of linguopalatal contact and the widening of the occlusion surface reflects a greater elevation of the tongue in initial position (and maybe some backing/fronting depending on the consonant's place of articulation). For consonants, this could result from an increased contraction of the elevator muscles of the tongue (genioglossus, palatoglossus and styloglossus). For the close vowel /i/, it depends on which muscles are considered to be mainly involved in its production. Straka considered that all vowel articulations were essentially governed by the activity of the lowering muscles. However, it has been empirically shown since then that elevator muscles (particularly the posterior genioglossus) are mainly responsible for the protrusion of the tongue to the front of the mouth (e.g.. Honda, Hirai & Kusakawa 1995, Payan & Perrier 1996). An increased contraction of this muscle would therefore increase the tongue protrusion, which in turns could result in an increase in linguopalatal contact.

(2) the decrease in nasal flow observed for nasals in initial position of higher constituents in this study, as well as the higher position of the velum of word-initial oral and nasal consonants observed in the literature, could also be explained by an articulatory strengthening. The elevation and the lowering of the velum position seem to be mainly governed by the activity of the levator-palatini, its contraction and its decontraction, respectively. An articulatory strengthening for non-nasals in initial position would thus increase the levator palatini contraction, resulting in a higher velum position. Conversely, for nasals, this strengthening would reduce the decontraction of the levator-palatini, resulting in a smaller lowering of the velum (see Straka 1963, Fujimura 1990).

(3) vowel glottalization observed in initial position in higher constituents could also result from articulatory strengthening if an increase in laryngeal muscles' contraction affects the constriction of the arytenoids, resulting in a perturbation of vocal fold vibrations.

Even though the definition adopted here remains to be tested empirically, articulatory strengthening appears as a conceivable unifying physiological mechanism to account for the articulatory modifications observed for various segments in domain-initial position. Articulatory strengthening seems here to be locally applied to the initial position of a constituent, and to be a function of the height of that constituent. Therefore, the distinction between prosodic boundaries seems to consist of a gradual increase in articulatory strengthening from lowest to highest prosodic constituents. In that sense, it is conceivable that articulatory strengthening is one of the physiological correlates of the prosodic phrasing of an utterance. However, considering the variability observed in the realization of the articulatory properties in each of the domain-initial position studied, it seems that articulatory strengthening should be, at this point, considered mainly as a trace of the prosodic encoding of utterances rather than an intended cue to phrasal organization.

CONCLUSION

While the segmental and suprasegmental aspects of speech have been most often studied separately, it becomes more and more obvious that these two aspects are closely connected. The signaling of prosodic boundaries is undoubtedly multi-parametric. Final lengthening and melodic contours may be the most robust and the most important acoustic cues for the perception of phrasing. However, this study of French confirms the fact that the prosodic phrasing of an utterance is also reflected in the articulation of initial segments. The relevance of this finding for perception, as well as its linguistic function, remain to be explored.

Acknowledgments

I would like to thank my four subjects, and particularly Laurent Girard, for participating in these experiments. I am particularly grateful to Pat Keating for her help and motivating spirit that accompanied me during my stay at UCLA, and to all the other members of the UCLA Phonetics Lab for comments on various presentations of this work. This study is part of a dissertation defended in Paris III in June 1998. This work was supported by a Research Allocation from the French Research Ministry and by NSF grant SBR-9511118 to P. Keating. The author is currently funded by grant 1113-049698.96 from the Fond National de Recherche Suisse.

5. References

Beckman, M. & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P.A. Keating (Ed.) Papers in Laboratory Phonology III: Phonological structure and phonetic form, 7-33. Cambridge UK: Cambridge U. Press.

Beckman, M., Edwards, J., and Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. Docherty and D. R. Ladd (Eds.) Papers in Laboratory Phonology II: Gesture, Segment, Prosody, 68-86. Cambridge UK: Cambridge U. Press.

Bell, A. & Hooper, J. (1978). Issues and evidence in syllabic phonology. In A. Bell & J. Hooper (Eds.) Syllabes and segments, 3-24. Amsterdam: North-Holland.

Bell-Berti, F. (1993). Understanding velic motor control: studies of segmental context. In M. K. Huffman & R. A. Krakow (Eds.), Nasals, nasalisation and the velum, 63-86. San Diego: Academic Press.

Benguerel, A.-P. (1977). Velar coarticulation in French: a fiberscopic study. *J. of Phonetics*, **5**, 145-158.

Bourciez, E. & Bourciez, J. (1967). Phonétique française, étude historique. Coll. Traditions de l'Humanisme, Klincksieck.

Brunot, F. & Bruneau, C. (1937). Précis de grammaire historique de la langue française. Paris: Masson & Cie.

Byrd, D. (1994). Articulatory timing in English consonant sequences. PhD. diss. UCLA. (distributed as *UCLA W.P.*, **86**)

Byrd, D., Kaun, A., Narayanan, S. & Saltzman, E. (1996). Phrasal signatures in articulation. Papers in Laboratory Phonology V, (to appear). Paper presented at the 5th conference on laboratory phonology, july 1996.

Cho, T. & Keating, P.A. (1999) Articulatory and acoustic studies of domain-initial strengthening in Korean. *This volume.*

Cooper, A. (1991). Laryngeal and oral gestures in English /p, t, k/. *Proceedings of the 12th Int. Congress of Phonetic Sciences*, **2**, 50-53.

Debrock, M. (1977). An acoustic correlate of the force of articulation. *J. of Phonetics*, 5, 61-80.

de Jong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *JASA*, **97**, 491-504.

Delattre, P. (1940). La force d'articulation consonantique en français. *The French Review*, **t. XIV**, 220-232

Delattre, P. (1965). Comparing the phonetic features of English, French, German and Spanish: an interim report. New York: Chilton Books

Dilley, L. & Shattuck-Hufnagel S. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *J. of Phonetics*, **24**, 423-444.

Edwards, J. & Beckman, M. (1988). Articulatory timing and the prosodic interpretation of syllable duration. *Phonetica*, **45**, 156-174.

Farnetani, E. &. Vayra. M. (1996). The role of prosody in the shaping of articulation in italian CV syllables. *Proceedings of the 1st ESCA Workshop on Speech Production Modeling, 4th Speech Production Seminar,* 9-12, Autrans.

Flege J.E. (1982). Laryngeal timing and phonation onset in utterance-initial English stops. *J. of Phonetics*, **10**, 177-192.

Fletcher, S. G. (1989). Palatometric specification of stop, africate, and sibilant sounds. *J. of Speech and Hearing Research.*, **32**(736-748).

Fougeron, C. & Keating, P. (1997). Articulatory strengthening at edges of prosodic domain. *JASA* **106 (6)**, 3728-3740.

Fougeron, C. (1998). Variations articulatoires en début de constituants prosodiques de différents niveaux en français. Thèse de doctorat, Univ. Paris III.

Fujimura, O. (1977). Recent findings on articulatory processes. In Carré, Descout & Wajskop (Eds.) Articulatory modeling and phonetics, 115-126. GALF.

Fujimura, O. (1990). Methods and goals of speech production research. *Language and Speech*, **33(3)**, 195-258.

Giot J. (1977). Etude comparative de syllabes accentuées et prétoniques du français sur les plans articulatoire et acoustique. *Travaux de l'Inst. de Phonétique de Strasbourg*, **9**, 89-169.

Gordon, M. (1996) The effect of stress and prosodic phrasing on duration, acoustic amplitude and air flow of nasal consonants in Estonian. *UCLA WP*, **92**, 151-159

Honda, K., Hirai, H. & Kusakawa, N. (1995). Modeling vocal tract organs based on MRI and EMG observations and its implication on brain function. *R.I.L.P.*, **27**, 37-50.

Hsu, C-S & Jun, S-A (1997) Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic?. (unpublished manuscript).

Jun, S.-A. (1993) The Phonetics and Phonology of Korean Prosody. PhD diss. OSU

Kahn, D. (1976). Syllable based generalizations in English phonology. Indiana Univ. Linguistics Club, Bloomington, Indiana.

Keating, P.A., Cho, T., Fougeron, C. & Hsu, C.-S. (1998). Domain-initial articulatory strengthening in four languages. To appear in Papers in Laboratory Phonology 6.

Keating, P.A., Wright, R. & Zhang J. (1999) Word-level asymmetries in consonant articulation. *This volume.*

Kent, R.D. & Netsell, R. (1971). Effects of stress contrast on certain articulatory parameters. *Phonetica*, **24**, 23-44.

Klatt, D.H. (1975). Vowel lengthening is syntactically determined in connected discourse. *J. of Phonetics*, **3**, 129-140.

Kohler, K. .J. (1994). Glottal stops and glottalization in German. *Phonetica*, **51**, 38-51

Krakow, R. A. (1989) The Articulatory Organization of Syllables: A Kinematic Analysis of Labial and Velic Gestures. PhD diss. Yale Univ.

Krakow, R. A., Bell-Berti, F., & Wang, Q.E. (1994). Supralaryngeal declination: evidence from the velum. In F. Bell-Berti and L. Raphael (Eds.) Producing speech: A Festschrift for Katherine Safford Harris, 333-353. Woodbury NY: AIP Press.

Lehiste, I. (1960) An acoustic-phonetic study of internal open juncture. *Supplement to Phonetica,* **5.**

Lehiste, I. (1964) Juncture. *Proceedings of the 5th Int. Congress of Phonetic Sciences*, 172-200

Lindblom, B (1963). Spectrographic study of vowel reduction. *JASA*, **35(11)**, 1773-81.

Lisker, L. & Abramson, A. (1967). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, **20**, 384-422.

Malmberg, B. (1943) Le système consonantique du français moderne. Leipzig: Harrassowitz.

Marchal, A. (1979). Etude des occlusives t/d du français par l'electroplalatographie. *Proceedings of the9th Int. Congress of Phonetic Sciences,* **1**, 203.

Ohala, J. & Kawasaki, H. (1984) Prosodic phonology and phonetics. *Phonology Yearbook*, **1**, 113-127

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *JASA*, **54(5)**, 1235-1247.

Payan, Y & Perrier, P (1996) Controle de la langue en parole: quelques propositions testées sur une modélisation biomécanique. *Actes des XXIeme JEP*, 79-82.

Pierrehumbert, J. & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty and D. R. Ladd (Eds.) Papers in Laboratory Phonology II: Gesture, Segment, Prosody, 90-117. Cambridge UK: Cambridge U. Press.

Rousselot, P.J. (1901). Principes de phonétique expérimentale. Paris: Welter.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook*, **3**, 371-405.

Shadle, C.H. & Scully, C. (1995). An articulatory-acoustic-aerodynamic analysis of [s] in VCV sequences. *J. of Phonetics*, **23**, 53-66.

Simon, P. (1967) Les consonnes françaises : mouvements et positions articulatoires à la lumière de la radiocinématographie. Thèse de doctorat, Université de Strasbourg.

Stone, M., Faber, A., Raphael, L.J. & Shawker, T.H.(1992). Cross-sectional tongue shape and linguoplalatal contact patterns in [s], [ʃ], and [l]. *J. of Phonetics*, **20**, 253-270.

Straka, G. (1963). La division des sons du language en voyelles et consonnes peut-elle être justifiée? *Travaux de Ling. et de Littérature, Univ. de Strasbourg*, **1**, 17-99.

Straka, G. (1964). L'évolution phonétique du latin au français sous l'effet de l'énergie et de la faiblesse articulatoires. *Travaux de Ling. et de Littérature, Univ. de Strasbourg*, **2**, 17-98.

Vaissière, J. (1983). Language-Independent Prosodic Features. In A. Cutler & D.R. Ladd (eds) Prosody: Models and Measurements. Berlin: Springer-Verlag, 53-66.

Vaissière, J. (1986). Variance and Invariance at the Word Level. Invited Contribution, in J.S. Perkell & D.H. Klatt (Eds.), Invariance and variability in speech processes, 534-53. Lawrence Erlbaum Ass.

Vaissière, J. (1988). Prediction of velum movement from phonological specifications. *Phonetica*, **45**, 122-139.

Vaissière, J. (1992). Rhythm, accentuation and final lengthening in French. In J. Sundberg, L. Nord & R Carlson (Eds.), Music, Language, Speech and Brain, 108-120. Stockholm: Wenner-Gren Int. Symposium Series.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., Price, P.J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *JASA*, **91(3)**, 1707-1717.

Articulatory and acoustic studies of domain-initial strengthening in Korean

Taehong Cho and Patricia Keating

1. Introduction

A traditional phonetic distinction (made, for example, by the IPA) is that between *segmental* and *suprasegmental* properties. Segmental properties are those having to do with the intrinsic featural content of individual consonants and vowels. Suprasegmental properties, in contrast, hold above the individual segment, characterizing a sequence of segments, or one segment relative to another, or the position in which some segment occurs. Lehiste (1970) distinguishes segmental from suprasegmental as paradigmatic vs. syntagmatic. In the IPA tradition, suprasegmental properties include quantity, stress, tone, and juncture.

The fundamental idea behind this distinction is that each segment has some basic or inherent phonetic characterization, to which the suprasegmental properties of a particular utterance can be added. Some articulatory or acoustic dimensions are thought to serve primarily segmental purposes—for example, oral constrictions or formant frequencies. Other dimensions are thought to serve primarily suprasegmental purposes—for example, f0 or loudness. A few dimensions are considered to be available for both purposes—for example, segment duration.

Nonetheless, it has become clear that there is no simple distinction between segmental and suprasegmental dimensions. Consider, for example, f0. Although the most important and salient uses of f0 are suprasegmental—for example, tone and intonation—f0 also varies as a function of segmental identity. Thus, high vowels generally have higher "intrinsic" f0 than low vowels (Lehiste, 1970). Conversely, a prototypically segmental property like vowel quality is known to be partly a function of suprasegmental factors such as stress and length (Lehiste, 1970). Indeed, recent work has established that many segmental properties may be affected by suprasegmental, or prosodic, structure. (By *prosody* we mean all aspects of the higher-level (suprasegmental) organization of speech; see Shattuck-Hufnagel & Turk, 1996.)

As an example of this more recent literature, Jun (1993) showed that the Voice Onset Time (VOT) of Korean aspirated /p^h/ depends on the position of /p^h/ in its word and phrase. In her study, VOT was systematically longer at the beginning of a word than medially in a word, and longer still at the beginning of a small phrase. Thus, while Korean aspirated stops are always aspirated, the degree of aspiration varies prosodically. Jun suggested that this variation reflected a strength hierarchy of prosodic positions. Such VOT variation can be interpreted as reflecting prosodically-conditioned differences in the degree of glottal opening during spread-glottis consonants (Pierrehumbert & Talkin, 1992), as was recently confirmed for Korean by Jun, Beckman and Lee (1998). That is, it seems that in Korean at least one segmental dimension, glottal opening for an aspirated consonant, is in part also suprasegmental, controlled by prosody.

Fougeron and Keating (1997) extended this kind of result to an oral articulation, the oral constriction for an alveolar stop consonant in English. They showed that the extent of constriction for English /n/ depends on the position of the /n/ not only in its word and phrase, but even in different kinds of phrases. Consonants at beginnings of phrases were more constricted than consonants in the middles of phrases. Furthermore, consonants at the beginnings of larger phrases were more constricted than consonants at the beginnings of smaller phrases, or words. English speakers in the study sometimes tuned their articulations so as to distinguish four different phrasal positions. Fougeron and Keating, like Jun, interpreted their result as a kind of strengthening of a consonant's oral articulation according to the strength of its prosodic position — the stronger the position, the stronger the articulation — and referred to their result as *domain-initial strengthening*. Similar results were also found by Fougeron & Keating (1996) and Fougeron (1998) for French, and Hsu & Jun (1998) for Taiwanese. Fougeron & Keating (1996) and Fougeron (1998) also found prosodically-conditioned variation in the nasalization of French /n/. On this view, then, a variety of phonetic properties, whether traditionally considered "segmental" or "suprasegmental," may be finely tuned by prosodic structure.

In some respects Korean would be expected to replicate the results from these previous studies. Jun has already shown that Korean VOT, a laryngeal dimension, is sensitive to at least some aspects of prosodic structure. Furthermore, the Korean rule of "Lenis Stop Intervocalic Voicing," by which lenis stops /p t k/ are voiced to [b d g], is one of the best-known prosodically-conditioned segmental lenitions of the world's languages; its effects are so strong that they are audible without any instrumental analysis. Jun (1993; 1998) found that this voicing is generally constrained by the consonant's position in a phrasal domain, in that consonants voice when they are inside an Accentual Phrase. In addition to these prosodic effects on Korean laryngeal activity, Jun also showed that various other Korean segmental processes are sensitive to prosodic domain boundaries. For these reasons we might well expect Korean to show other, less audible, prosodic effects on consonant articulations. Therefore we hypothesize that Korean should show clear effects of position in prosodic domain on oral consonant articulations, as well as on glottal articulations. On the other hand, at the same time Korean presents an interesting apparent counter-example to the general finding that Word-medial position is weaker than Word-initial position. This counter-example comes from the behavior of the Korean tense (fortis) stops. In Korean, it is generally reported that Word-medial tense stops lengthen (e.g., Silva, 1992; Han, 1996), which could perhaps be correlated with more forceful articulation than for Word-initial tense stops. Thus Word-medial position appears to be a weak position for lenis and aspirated stops (which show more voicing and less aspiration, respectively) but not for tense stops, in Korean. Perhaps these tense consonants are differently affected in other prosodic positions as well.

In this paper we extend Jun's 1993 work on Korean VOT in several ways. First, we look at additional phrasal prosodic domains in Korean, following Fougeron & Keating (1996; 1997). Second, like Fougeron & Keating we include a measure of oral constriction of stop consonants, thus replicating in Korean their study of prosodic effects on oral constrictions. Third, we examine not just VOT, but also several other acoustic measures. Fourth, we consider more than one consonant type — all four Korean dental stops, oral and nasal. Our hypotheses are that Korean will show domain-initial strengthening of oral consonant articulations and of at least some acoustic measures; that this strengthening will differentiate several phrasal prosodic domains; but that the tense dental stop might not show such strong effects.

Studying prosodic effects on articulation requires an independent scale of prosodic position and strength. A well-known scale of this kind is the *prosodic hierarchy*. The hypothesis of a prosodic hierarchy is that speech utterances are hierarchically organized, with higher (or larger) units being decomposed into lower (or smaller) constituents. These prosodic constituents, or domains, can be in part derived from syntactic constituents (Selkirk, 1984, 1986; Nespor & Vogel, 1986) and can be identified on the basis of segmental phonological rule application (e.g., Nespor & Vogel, 1986; Jun, 1993) and/or intonation (e.g., Beckman & Pierrehumbert, 1986; Jun, 1993).

In this work we use the model of Korean prosodic structure of Jun (1993) and Beckman & Jun (1996), which builds on earlier work concerned with syntactic bases for prosody (e.g., Y. Cho, 1990; Silva, 1992; Kang, 1992) and with its intonational correlates (e.g., de Jong, 1989; Lee, 1989). For Korean, Jun (1993; 1998) has shown that the prosodic hierarchy must include at least two phrasal prosodic domains (Accentual Phrase, Intonational Phrase) as well as the prosodic Word plus any intra-Word domains. We also add a possible higher prosodic constituent, the *Utterance*. This model is shown in Fig. 1, which gives a sample structure showing the hierarchical organization of the prosodic domains.

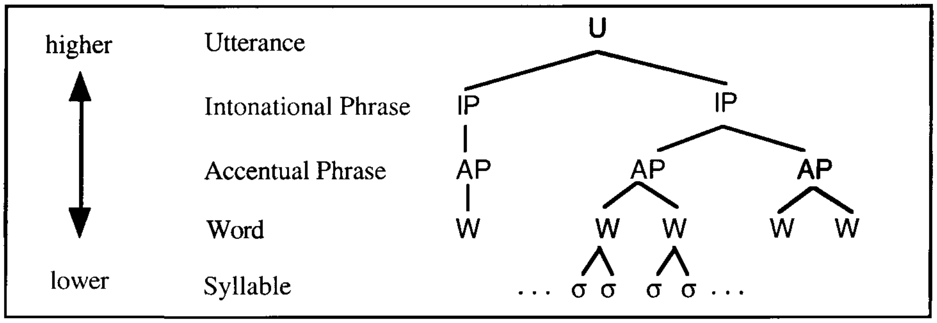


Figure 1. Prosodic structure of Korean (adopted from Jun, 1993 and Beckman & Jun, 1996, with Utterance level added in this study)

As seen in the figure, *Syllables* (s) are grouped into *Words* (W); Words are grouped into *Accentual Phrases* (AP); *Accentual Phrases* are grouped into *Intonational Phrases* (IP). An Accentual Phrase is usually marked by a (LH)LH tonal pattern; an Intonational Phrase, by a final tonal pattern along with a substantial final lengthening. The highest level *Utterance* (U, usually punctuated by a period), into which Intonational Phrases are grouped, is distinguished from IP by a pause, and by a more marked final intonation. The domain U was added in order to see if an orthographic period triggering a substantial pause causes any articulatory and/or acoustic differences compared to IP. Following the Strict Layer Hypothesis (Selkirk, 1986), it is assumed that the beginning and end of each higher domain is also the beginning and end of lower domains.

2. Method

2.1. *Electropalatography (EPG)*

In our articulatory experiment we examine variation in the oral articulation of consonants. The articulatory measure for consonants was linguopalatal contact, contact between the tongue (especially tongue blade and front) and the hard palate. The amount of linguopalatal contact indicates the degree of overall oral constriction: the more linguopalatal contact, the greater oral constriction with a greater articulatory magnitude. Linguopalatal contact was measured by electropalatography (EPG) using the Kay Elemetrics Palatometer 6300. The Kay Palatometer uses custom-fabricated pseudo-palates made of a thin acrylic, held in place by wrapping around the upper teeth. A pseudo-palate has 96 electrodes, as shown in Fig. 2 (in section 2.3), covering the entire hard palate, the entire inside surfaces of the upper molars, and part of the inside surfaces of the upper front teeth. In order to capture more denti-alveolar contact, a custom configuration of electrodes was designed in which two electrodes were located lower on the upper front teeth; these are shaded in the figure. When an electrode is contacted by the tongue, a circuit is completed and the contact is recorded by the Palatometer. Each sweep of the 96 electrodes takes 1.7 ms, and the sampling interval is 10 ms.

2.2. *Test sentences and Procedure*

The test consonants are the four anterior coronal stops /n, t, t^h, t*/ (where /t*/ represents the fortis, or tense, stop, for which there is no official IPA transcription). Each test consonant was placed in a fixed segmental context within a set of sentences. Example sentences for /n/ are shown in Table I. The sets of sentences were constructed to vary in their likely phrasing, so that the prosodic context of the test consonants would vary. When subjects produced these sentences with the expected phrasings, then the test consonant was initial in a prosodic domain that varied systematically from high (Utterance) to low (Syllable). Table I characterizes the prosodic position of the test consonants as the highest prosodic domain in which the consonant is initial. Thus "Ui" means that the highest domain in which this consonant is initial, is the Utterance (U). The number of syllables preceding each test consonant was also controlled, in order to factor out the possibility of the articulatory declination (Krakow, Bell-Berti, & Wang, 1994) which may induce differences in articulatory magnitude (though this was not seen in Fougeron and Keating's (1997) study of English).

TABLE I. Test sentences for the target consonant /n/. (See Appendix 1 for /t, t^h, t*/.)

(a) Higher levels above Word for /n/ in /a#_a/. (# = a prosodic boundary)

U-Initial (Ui)	igosɨn patakk**a.** [$_{U}$ **na**mdʒuga jəgisə sanda.] This place seashore NAMJOO here lives 'This place is the seashore. NAMJOO lives here.'
IP-Initial (IPi)	igosɨn patakk**a,** [$_{IP}$ **na**mdʒue kohjaŋida.] This place seashore NAMJOO's hometown 'This place is the seashore, which is NAMJOO's hometown.'
AP-Initial (APi)	igosɨn patakk**a** [$_{AP}$ **na**mtʃ*oge] itt*a This place seashore south_Loc be 'This place is located to the south of the seashore.'
Word-Initial (Wi)	igosɨn [$_{AP}$ patakk**a** ($_{W}$ **na**mdʒaga)] sanɨn goʃida. This place seashore man-Nom live-REL place-Dec. 'This place is where the seashore man lives.'

(b) Word and Syllable levels for /n/ in /o#_ɛ/. (# = a prosodic boundary)

Word-Initial (Wi)	kɨjədʒanɨn [**AP** maɾɨmmo (**W** **n**ɛgiɾɨl)] tʃɛanhɛtt*a. the woman-Top. parallelogram betting-acc. suggested 'The woman suggested betting with the parallelogram (on something).'
Syllable-Initial (IPi)	kɨjədʒanɨn [**AP** jəɾɨm (**W** mo(**S** **n**ɛgiɾɨl)] tʃɛanhɛtt*a. the woman-Top. summer harvest-acc. suggested 'The woman suggested fall harvest.'

As it was difficult or impossible to construct meaningful sentences with the same segmental context and syllable count for all prosodic conditions, the sentences for each test consonant were constructed in two subsets. Those shown in (a) in Table I provide matched comparisons at and above Word-initial position; those in (b) provide a comparison at and below Word-initial position.

Two male and one female Seoul Korean speakers participated in the experiment. Speakers NHL and JYY (students at UCLA) had been in America for 2 and 1 years, respectively, at the time of recording and Speaker THC (one of the authors) for 3 years. Each test sentence was repeated 20 times. The audio and EPG signals, with 12.8 kHz and 100 Hz sampling rates respectively, were recorded directly into the computer through Kay Elemetrics's Computerized Speech Lab (CSL) and Palatometer. In total, 960 sentences (4 levels x 4 consonants x 3 speakers x 20 repetitions) were analyzed for higher level comparisons (U, IP, AP, W) and 480 sentences (2 levels x 4 consonants x 3 speakers x 20 repetitions) were analyzed for the lower level comparisons (Word, Syllable).

Subjects were told in part how to phrase each sentence for some prosodic boundaries. They were asked to pause after a period (which marked an Utterance boundary) but not to pause after a comma (which was used to induce an IP boundary). No instructions other than this were given to the speakers. For the AP and Word boundaries, subjects in general made the intended intonational contours with an appropriate break without being told to do so. Each uttered sentence was checked during the recording session by the first author who is a trained K-ToBI (Beckman & Jun, 1996) transcriber. Whenever subjects made a mistake or produced an unintended rendition, they were asked to read it again for several times from which the best rendition was chosen. The criteria used for partial prosodic coding (i.e., for coding the type of prosodic domains around which the target consonants occur) are summarized in Table II.

TABLE II. Criteria used for partial prosodic coding, based on K-ToBI (Beckman & Jun, 1996).

U	Period mark used for triggering a pause as well as IP boundary. • Boundary Tone (L%) • Preceded by a pause (greater than break index number 3 in K-ToBI)
IP	Comma used for triggering IP boundary • Boundary Tone (usually H% but sometimes HL% or L%) • Considerable final lengthening with break index number 3
AP	Adverbial Phrase used for triggering AP boundary • (LH)LH or (HH)LH phrasal tones • Break index number 2
W	Second word of two successive words which are grouped into AP • No tonal specification, but in general, initial syllables are associated with AP phrasal L • Virtually no perceived break (break index number 1)
σ	Second syllable of the second word of two successive words which are grouped into AP • No tonal specification • Break index number 1, same as Wi except that the break is lexically determined

2.3. *Measurement*

For analysis of linguopalatal contact, three different regions on the palate were defined as shown in Fig. 2. The entire region covers all 96 electrodes; the front region, which is a region of primary coronal consonant contact, includes the front 47 electrodes covering the region from the front teeth to the alveolo-palatal area; and the back region, which is a region of vowel contact, includes the back 49 electrodes covering mainly the mid-palatal area. Analyses are primarily based on contact patterns in the entire region, unless otherwise specified. Linguopalatal contacts in the front and back regions will be compared to those in the entire region to see which part of the tongue and palate contributes more to the overall variations.

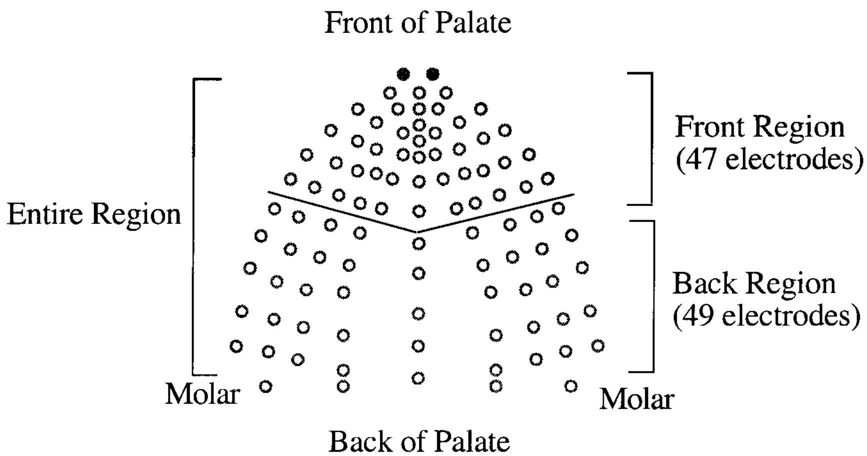


Figure 2. Placement of 96 electrodes with three analysis regions.

2.3.1. *Linguopalatal Contact*

EPG data were analyzed by computing the percent of the electrodes in a region contacted in each data frame (see Byrd, Flemming, Mueller, & Tan, 1995 for detailed method). For the entire region of 96 electrodes, one contacted electrode is thus approximately equal to one percentage point. For each consonant, peak linguopalatal contact was measured in the single frame that shows the most extreme contact for that segment. Fig. 3 shows data for sample tokens of /n/ by Speaker NHL with variations in the amount of peak linguopalatal contact as a function of prosodic position. As shown in the figure, the peak contact measure captures variations in the amount of linguopalatal contact of domain-initial consonants at different prosodic levels. It does not, however, capture other aspects of the variation, such as the shift in place of articulation seen here.

2.3.2. *Seal Duration*

The duration between the first and the last frames in which the oral cavity was completely sealed, called the Seal Duration, was measured.

2.3.3. *Skewness of Articulatory Movement*

Overall contact trajectories (Barry, 1991) as a function of time within each token were treated as distributions, and the skewness of each trajectory's shape was measured. A movement trajectory is said to be skewed if the highest point of the trajectory does not correspond to the center of the distribution (i.e., median), thus causing one of the 'tails' in the distribution to be longer than the other. The trajectory's shape is 'skewed to the right' or 'positively skewed' with a longer right 'tail' and 'skewed to the left' or 'negatively skewed' with a longer left 'tail.' Byrd *et al.* (1995) suggest that the degree of asymmetry between the onset and offset of articulatory movement can be measured by the skewness of the linguopalatal contact profile. A positive skew indicates that the consonant closing gesture occurs more quickly than its opening or releasing gesture, for the portion of the gesture during which linguopalatal contact occurs. On the other hand, a negative skew indicates that the consonant closing gesture occurs more slowly than its opening or releasing gesture. Fougeron & Keating (1997: 3737) discuss the possibility that higher movement velocity would result in a "greater impact of the tongue against the palate at closure" which may in turn result in a greater linguopalatal contact. If this is true, we would expect relatively greater positive skewness for a consonant with greater linguopalatal contact. So, the measurement of skewness examines whether any variation in the amount of linguopalatal contact can be partly attributable to different speeds of closing formation, relative to the releasing formation. To get the contact profile from which the skewness was calculated, an approximation of the articulatory movement was identified for each consonant: from the first frame where the tongue, in beginning to contact the hard palate area progressively forward, reaches 20% contact, to the last frame where the tongue is moving away from the consonant contact, but contact remains above 20%. The 20% threshold was used because for some cases, especially lower domains, it was difficult to locate an exact onset or offset of the movement. Thus, the approximated articulatory movement whose skewness was calculated, is within the interval between the onset of the consonant closing gesture and the offset of the consonant opening gesture. Fig. 4 shows a sample contact profile with a negative skewness (skew = -1.384). This particular sample profile suggests that the closing gesture towards the maximum linguopalatal contact is slower than the releasing gesture. The skewness was measured for three stops /t, t^h, t*/ for the higher levels, Ui to Wi.

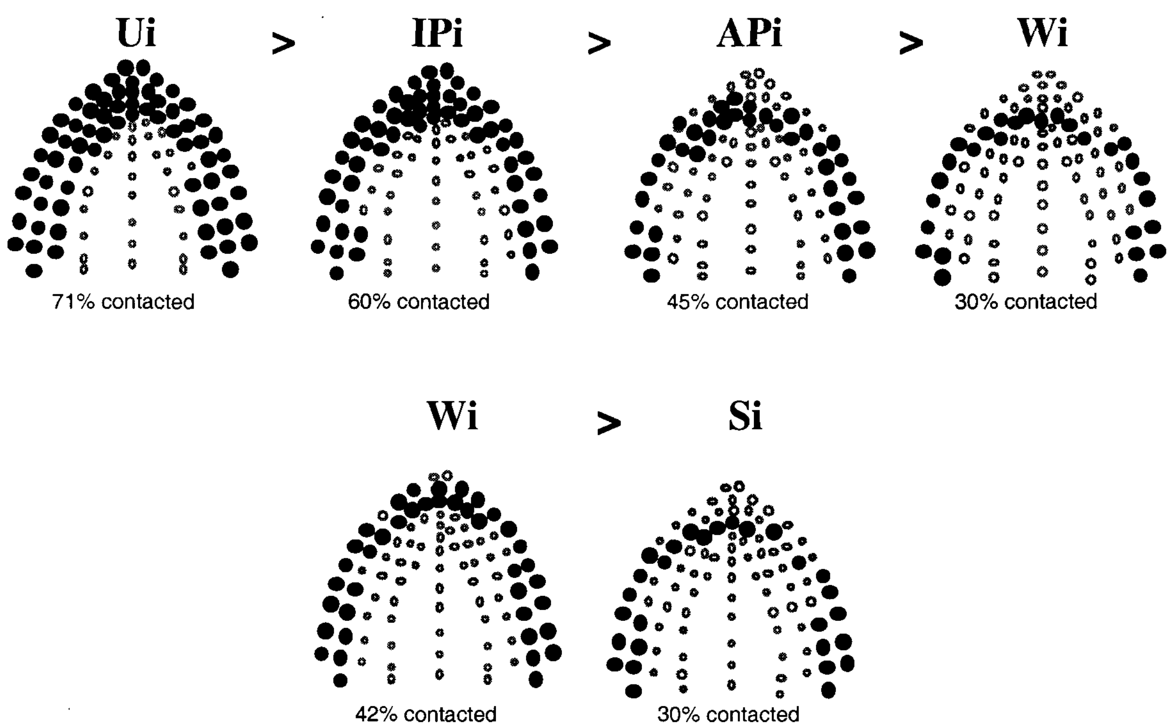


Figure 3. Sample tokens of linguopalatal contacts for /n/ by Speaker NHL, from 2 experiments. Wi vs. Si comparisons were made separately from the other comparisons due to the difficulties of controlling for segmental context. For this lower level comparison, contact in Wi is greater and more fronted than that in Wi for higher level comparisons, because the following vowel is front.

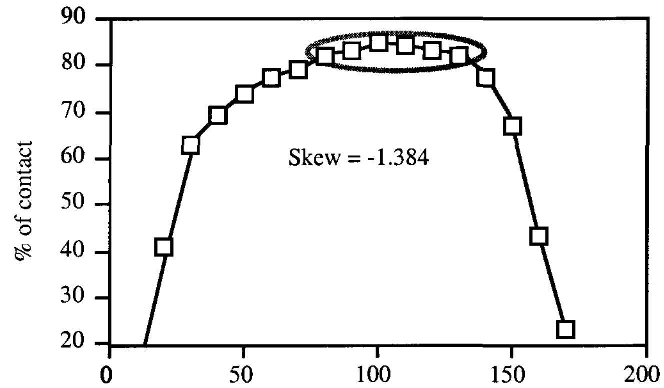


Figure 4. Sample contact profile with a negative skewness for /t*/ in U-initial position. Note that the circled maximum contact area is formed to the right of the median along the time dimension.

2.3.4. *Acoustic Parameters*

Several standard acoustic measures were made from the audio signal:

• *V1 and V2 duration:* The durations of the vowel before a boundary (V1) and the vowel after a boundary (V2) were measured. Note that V1 is always Word-final (in an open syllable) and therefore likely to show any pre-boundary lengthening.

• *Closure Duration:* Acoustic closure duration for all test consonants was taken from spectrograms. For oral stops this measure included both voiced and voiceless portions of closure, from the offset of F2 in V1 to the beginning of the stop burst. For /n/ this measure was nasal duration, taken from the onset to the offset of nasal energy displayed in the spectrogram.

• *Voicing During Stop Closure:* Voicing residue interval during stop closure, indicated by the voicing bar (glottal pulsing) at the bottom of the spectrogram, was taken from the offset of F2 in V1 to the onset of the completely voiceless closure. Then, the percent of this interval relative to the entire closure duration was calculated, for /t, t^h, t*/, for the higher levels Ui to Wi. The presence of voicing at the beginning of the closure of a voiceless stop indicates that the vocal cords have not yet completely abducted; to some extent it also reflects vocal tract tension, since closure voicing is more likely with less vocal tract tension.

o *Voice Onset Time:* VOTs for /t, t^h, t*/ were taken from the point of the stop release to the voice onset of the following vowel, as seen in spectrograms in F2 and above. Thus for the lenis stops, any breathy voicing with only low-frequency harmonics was included in the VOT.

• *Total Voiceless Interval:* The interval from the onset of the voiceless portion of closure to the voice onset (i.e. voiceless closure plus VOT) was measured for just the aspirated stop /t^h/. This measure reflects the total time of glottal opening during the stop.

• *RMS Burst Energy:* The acoustic energy at the burst was measured from an FFT spectrum giving the RMS value over all frequencies. A 10 ms (124 point) window was centered over the release of the stops /t, t^h/; for the tense stop /t*/ which has a very short VOT, a smaller window (5 ms, 64 point) was used in order to prevent the window from including the following vocalic energy. The burst energy for Word-initial lenis stop /t/ could not be measured in a comparable fashion, due to the presence of voicing throughout the consonant; therefore only tokens from higher levels were measured.

• *Nasal Energy Minimum*: The energy during nasal consonants was measured as the lowest point (valley) of the RMS acoustic energy profile. The minimum value was measured because the energy profile of a nasal consonant has less energy than in surrounding vowels. (Peak energy is always found at the edges of a nasal, next to surrounding vowels, and thus does not constitute an independent measure.) However, for nasals adjacent to a pause, the minimum energy value is always associated with the pause, and thus nasal energy was not measured for Utterance-initial (pause-adjacent) tokens.

The coded data were submitted to statistical analyses using StatView (Abacus Concepts, 1992). For most cases, analyses are based on one-factor analysis of variance (ANOVA) and Fisher's protected least significant difference (PLSD) *post-hoc* pairwise comparisons, with 0.05 as the significance level for testing a variety of factors. In general, each speaker was considered separately because of the interspeaker variations. Descriptions of individual tests used in the present study will be given in detail in each results section below.

3. Results

3.1. *Variation in Articulatory Parameters as a function of prosodic position*

3.1.1. *Linguopalatal Contact*

We first examine the extent to which linguopalatal contact over the entire pseudo-palate varies with the hierarchically-nested prosodic positions (Word-initial and higher), specifically to test the hypothesis that a higher-domain-initial consonant is produced with a greater linguopalatal contact compared to a lower-domain-initial one. Two-factor ANOVA with factors Prosodic Position and Speaker showed that the peak linguopalatal contact varies significantly depending on the prosodic position ($F(3, 952) = 191.806$, $p < 0.0001$). Fisher's PLSD *post hoc* pairwise comparison confirmed that all domains are differentiated from one another by peak contact (at a significance level of 0.05) in decreasing order Ui, IPi, APi, and Wi, as shown in Fig. 5a. There is also a significant interaction between the two factors due to differences among speakers. For each speaker, there is a significant effect of prosodic position on linguopalatal contact ($F(3, 316) = 94.54$, $p < 0.0001$ for NHL; $F(3, 316) = 91.127$, $p < 0.0001$ for THC; $F(3, 312) = 124.848$, $p < 0.0001$ for JYY). According to *post hoc* pairwise comparison two speakers, NHL and THC, make a four-way distinction among Ui, IPi, APi, and Wi, while Speaker JYY makes a only three-way distinction, with no difference between APi and Wi. This general pattern is also maintained when each consonant is considered separately, as shown in Fig. 6a. All levels for each consonant were found to be significantly differentiated by contact (at a significance level of $p < 0.01$), except in the comparison between APi and Wi for /t/.

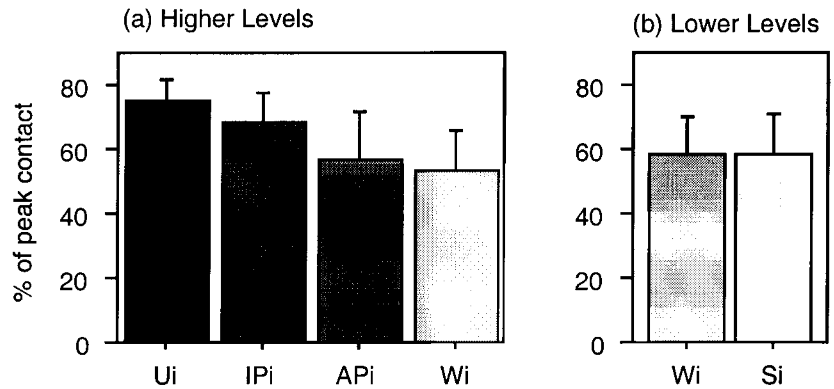


Figure 5. Peak contact across speakers and consonants. (a) higher levels from Ui through Wi and (b) lower levels between Wi and Si. Error bars = standard deviations.

A similar effect is observed when the front and back palate regions were considered separately, as shown in Figs. 6b-c. Four separate one-factor ANOVAs (4 consonant types) in both the front and back regions confirm significant effects of prosodic position on linguopalatal contact at $p < 0.0001$. For each case, the results of *post hoc* pairwise comparisons show that at least a three-way distinction is made. However, as shown in the figure, the degree of variation in the front region (roughly, tongue blade contact) is similar to that in the entire region, while the back region (roughly, tongue body contact) shows less variation. This difference shows that it is the front region that contributes more to the variations of linguopalatal contact as a function of prosodic position ($F(3,952) = 171.626$ for Front Region; $F(3,952) = 120.691$ for Back Region).

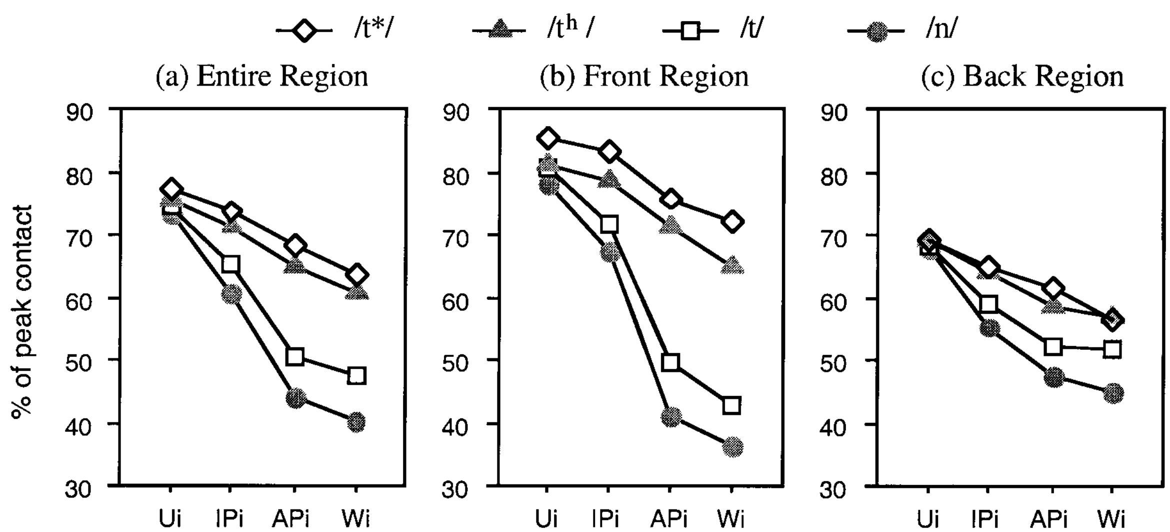


Figure 6. Peak contact for each consonant. (a) contact over the entire pseudo-palate area; (b) contact in the front region and (c) contact in the back region (see Fig. 2). Data pooled across speakers.

We also examined EPG displays (and the locations of electrodes on the individual speakers' pseudo-palates) to see how consistently the place of articulation of the stops varied with prosodic position. It was noted above that Fig. 3 shows a shift in place, a backing as the stop moves from higher to lower domains. This difference is consistent for all three speakers for /n/. When this consonant has more contact, its place is dentialveolar, but when it has less contact, its place is palato-alveolar. There is a similar, but less dramatic, effect for /t/: when /t/ has less contact, its place is alveolar. However, there is no equivalent place shift for /t^h/ and /t*/.

Fig. 6 shows that all four test consonants are affected by prosodic position; this was also tested separately for each speaker. Twelve separate one-factor ANOVAs (3 speakers x 4 consonant types) were generated with accompanying Fisher's PLSD *post hoc* pairwise comparisons. All 12 cases were found to show significant effects of prosodic position on the linguopalatal contact at $p <$ 0001. The results of *post hoc* pairwise comparisons are given in Table IIIa. What is common for all consonant types is that speakers make at least a three-way distinction between Ui and Wi. Speaker NHL makes a four-way distinction for /n, t/ and a three-way distinction for /t^h, t*/ (with no difference between Ui and IPi); Speaker THC makes a four-way distinction for all consonants; and Speaker JYY makes a three-way distinction for /n, t*/, and a four-way distinction for /t^h, t/. In addition, speakers always make a distinction between APi and the higher domains IPi and Ui. However, Speaker JYY has greater linguopalatal contact for Word-initial /t/ than for AP-initial /t/, which is opposite to the general trend. In general, similar patterns are observed across consonants. The major differences are whether distinctions between APi and Wi (for JYY) or between UI and IPi (for NHL) are made. These differences are responsible for the interspeaker variation.

Fig. 5b shows the comparison between lower levels Wi and Si. An unpaired t-test showed no significant difference between Wi and Si when data were pooled across speaker and consonant type. However, in three-factor ANOVA with factors prosodic position, consonant type and speaker, a significant interaction was found for each combination of factors (all $p < 0.0001$). Interestingly enough, while Wi is associated with a greater contact than is Si for /n, t/, the reverse is true for /t^h, t*/, as shown in Fig. 7a. As was the case in the comparisons among higher levels, the degree of variation is greater in the front region than in the back region. While a significant difference in the front region was maintained for each consonant except /t/, no significant difference in the back region was found for any consonant.

There is also some inter-speaker variation, as shown in Fig. 7b. Results of detailed *post hoc* comparisons are given in Table IIIb. All three speakers distinguish Wi and Si for /n/ at $p < 0.05$, with the linguopalatal contact being greater Word-initially than Syllable-initially (or Word-medially), in each region. For /t/, only Speaker THC showed a significant difference. On the other hand, for /t^h/, a significant difference was found for Speakers NHL and JYY, with a greater contact for Si than Wi. Finally, for /t*/, Speaker NHL shows a significant difference between Wi and Si with greater contact Word-initially. The reverse is true for the other two speakers, who produce /t*/ with greater contact Syllable-initially (Word-medially) than Word-initially.

In sum, the data show that, overall, all consonants are produced with greater linguopalatal contact in higher domain-initial positions, especially for higher level comparisons from Ui to Wi, with just small interspeaker and interconsonantal differences. For lower level comparisons, all speakers make distinctions between Wi and Si for /n/, but not always for other consonants. For /n, t/, Wi tends to have a greater contact than Si, while for /th, t*/, greater contact was found for Si (Word-medial).

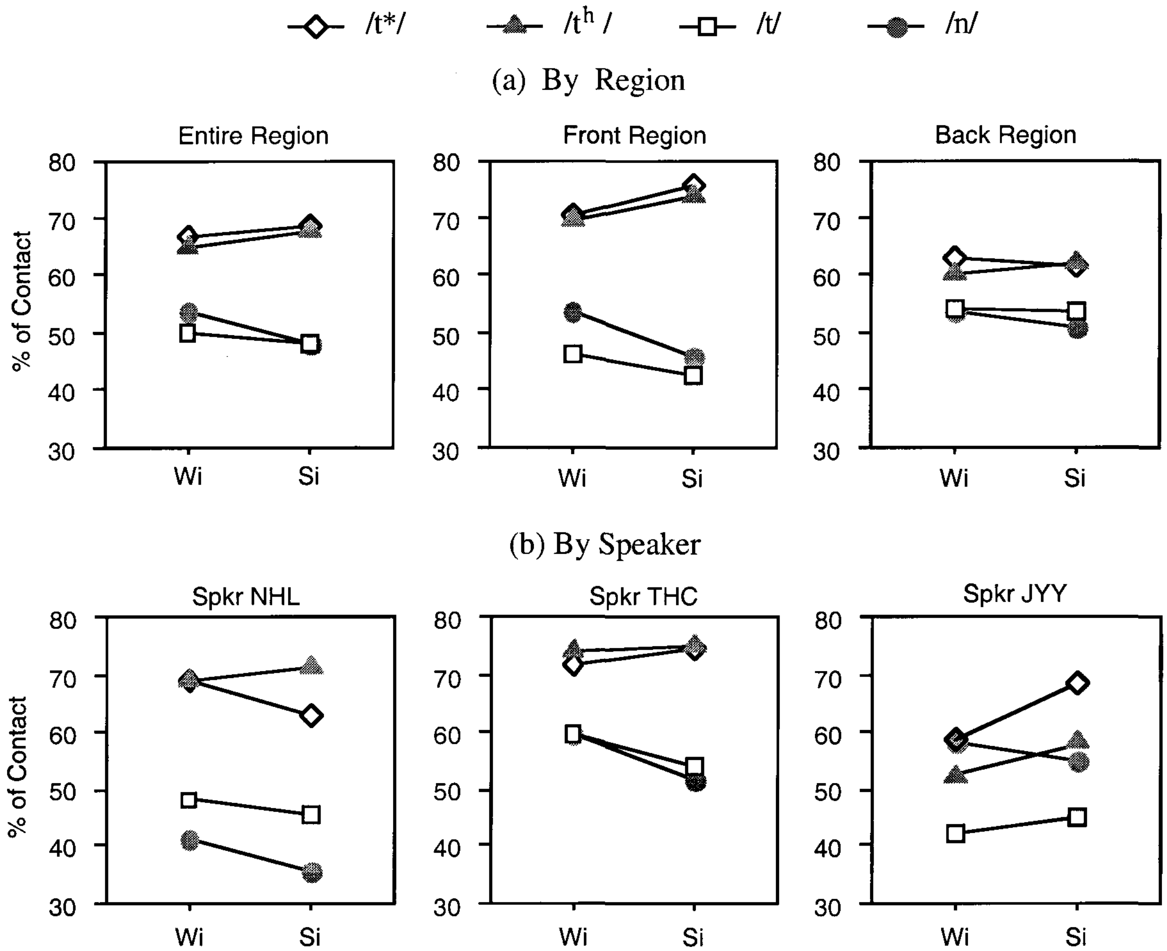


Figure 7. Contact difference between Wi and Si. (a) for each consonant and (b) for each speaker.

TABLE III. Results of statistical comparisons for linguopalatal contact, seal duration, gesture duration. Fisher's PLSD *post hoc* comparison of percent of linguopalatal contact for higher levels and T-tests for lower levels. ('>' (or '<') refers to $p < 0.05$ and '=' refers to $p > 0.05$.)

(a) Higher Level Comparisons

Spkr NHL	Maximum Contact	Seal Duration
/n/	Ui > IPi > APi > Wi	Ui > IPi > APi = Wi
/t/	Ui > IPi > APi > Wi	Ui > IPi > APi > Wi
/t^h/	Ui = IPi > APi > Wi	Ui > IPi > APi > Wi
/t*/	Ui = IPi > APi > Wi	Ui > IPi > APi > Wi
Spkr THC	**Maximum Contact**	**Seal Duration**
/n/	Ui > IPi > APi > Wi	Ui > IPi > APi > Wi
/t/	Ui > IPi > APi > Wi	Ui > IPi > APi = Wi
/t^h/	Ui > IPi > APi > Wi	Ui > IPi > APi > Wi
/t*/	Ui > IPi > APi > Wi	Ui > IPi > APi > Wi
Spkr JYY	**Maximum Contact**	**Seal Duration**
/n/	Ui > IPi > APi = Wi	Ui > IPi > APi = Wi
/t/	Ui > IPi > APi < Wi	Ui > IPi > APi = Wi
/t^h/	Ui > IPi > APi > Wi	Ui > IPi > APi > Wi
/t*/	Ui > IPi > APi = Wi	Ui > IPi > APi = Wi

(b) Lower Level Comparisons

Spkr NHL	Maximum Contact	Seal Duration
/n/	Wi > Si	Wi > Si
/t/	Wi = Si	Wi = Si
/t^h/	Wi < Si	Wi < Si
/t*/	Wi > Si	Wi < Si
Spkr THC	**Maximum Contact**	**Seal Duration**
/n/	Wi > Si	Wi > Si
/t/	Wi > Si	Wi > Si
/t^h/	Wi = Si	Wi < Si
/t*/	Wi < Si	Wi < Si
Spkr JYY	**Maximum Contact**	**Seal Duration**
/n/	Wi > Si	Wi = Si
/t/	Wi = Si	Wi < Si
/t^h/	Wi < Si	Wi < Si
/t*/	Wi < Si	Wi < Si

3.1.2. *Seal Duration*

For all speakers, seal duration is significantly different depending on the prosodic position (when pooled across consonants) at $p < 0.0001$. The variation across prosodic position is shown in Fig. 8a. The common result obtained from *post hoc* comparisons is that speakers make at least a three-way distinction between Ui and Wi, as in the case of linguopalatal contact. Another interesting observation is that unlike for linguopalatal contact, speakers always distinguish between Ui and IPi by seal duration, thus making the three-way distinction Ui - IPi - APi all the time. So, the only difference across speakers and consonants is whether distinctions between APi and Wi are made. The detailed statistical results appear in Table IIIa.

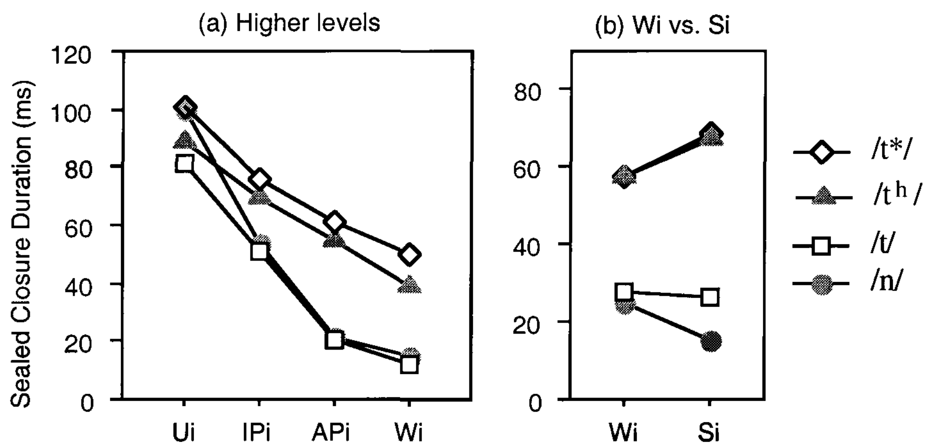


Figure 8. Seal duration as a function of prosodic position and consonant type. Data pooled across speakers.

For the lower level comparisons, as shown in Fig. 8b, Wi /t, n/ tend to have longer seal duration than Si (Word-medial) ones. On the other hand, Si /t^h, t*/ have longer durations than Wi ones, the same pattern as in the case of linguopalatal contact. However, there is interspeaker difference, as shown in Table IIIb. Most notably, Speaker JYY did not make a significant distinction for /n/. This speaker also produced /t/ with longer duration for Si than Wi, which is opposite to the other two speakers. Also, Speaker NHL made /t*/ longer when Wi than Si, the reverse of the other two speakers. But, for /t^h/, all speakers showed a significantly longer seal duration for Si than Wi.

3.1.3. *Skewness*

Recall that skewness here is a measure of the shape of the contact profile when at least 20% of the electrodes are contacted. In general, the skewness value is smaller in higher domain-initial positions, as shown in Fig. 9. This direction is significantly maintained for all consonants. Most of the cases are associated with negative skewness. As Byrd *et al.* (1995) suggested, this direction of skewness shows that the consonant closing is made relatively slowly - here, in higher domain-initial positions compared to lower ones. This direction of skewness also indicates that the peak of the contact is formed relatively later in higher domains.

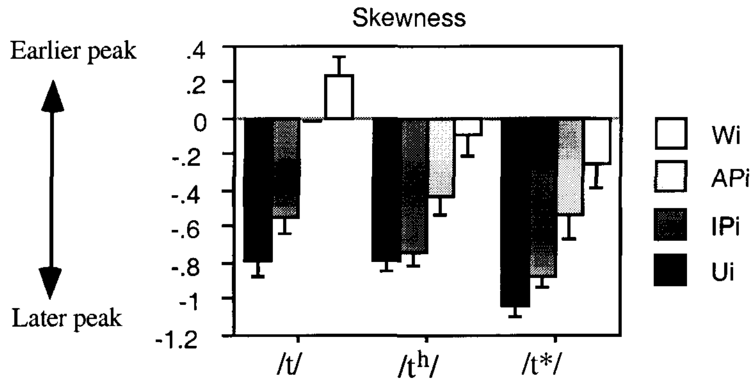


Figure 9. Skewness of linguopalatal contact profile with data pooled across speakers. Error bars = standard deviations.

3.2. *Acoustic parameters*

3.2.1. *Acoustic Closure Duration*

Fig. 10 shows the acoustic closure duration for the oral stops which are initial in an IP, AP, or Word. (Acoustic closure duration could not be measured for stops in Utterance-initial position due to the substantial silent pause.) These data can be compared to the measurements of articulatory seal duration, shown above in Fig. 8 (section 3.1.2.). In general, acoustic closure duration patterns similarly to articulatory seal duration, with a high correlation (R^2 = 0.711, R = 0.843 across speakers and consonants) between the two measures. However, the acoustic duration was generally longer than the seal duration. There are at least two reasons for this difference. First, there were quite a few tokens with no visible seal in the EPG record — i.e., 0 msec seal duration — which nonetheless showed a closure interval in the acoustic signal — i.e. a closure duration greater than 0 msec. (It must be that in these tokens, whatever air exited the vocal tract did not produce any appreciable acoustic noise.) Second, since seal duration was counted as the number of EPG frames showing the seal, any time between the last such frame and the actual release was excluded.

As shown in Fig. 10, two speakers, NHL and THC, always make a three-way distinction in closure duration from IPi to Wi; Speaker JYY makes a two-way distinction with no difference between APi and Wi for /t*/. Speaker JYY also shows the opposite trend between APi and Wi for /t/, the same pattern that was observed for linguopalatal contact. For comparisons between Wi and Si for /t^h, t*/ (figures not provided), each speaker makes a significant distinction by making Syllable-initial stops longer than Word-initial ones, as was found in the case of seal duration. Closure duration for /t/ was not measured in these positions due to the difficulty caused its frequent pronunciation as an approximant , with no acoustic discontinuity, in which case closure duration is 0 ms.

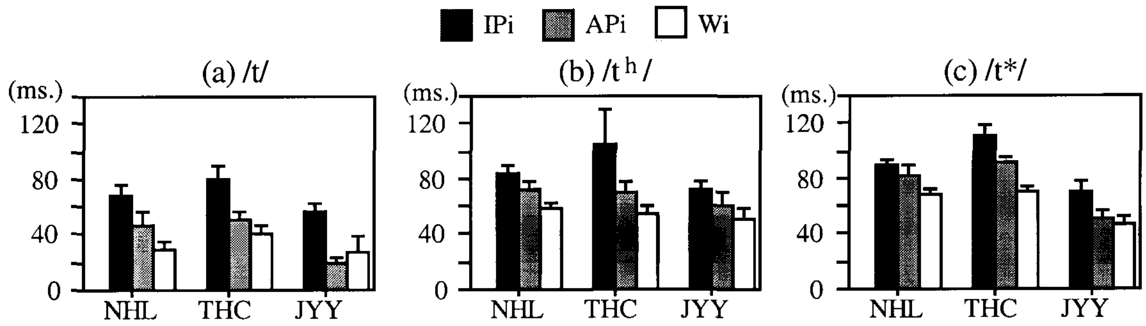


Figure 10. Variation in acoustic closure duration (ms.) for oral stops /t, t^h, t*/. Error bars = standard deviations.

3.2.2. *Voicing During Acoustic Closure*

The duration of voicing during the acoustic closure was found to be significantly influenced by the prosodic position ($F(2, 176) = 18.929, p < 0.0001$ for /t/, $F(2, 176) = 12.469, p < 0.0001$ for /t^h/, and $F(2, 176) = 19.384, p < 0.0001$ for /t*/.) As summarized in Fig. 11, the duration of voicing during the closure is, in general, longer for Wi, and shorter for IPi. Pairwise *post hoc* comparisons show that a three-way distinction was made for /t/ at $p < 0.05$ while a two-way distinction was made for /t^h/ (i.e., IP < (APi = Wi)) and /t*/ (i.e., (IP = APi) < Wi)).

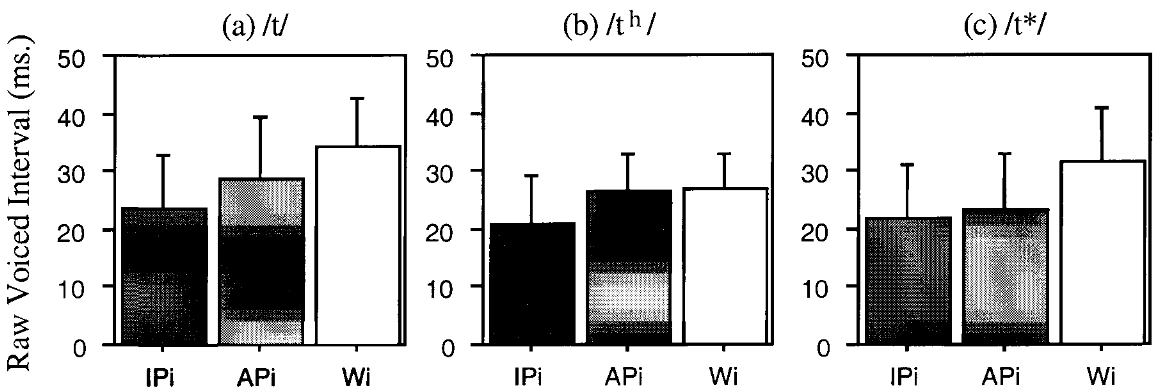


Figure 11. Voicing (in ms) during the acoustic stop closure for each oral stop. Data pooled across speakers. Error bars = standard deviations.

The effect of prosodic position on stop voicing is even clearer when we consider closure voicing as a percentage of the entire stop closure duration, as shown in Fig. 12. Since total closure is longer while closure voicing is shorter in higher prosodic positions, %voicing is much less. Pairwise *post hoc* comparisons show that all speakers make a three-way distinction from IPi through Wi at $p < 0.001$ or less, with proportionately less voicing in higher positions. (Recall that measurements for Ui were not made because of its substantial silent pause.)

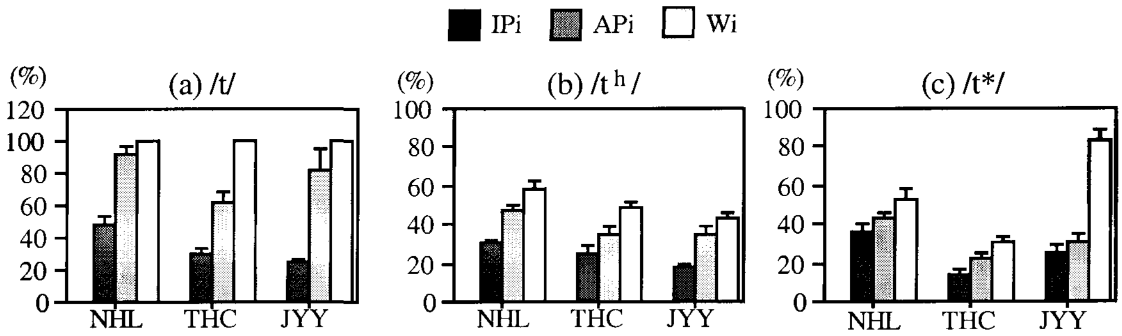


Figure 12. %voicing during the acoustic stop closure for oral stops /t, t^h, t*/. Data pooled across speakers. Error bars = standard deviations.

3.2.3. *VOT*

In general, as shown in Fig. 13a, VOT is longer when a stop occurs in a higher domain-initial position, especially for /t, t^h/. VOT is not shown for the tense stop /t*/, as systematic differences were not found (VOT is always quite short and there is little variation). One-factor ANOVA with data pooled across speakers shows that there is a significant effect of prosodic position on VOTs for both /t, t^h/ (for /t/, $F(3, 232) = 67.824$, $p < 0.0001$; for /t^h/, $F(3,232) = 123.948$, $p < 0.0001$). Pairwise *post hoc* comparisons calculated separately for each speaker confirm that for /t/, all speakers make a four-way distinction (Ui > IPi > APi >Wi); for /t^h/, Speaker THC makes a four-way distinction while the other two speakers make a three-way distinction with no difference either between APi and Wi (for Speaker JYY) or between IPi and APi (for Speaker NHL).

In the comparison between Wi and Si for /t^h/, Wi has a longer VOT than Si ($t = 5.201$, $p < 0.0001$), as shown in Fig. 13b. (Note that VOTs for /t/ for Wi and Si were not plotted due to the intervocalic voicing throughout the closure in most of the tokens, which makes measured VOT zero.) However, t-tests done separately for each speaker show that a significant difference was made by Speakers NHL and JYY, but not by Speaker THC.

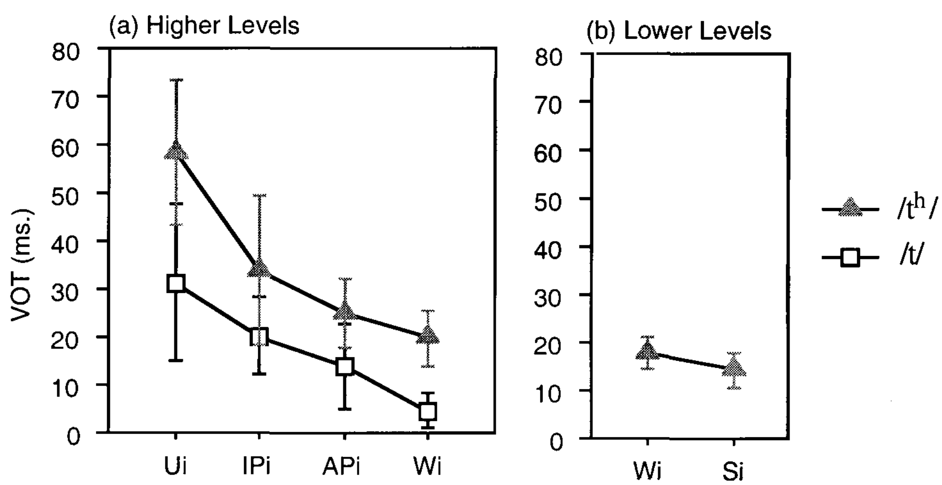


Figure 13. Variations in VOT. (a) higher levels and (b) lower levels. Error bars = standard deviations.

3.2.4. *Total Voiceless Interval*

Fig. 14a shows the acoustic Total Voiceless Interval, which combines voiceless closure duration and VOT, thus indexing the duration of glottal opening. Only the aspirated stop /t^h/ was examined since only it is produced with a fair amount of glottal opening during the closure. As noted above, the voiceless closure duration for Ui could not be measured because it could not be distinguished from any silent pause. However, a rough gauge of this comparison can be observed from Fig. 14b, in which voiceless closure duration is replaced by seal duration. All speakers make clear three-way (in Total Voiceless Interval) or four-way (in seal duration plus VOT) distinctions at $p < 0.0001$. Thus this measure of the duration of glottal opening more consistently distinguishes the prosodic domains than does either VOT or closure duration alone.

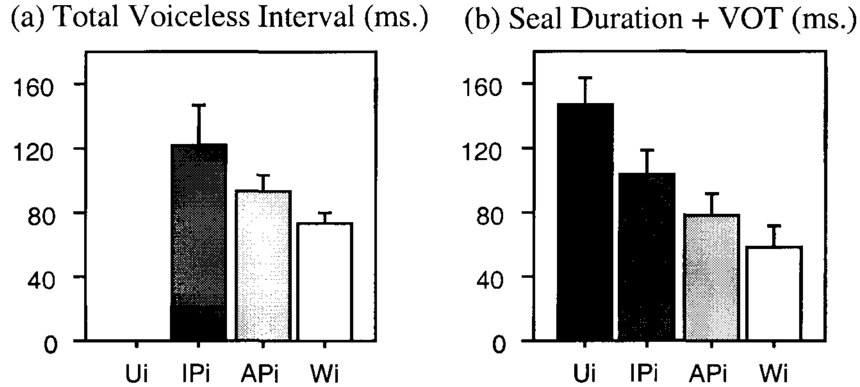


Figure 14. Total voiceless interval (ms.) for aspirated stop. (a) is voiceless closure duration plus VOT and (b) is seal duration plus VOT. Error bars = standard deviations.

3.2.5. *RMS Burst Energy*

This measurement was not made for the Word-initial lenis stop /t/ due to the presence of voicing throughout the consonant. One-factor ANOVA with data pooled across both speakers and the other three consonants shows a significant effect of position on RMS burst energy ($F(3, 643) = 7.19$, $p < 0.0001$) with a tendency for burst energy to be smaller in stops in a higher position. However, as can be seen from Fig. 15, there are significant variations across speaker and consonant type, and the directions of difference are inconsistent. Out of nine one-factor ANOVAs (3 speakers x 3 stops), three cases confirmed that there is no significant effect of position on RMS burst energy; Speaker NHL shows no difference for /t/, and Speaker THC for /t^h, t*/. Results of *post hoc* comparisons also show a great amount of interspeaker and interconsonant variations along with the effect of prosodic position. For example, for /t*/, while Speaker NHL shows progressively increasing burst energy from Ui to Wi, Speaker THC shows virtually no difference at all. Considering only the higher three domains, Ui, IPi, and APi, we found a tendency for burst energy to increase from Ui to APi only for /t^h/ by Speakers NHL and JYY and for /t*/ by all speakers. However, this progressively increasing pattern is broken with the inclusion of Wi.

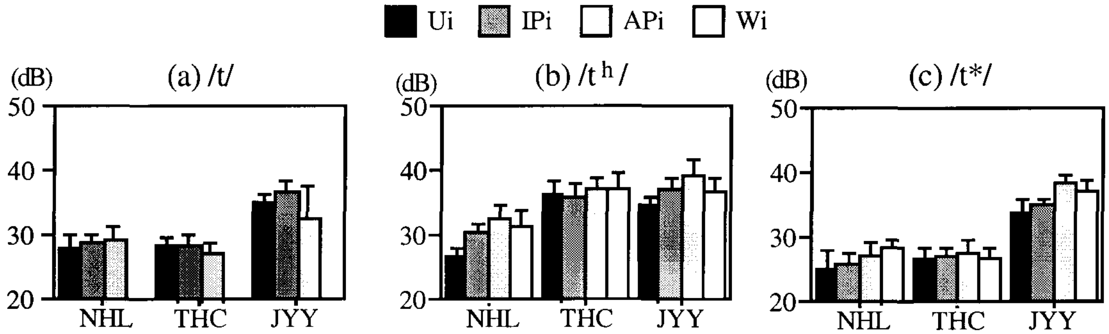


Figure 15. RMS burst energy (dB) for each oral stop. Error bars = standard deviations.

3.2.6. *Nasal Duration*

This is the acoustic measure of consonant duration for the nasal test consonant /n/. There is a significant effect of prosodic position on nasal duration ($p < 0.0001$) for all speakers. As can be seen from Fig. 16a, Speaker NHL makes a four-way distinction in the order of IPi, Ui, APi, and Wi. For the other two speakers, /n/ duration is shortest for Ui. (Shortest nasal duration for Ui was also found in Fougeron & Keating (1997).) Excluding Ui, two speakers (NHL and THC) make three-way distinctions with the nasal duration in the decreasing order of IPi, APi, and Wi while there is no difference at all for Speaker JYY. For comparisons between Wi and Si, Wi nasals are longer than Si nasals. Significant differences were found for NHL and THC ($p <. 0001$) but not for JYY with the same direction still maintained. Thus, while there is quite a significant effect of prosodic position on the acoustic duration of /n/, the effect does not seem to be as strong as for oral stops.

3.2.7. *Nasal Energy Valley*

This measure reflects the minimum nasal airflow during a nasal consonant. Recall that nasal energy for Ui was not measured due to its asymmetrical timecourse after the preceding silent pause. In general, nasal energy is lower in higher domain-initial positions, as shown in Fig. 16b. What is true for all speakers is that nasal energy is lowest for IPi. Speaker JYY makes a three-way distinction (IPi < APi < Wi) while the other two make a two-way distinction (IPi < (APi = Wi)). For Wi vs. Si, all speakers have smaller nasal energy for Wi compared to Si, but the difference was not significant for JYY.

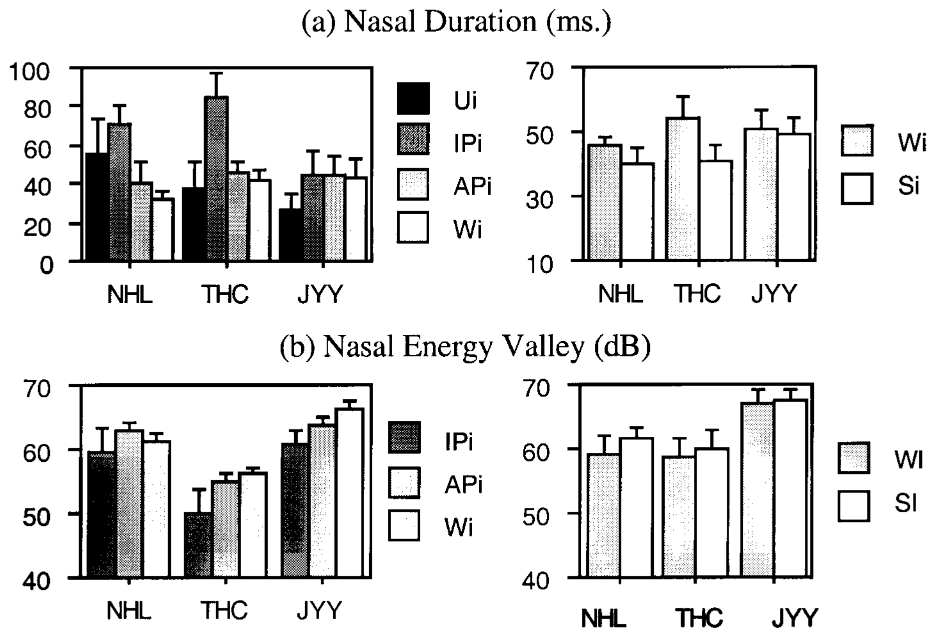


Figure 16. (a) nasal duration and (b) nasal energy valley. Error bars = standard deviations.

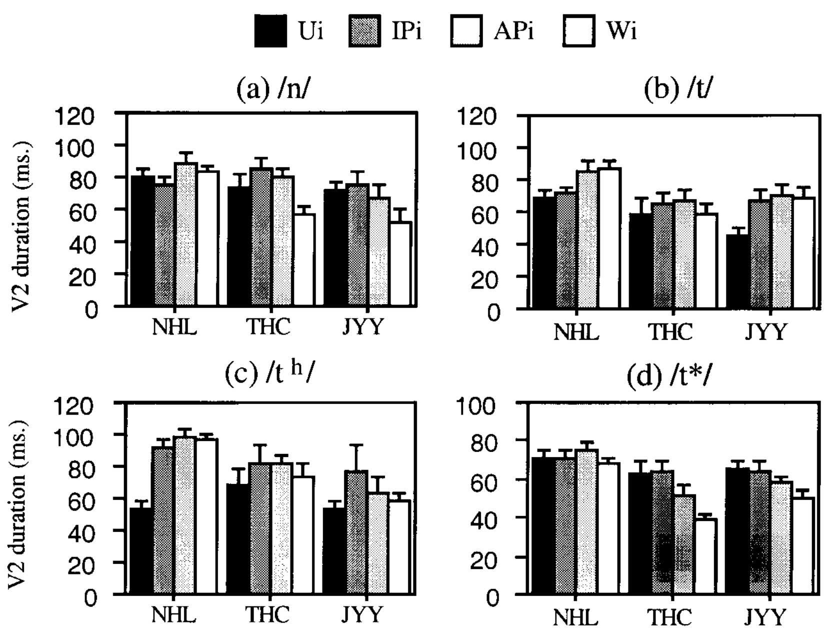


Figure 17. Vowel duration in domain-initial CV. Error bars = standard deviations.

3.2.8. *Duration of V in domain-initial CV* (V2)

This is a measure of post-boundary vowel lengthening (note however that the segment immediately after the boundary is the initial C; V2 follows that C). One-factor ANOVAs show that the factor Prosodic Position significantly influences V2 duration at $p < 0.0001$ for each speaker and consonant context, except for one case where the significance is maintainted at $p < 0.05$. However, the direction varies across speakers and consonant contexts, as can be seen in Fig. 17. For example, after /t/, there is a trend for V2 to be shorter in a higher position, especially true when excluding Wi for Speakers THC and JYY. For /t^h/, the same trend can be observed for Speakers NHL and THC. For these two stops, what is true for all speakers is that vowel duration is shortest for Ui compared to all other position. In contrast, for /t*/, there is no significant difference in V2 duration between Ui and IPi for all speakers (confirmed in pairwise *post hoc* comparisons). However, for two speakers (THC and JYY), V2 after /t*/ is generally longer in a higher domain. A similar pattern is observed for /n/ for these two speakers. That is, in cases where VOT does not vary with prosodic position (/t*, n/), V2 is longer, but where VOT does vary with prosodic position, V2 tends to be shorter. Recall that V2 duration is the duration of only the voiced part of V2, excluding the Voice Onset interval. Thus there appears to be a compensatory relation between these two measures. This lack of any overall consistent pattern accords with results of Jun and Lee (1998), who found that in Korean, vowels are not lengthened under focus. That is, Korean vowels seem to show less durational variation than vowels in some other languages.

3.2.9. *Duration of vowel in domain-final position (V1)*

Results for V1 are rather different than for V2. V1 duration is a measure of domain-final lengthening. (Pause duration is another such measure, but not included here.) One-factor ANOVAs separately by consonant show that there is a significant effect of prosodic position on domain-final vowel duration. As can be seen in Fig. 18, the common result from pairwise *post hoc* comparisons is that across consonants the domain-final vowel is longest for either U-final or IP-final, intermediate for AP-final and shortest for W-final. 12 separate *post hoc* comparisons show that each speaker makes at least a three-way distinction among U-final/IP-final, AP-final, and W-final. There is a great and significant difference between U-final/IP-final and AP-final/W-final; and a relatively small but significant difference between AP-final and W-final. In general, the higher the prosodic position, the longer the domain-final vowel.

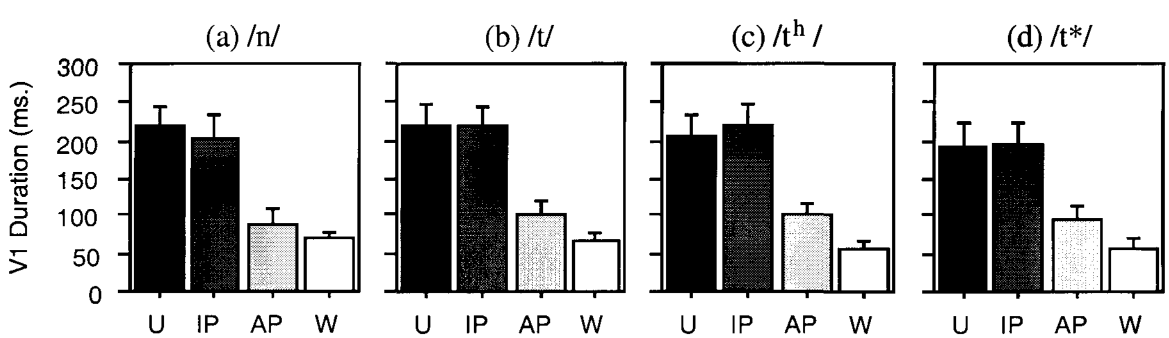


Figure 18. Domain-final vowel duration for higher level comparisons. Each panel shows data for one consonant after the boundary. Error bars = standard deviations.

The difference between AP-final and Word-final vowels, though statistically significant, is subtle. The histogram in Fig. 19a shows that vowel durations divide into two quite distinct distributions, such that IP-final vowels are categorically longer than AP-final vowels -- this is the difference also revealed in Jun (1993; 1995) -- while AP-final vowels are only gradiently longer than Word-final ones.

Fig. 19b shows a histogram of the %contact data. It can be seen that no obvious categories present themselves in this figure, though in section 3.1.1 we saw that the contact in the different prosodic domains is reliably different. From this figure, it seems clear that %contact varies gradiently. Thus, preboundary lengthening may be at least somewhat categorical, but post-boundary increases in linguopalatal contact appear to be continuous.

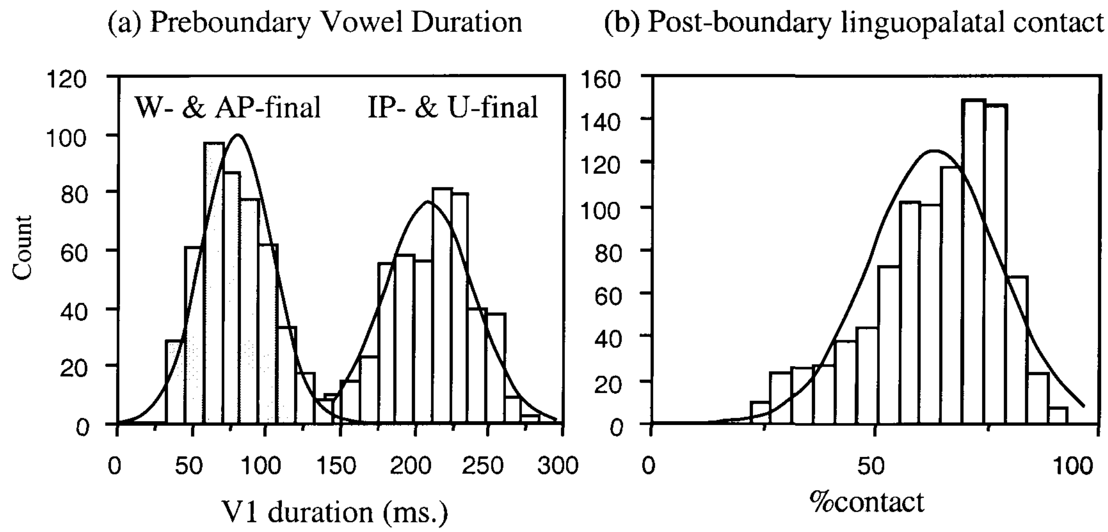


Figure 19. Distribution of preboundary vowel duration measurements (a) and post-boundary linguopalatal %contact measurements (b).

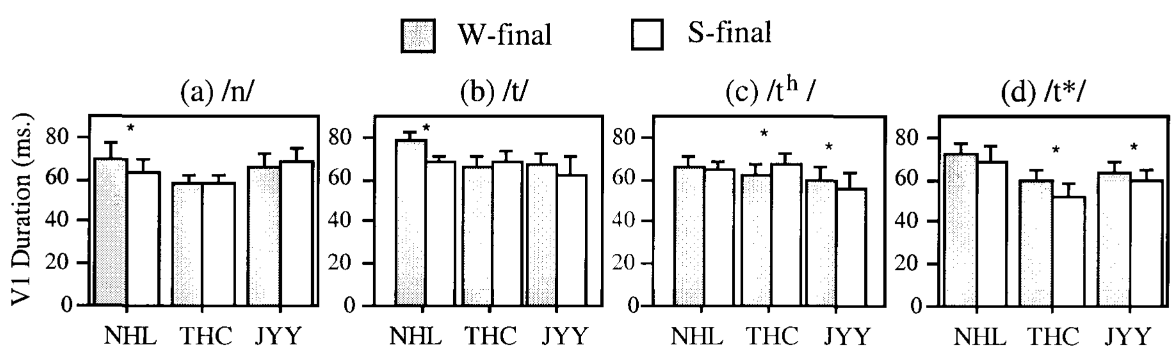


Figure 20. Domain-final vowel duration for lower level comparisons. Pairs with significant differences ($p < 0.05$ in T-tests) are marked with '*'. Each panel shows data for one consonant after the boundary. Error bars = standard deviations.

In the lower level comparisons between W-final and S-final (W-medial) vowels, patterns vary with speaker and consonant, as shown in Fig. 20. When data were pooled across following consonant contexts, only one speaker, NHL, produces W-final vowels significantly longer than S-final (W-medial) ($t = 5.069, p < 0.0001$). When each following consonant context was considered separately, two speakers, THC and JYY, make a significant distinction for /t*/ with W-final being longer than S-final, with the same direction shown by Speaker NHL. For /n,t/, one speaker, NHL, makes a distinction; and for /t^h/, Speaker THC makes a distinction but in the opposite direction. Thus the W-internal differences are not at all consistent in the way the higher level ones are.

3.3. *Variation as a function of consonant type*

The four different consonants studied here are produced with different amounts of linguopalatal contact, in decreasing order /t*, t^h, t, n/. This can be seen in Fig. 6, where the consonants appear to divide into two groups. In order to examine the consonantal effect and its interaction with prosodic position, two-factor ANOVAs with factors of Consonant Type and Prosodic Position were generated for each speaker. First, there is a significant effect of consonant type on linguopalatal contact for each speaker ($F(3, 304) = 520.499, p < 0.0001$ for Speaker NHL; $F(3, 304) = 207.635$, $p < 0.0001$ for Speaker THC; $F(3, 300) = 159.771$, $p < 0.0001$ for Speaker JYY). *Post hoc* pairwise comparisons confirm that all consonants are differentiated by linguopalatal contact ($n < t < t^h < t^*$) by two speakers, NHL and THC but a two-way distinction was made ($n = t < t^h = t^*$) by Speaker JYY. Thus, for all speakers, tense and aspirated stops have greater linguopalatal contact than do nasal and plain stops. Second, there is also a significant interaction between consonant type and prosodic position for each speaker ($F(9, 304) = 70.574$, $p < 0.0001$ for Speaker NHL; $F(9, 304) = 29.022$, $p < 0.0001$ for Speaker THC; $F(9, 300) = 20.305$, $p < 0.0001$ for Speaker JYY).

Therefore, in order to see if this overall pattern is maintained throughout all prosodic positions, 12 separate one-factor ANOVAs (3 speakers x 4 prosodic positions) with the independent variable Consonant Type were conducted. A summary of results is given in Table IV. In general, fewer distinctions are made between consonant types in higher positions (i.e., Ui and IPi) than in lower positions (i.e., APi and Wi). The fact that the higher the position, the smaller the F-values, suggests that the effect of consonant type is greater in lower domain-initial positions. It is also noteworthy that at least a two-way distinction, separating /n, t/ and /t^h, t*/ into different groups, is made throughout all prosodic positions, except Ui for Speakers NHL and THC.

TABLE IV. A summary of 12 one-factor ANOVAs with factors Linguopalatal Contact (dependent) and Consonant Type (independent). Fisher's PLSD *post hoc* comparisons for percent of linguopalatal contact are also reported. ('$>$' (or '$<$') refers to $p < 0.05$ and '$=$' refers to $p > 0.05$.)

Spkrs	Position	*F*-Values	*P*-values	*Post hoc* comparisons
NHL	Ui	$F(3, 76) = 3.776$	$p = 0.0139$	$n = t = t^h$; $t = t^h < t^*$; $n = t^*$
	IPi	$F(3, 76) = 138.098$	$p < 0.0001$	$n < t < t^h = t^*$
	APi	$F(3, 76) = 138.868$	$p < 0.0001$	$n < t < t^h < t^*$
	Wi	$F(3, 76) = 476.844$	$p < 0.0001$	$n < t < t^h = t^*$
THC	Ui	$F(3, 76) = 12.375$	$p < 0.0001$	$n = t = t^h < t^*$
	IPi	$F(3, 76) = 26.437$	$p < 0.0001$	$n = t = t^h < t^*$; $n < t^h$
	APi	$F(3, 76) = 98.806$	$p < 0.0001$	$n < t < t^h < t^*$
	Wi	$F(3, 76) = 147.632$	$p < 0.0001$	$n < t < t^h < t^*$
JYY	Ui	$F(3, 76) = 19.902$	$p < 0.0001$	$n < (t = t^*) < t^h$
	IPi	$F(3, 76) = 34.078$	$p < 0.0001$	$n = t < t^h = t^*$
	APi	$F(3, 76) = 93.242$	$p < 0.0001$	$n > t$; $n < t^h$; $t < t^h = t^*$
	Wi	$F(3, 76) = 49.398$	$p < 0.0001$	$n = t < t^h = t^*$

4. Discussion

4.1. *Domain-initial strengthening as cumulative*

The articulatory and acoustic parameters examined in the current study strongly support the hypothesis that there is domain-initial strengthening in Korean, and that it is cumulative. Recall that in our corpus, a Syllable-initial consonant was Word-medial; a Word-initial consonant, AP-medial; an AP-initial consonant, IP-medial; and finally an IP-initial consonant, U-medial. Almost all domain-initial strengthening found in a higher position was greater than any strengthening found in a lower position. For example, the articulatory magnitude as manifested in linguopalatal contact becomes progressively greater as a test consonant moves up in the prosodic hierarchy. This progressively increasing trend was also found in domain-initial lengthening of seal duration, VOT, and Total Voiceless Interval. Similarly, a progressively decreasing trend was found in nasal energy and RMS burst energy, though the latter was not so consistently cumulative. The systematic exception to this overall cumulative pattern is discussed in section 4.5 below.

4.2. *Nature of Articulatory Strengthening*

Fougeron & Keating (1997) presented several possible mechanisms for domain-initial strengthening. We will discuss two of these in light of the present study.

First, Fougeron & Keating suggested that articulatory strengthening may be due to longer durations in stronger positions. Longer durations would result in less articulatory undershoot, so that the articulatory target of a consonant would be more closely approximated (cf. Lindblom, 1963; Moon & Lindblom, 1994). Although Fougeron and Keating did not find a strong relation between these variables in their English data, this possibility can be tested for Korean by examining the correlation between linguopalatal contact and seal duration, shown in Fig. 21. Though the linear relation between these two variables is strong for all three speakers, it can be seen in the graphs that the relation is not in fact linear, but instead asymptotes at the largest values of contact. Therefore the regressions shown in Fig. 21 are polynomial; the relations here (R^2 = 0.91 for Speaker NHL; 0.85 for Speaker THC; and 0.77 for Speaker JYY) are slightly stronger than linear regressions (not shown). A strikingly high proportion of the variance is accounted for by this single variable, especially for the first two subjects. Overall, then, we can reliably predict the amount of linguopalatal contact from the seal duration: the longer the seal duration, the greater the linguopalatal contact. However, at the longest durations, contact asymptotes. In particular, for durations above about 80 msec, peak contact remains similar (and large). This relation suggests that about 80 ms is enough time to reach a contact target, and there is some articulatory undershoot for durations shorter than this.

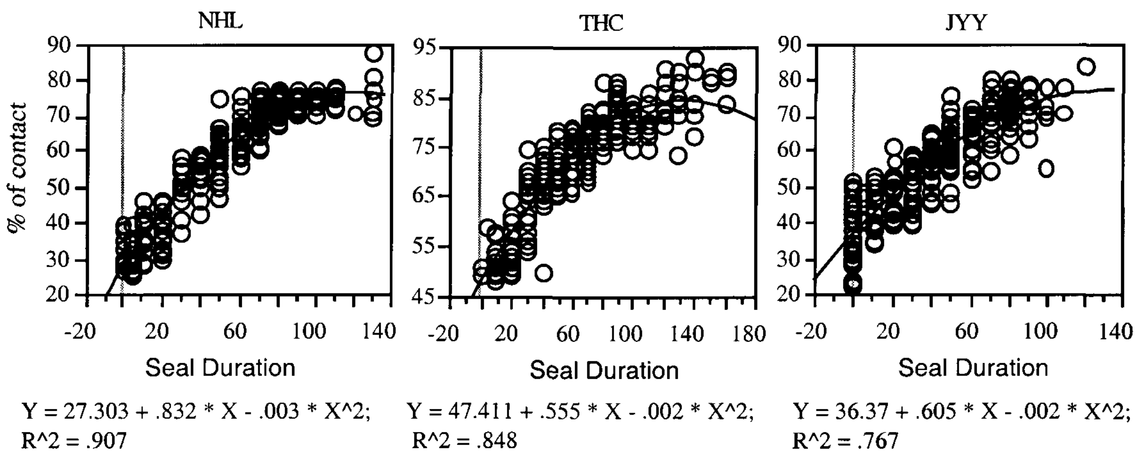


Figure 21. Polynomial regression plots with linguoplatal contact against seal duration. Note that y- and x-axes are scaled for each subject's data. Data are pooled across consonants.

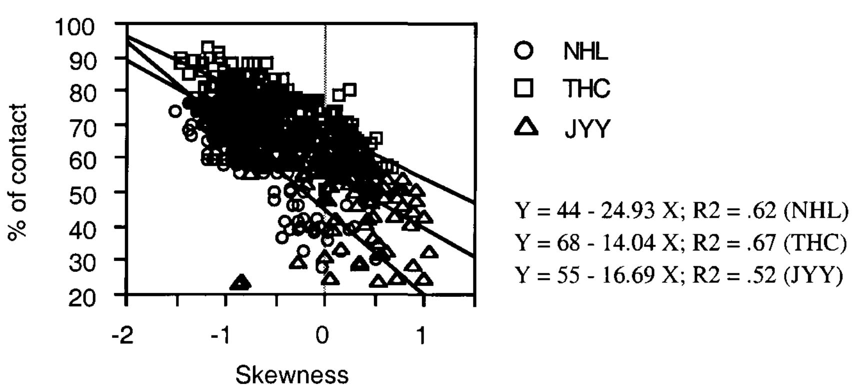


Figure 22. Simple regression plot with linguopalatal contact again skewness.

Korean is unusual among the languages studied to date in having such a close relation between lengthening and strengthening. Byrd et al. (1996) found lengthening, but essentially no strengthening, of Tamil /m/; as already noted, Fougeron and Keating (1997) found a weak correlation in English; Fougeron (1998) found moderate correlations in French, but weaker than those reported here. Thus there is no *necessary* connection between lengthening and strengthening.

The articulatory undershoot hypothesis is further supported by the skewness of the articulatory trajectory, which varies as a function of prosodic position (see Fig. 9). As can be seen in Fig. 22, there is a fairly strong inverse linear relationship between the linguopalatal contact and the skewness. In addition, there is also a fairly strong inverse relationship between the skewness and the seal duration (not shown; R2 = .67 for Speaker NHL, R2 = .81 for Speaker THC, and R2 = .66 for Speaker JYY), which suggests that the longer the duration, the smaller or more negative the skewness. Fig. 23 shows representative articulatory trajectories with skewness varying as a function of prosodic position. The articulatory trajectories are asymmetrical, being skewed further to the left with longer durations in higher positions. This suggests that the most extreme articulation occurs toward the end of the articulatory trajectory. On the other hand, the articulatory trajectories in lower positions (i.e., APi and Wi) show less linguopalatal contact. This is presumably because the time taken does not allow the articulator to reach its extreme target value, and this results in less skewness.

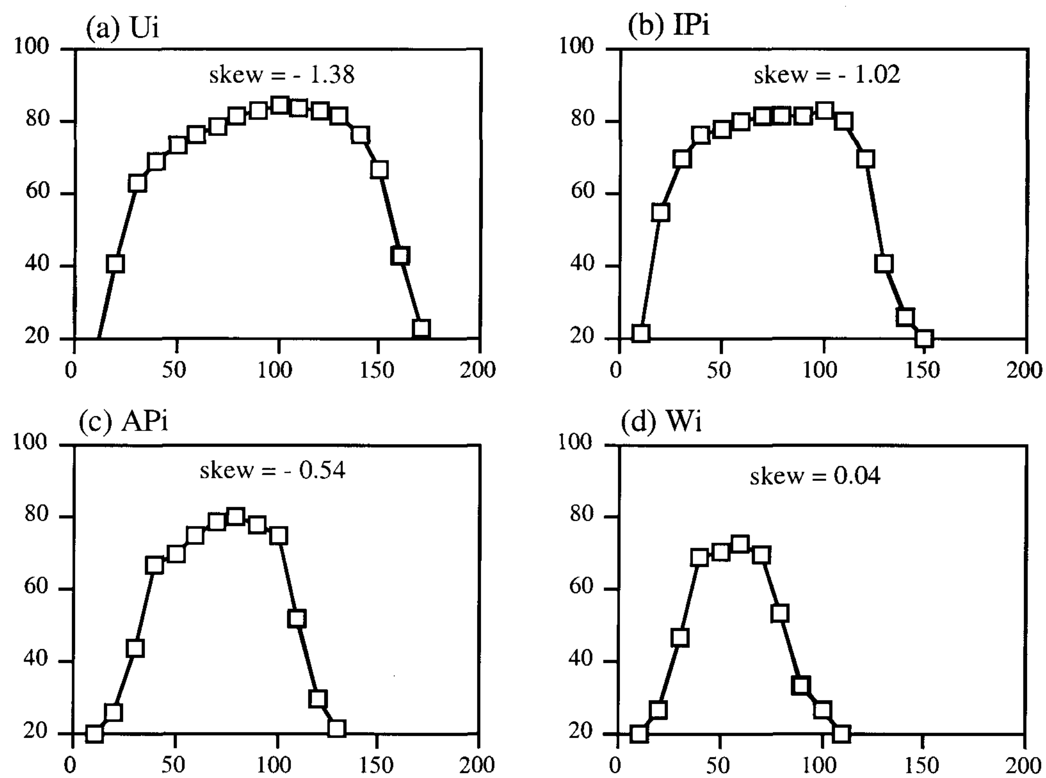


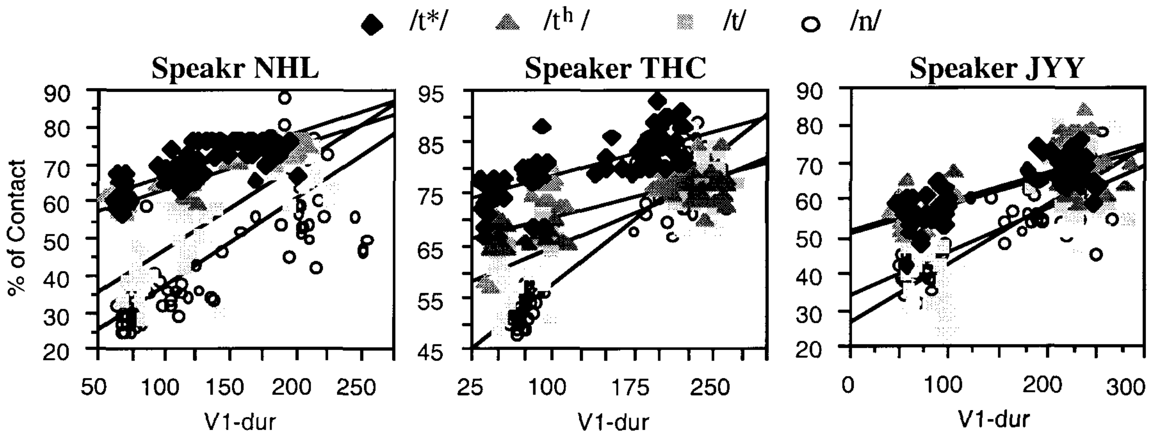
Figure 23. Representative articulatory trajectory with skewness varying as a function of prosodic position. Tokens are from /t*/ produced by Speaker THC.

The skewness pattern in these Korean data is consistent with Byrd (1994)'s findings for English coronal consonants. Byrd compared consonants in Word- (and Syllable-) initial position vs. Word- (and Syllable-) final position. She found that consonants had more positive skew in final position than in initial position. Since her Word/Syllable-initial position, which is prosodically stronger than her Word/Syllable-final position, has less positive skew, our result can be seen as extending this pattern to higher-level strong positions, which have more negative skew.

Second, Fougeron & Keating also suggested that initial strengthening may be attributable to a greater impact of the tongue against the hard palate in consonant formation. This greater impact would arise from higher articulator velocity, which would in turn result from a greater distance to be covered by the articulator (velocity proportional to displacement). The distance would be greater because final lengthening of a (non-high) vowel before the prosodic boundary would result in a more open vowel. In section 3.2.9 it was seen that in our data pre-boundary lengthening is fairly cumulative. However, we do not have contact data for the pre-boundary vowels in this study; this is because all our vowels are open vowels (unlike in Fougeron and Keating's study) and have little or no contact. Thus we do not have V-to-C displacement data, or movement velocity data. Nonetheless, the discussion above of skewness and its correlation with linguopalatal contact indirectly suggests that higher velocity is not the source of the increase in contact. As already shown earlier (Figs. 21, 22), for higher prosodic positions, the peak contact is reached later in the movement at the end of the stop closure. It seems unlikely that velocity differences early in the trajectory could account for contact differences which occur late in the trajectory. However, no firm conclusion can be drawn without the relevant direct measurements, ideally from articulatory movement data.

4.3. *Final Vowel Lengthening and Domain-initial Articulatory Strengthening*

As just noted, domain-final lengthening was found to be progressively greater as its prosodic position moves up in the prosodic hierarchy (see Figs. 18-20). This finding agrees with Jun (1993; 1995) that Korean IP-final vowels are lengthened, but it contradicts Jun's finding that there is no significant AP-final lengthening. One might suspect that the difference in vowel duration between AP-final and Word-final positions observed in our data may be due to the different segmental context associated with the AP-final vowel. All AP-final vowels in our study are preceded by voiced [g] while W-final vowels are preceded by voiceless (geminate) tense [k*], except in the /n/ corpus. If vowels are longer after [g] than after [k*], then the present result could be due to this confound. However, Cho (1996) showed that, in fact, vowels tend to be longer after voiceless tense stops than after voiced plain stops. If the vowel duration were determined by the preceding consonant type, then the results would be the opposite of what we found. Thus, the segmental context in our corpus even reinforces the strength of our result. Furthermore, the same AP-final lengthening was found for /n/, for which the segmental context is exactly the same across the test sentences. Therefore, the present study suggests that final lengthening in Korean may not be limited to IP- or Utterance-final positions but extends to AP-final position, as in French (Hirst & Di Cristo, 1984; Pasdeloup, 1990; Fougeron & Jun, 1998).



Y = 55 + .11 X; R2 = .64 (t*)
Y = 51 + .11 X; R2 = .81 (th)
Y = 23 + .23 X; R2 = .74 (t)
Y= 13 + .24 X; R2 = .61 (n)

Y = 72 + .05 X; R2 = .57 (t*)
Y = 65 + .05 X; R2 = .51 (th)
Y = 56 + .08 X; R2 = .64 (t)
Y = 40 + .16 X; R2 = .89 (n)

Y = 50 + .07 X; R2 = .63 (t*)
Y = 52 + .07 X; R2 = .57 (th)
Y = 26 + .15 X; R2 = .58 (t)
Y = 33 + .11 X; R2 = .72 (n)

Figure 24. Regression with %contact plotted against the duration of the preceding vowel which is domain-final.

To examine if there is a significant linear relationship between the linguopalatal contact of the domain-initial consonant and the duration of the preceding domain-final vowel, linear regression of peak contact against V1 duration is shown in Fig. 24. For each speaker, there is a significant relation between these two variables (R^2 is at least 0.51 or higher for all 12 cases (3 speakers x 4 consonants)). This result suggests that there is a close interaction between the preceding vowel duration and the domain-initial articulatory strengthening. In general, the longer the preceding vowel, the more consonantal contact. However, as noted in the previous section, this relationship is probably not due to velocity differences in the V-to-C movements; and as seen in Fig. 19 above, the distributions of the values are somewhat different.

We saw in the previous section that consonant duration accounts for much of the variance in consonant contact. Now we have seen that preceding vowel duration is also strongly related to consonant contact. These findings lead us to ask whether the dependent variable, %contact, can be better explained by these two independent variables together. As seen in Table VI (the first row of each consonant section), a series of multiple linear regressions reveals that the %contact for a consonant is reliably predictable by these two independent variables together. For Speakers NHL and JYY, these two variables account for almost all the variance in consonant contact. However, multiple regression values are not always greater than those of simple polynomial regressions. As marked by shading in Table VI, only 5 out of 12 comparisons show that preceding vowel duration and seal duration together account for the %contact variance better than does seal duration alone, and just as often the reverse is true. This suggests that preceding vowel duration and seal duration are correlated independently with %contact.

TABLE VI. A summary of multiple regressions of %contact with two independents, preceding vowel (V1) duration and seal duration. Parenthesized values are from simple polynomial regressions with a single independent, seal duration. Cells are shaded when multiple regression values are greater than simple polynomial regression values.

	Speaker NHL		Speaker THC		Speaker JYY	
	R	R^2	R	R^2	R	R^2
/n/	0.964 (0.970	0.929 0.942)	0.965 (0.970	0.930 0.942)	0.913 (0.881	0.833 0.776)
/t/	0.956 (0.961	0.914 0.923)	0.900 (0.961	0.809 0.923)	0.907 (0.917	0.822 0.840)
/th/	0.915 (0.829	0.837 0.687)	0.682 (0.829	0.674 0.687)	0.928 (0.920	0.862 0.846)
/t*/	0.902 (0.855	0.814 0.718)	0.808 (0.853	0.644 0.728)	0.940 (0.945	0.883 0.893)

4.4. *Acoustic cues and their correlation with articulatory strengthening*

The domain-initial articulatory strengthening revealed in linguopalatal contact will be linguistically more significant if it has acoustic or auditory consequences that are transferable to a listener. In the present study, we found that values along the acoustic parameters tested do vary as a function of prosodic position. That is, domain-initial positions are signalled to a significant extent by a number of acoustic correlates such as stop closure duration, VOT, Total Voiceless Interval, %voicing, vowel duration, RMS burst energy, nasal duration, and nasal energy. However, not every cue is directly linked to the strengthening in linguopalatal articulation. For example, the variation in %voicing and Total Voiceless Interval is surely linked to articulatory events at the glottis, while nasal energy and duration should be closely linked with the articulatory events at the velum. In what follows, we will summarize the acoustic results according to their relations to articulatory events.

4.4.1. *Acoustic correlates of the linguopalatal articulatory strengthening*

First, acoustic closure durations (which are roughly equivalent to seal durations for oral stops) are cumulatively longer in higher prosodic positions. As shown in Fig. 21, there is a significant and strong linear relationship between the linguopalatal contact and the seal duration. As discussed earlier, this durational difference can be related to articulatory undershoot, and thus seems to account for the articulatory strengthening. Since lengthening and strengthening are so strongly related in Korean, the variation in duration could be a significant perceptual cue for the articulatory strengthening.

Second, RMS burst energy tends to be smaller in oral stops in higher domain-initial positions for some cases. This result, however, must be taken to be only suggestive because of the inconsistent patterns across speakers and consonants. Although there is some effect of prosodic position on burst energy, the effect is not as strong as other acoustic and articulatory parameters. In fact, there is only a modest negative correlation between maximum linguopalatal contact and

RMS burst energy (R^2 = 0.134, R = 0.368 with data pooled across speakers and consonants). However, the correlation varies with consonant types and is notable only for /t*/ (R^2 = 0.096 for /t/, R^2 = 0.034 for /t^h/, and R^2 = 0.434 for /t*/). Thus, the results indicate that burst energy could be a weak acoustic cue for articulatory strengthening. For /t*/, the significant inverse relationship seems to agree with the prediction that the greater contact would give a longer release duration, and thus less peak burst energy (Stevens, Keyser, & Kawasaki, 1986). In any case, there is no support for the converse hypothesis, that the release of a strengthened articulation is faster and results in a louder burst.

Third, stops /t, t^h/ have longer VOTs in higher domain-initial positions (see Fig. 10). There is a significant linear relationship between linguopalatal contact and VOT as shown in Table V. As suggested by Pierrehumbert & Talkin (1992), and as discussed further below, it is most likely that the glottal opening gesture is strengthened domain-initially compared to domain-medially. However, it is also possible that the differences in linguopalatal contact result in aerodynamic differences at the release, which also affect VOT. If the oral pressure behind the constriction increases proportionally with longer duration — possible if the vocal tract walls remain fairly lax — but the speed of oral opening is constant, then the resulting greater oral pressure will take relatively longer to fall and allow an adequate transglottal pressure for the initiation of vocal fold vibration (Cho & Ladefoged, 1997). A problem with this explanation for our data, however, is that if oral pressure were higher but the release constant, burst energy would be expected to be greater, which it is not. Alternatively, a longer VOT can be associated with a larger contact area (Cho & Ladefoged, 1997; Stevens, 1999), if speed of release is proportional to contact. Cho and Ladefoged, following Stevens, explain that a slower release could be attributable to a Bernoulli effect over the large contact area pulling the articulators together. If the articulators come apart more slowly, a longer time will be needed before an appropriate transglottal pressure is achieved. The reliable linear relationship between linguopalatal contact and VOT appears to support this latter possibility, and the hypothesis of slower releases in higher prosodic positions receives support from studies of English (Byrd, 1994; Byrd & Saltzman, 1998).

TABLE V. Correlation between linguopalatal contact and VOT

	/t/		/t^h/	
NHL	R = 0.782	R^2 = 0.530	R = 0.604	R^2 = 0.365
THC	R = 0.640	R^2 = 0.410	R = 0.560	R^2 = 0.314
JYY	R = 0.711	R^2 = 0.506	R = 0.820	R^2 = 0.672

4.4.2. *Acoustic correlates of other articulatory events*

Total Voiceless Interval for the aspirated stop /t^h/ is cumulatively longer in higher domain-initial positions, marking all distinctions from IPi through Wi. Recall that this measurement serves as an index of the time during which the glottis is open. Thus it seems that the glottis is open longer in higher prosodic positions. Furthermore, the literature suggests that higher prosodic positions involve larger glottal openings. For example, Cooper (1991) found that English voiceless aspirated stops in Word-initial position have a larger glottal opening gesture compared to Word-medial position; and in a fiberscopic study of glottal configurations of Korean obstruents, Jun, Beckman, & Lee (1998) found that the glottal aperture is in general larger when lenis (plain) and aspirated obstruents /t, t^h, s/ occur in AP-initial position than when they occur in AP-medial position. We infer, then, that the longer Total Voiceless Intervals (and VOTs) in our data, especially for the aspirated stop /t^h/, reflect larger, as well as longer, glottal openings. If this is correct, then Jun *et al.*'s finding is extended to higher prosodic positions in Korean.

Second, the interval that is voiced during the stop closure is in general shorter in higher domain-initial positions (e.g., IPi) and shorter in lower domain-initial positions (e.g., Wi). We infer that glottal abduction occurs earlier in these higher positions. It is also likely that the glottal abduction is formed more quickly: Jun *et al.* (1998) found greater magnitude of the glottal opening gesture in AP-initial position, and the velocity of a movement is usually proportional to its magnitude. In addition, this difference in closure voicing may reflect differences in vocal tract tension — voicing into stop closure is associated with less vocal tract tension (Keating 1984). The shorter voicing interval in a higher prosodic position would thus suggest greater vocal tract tension as one of the articulatory properties characterizing prosodically strengthened consonants.

It is interesting to compare the effect of phrasal position, as in the present study, and stress accent on closure voicing. Keating (1984) showed that in English and Swedish, stress on a following vowel *increased* the closure voicing of (Syllable-initial) stop consonants, corresponding to predictions from an aerodynamic model of consonant production. The explanation was that the greater respiratory effort associated with stress caused subglottal pressure to increase more than it caused oral pressure to increase, thus favoring voicing. This difference between (English/Swedish) stress and (Korean) phrasal position suggests that domain-initial strengthening is not the same as stress, at least not in the sense of increased respiratory effort.

Third, acoustic energy for nasal /n/ tends to be lower in higher domain-initial positions (Fig. 19b). Most likely the velum is higher (resulting in less nasal energy) for a higher domain-initial /n/. Our data are like Fougeron (1998) for French, who has proposed that such variation in nasal energy is caused by differences in muscular tension — that by hypothesis, in higher prosodic positions, the levator palatini is more tensed and the velum is higher, regardless of whether the consonant is oral or nasal. She compares this hypothesis favorably with an alternative, that consonants in strong prosodic positions are generally produced with less sonority (e.g. less energy in lower harmonic components).

4.4.3. *Closure voicing and the Korean voicing rule.*

We saw above (Fig. 14) that %voicing during acoustic closure is smaller for stops in higher domain-initial positions. The proportion of closure which is voiced bears on whether a stop will be heard as voiced or voiceless. Thus this measure allows us to look at prosodic conditioning of the intervocalic voicing of the lenis stop /t/. Jun (1993; 1995) claimed that the well-known rule of Korean intervocalic voicing applies to lenis stops anywhere inside an AP, but not AP-initially. She also showed that when consonant duration was regressed against the duration of the following vowel, there was no distinct separation between fully voiced, partially voiced, and voiceless plain stops. She argued that this result showed that voicing is a gradient, phonetic rule. Docherty (1995) took exception to this conclusion. Our data allow us to revisit this issue, since we have data on lenis stop /t/ in a range of prosodic positions.

Fig. 25 shows the distribution of %voicing during closure in each domain-initial position in which it could be determined. (These are the same data as in Fig. 12.) In general, it can be seen that almost all the AP-medial (Wi) stops have 100% closure voicing (right panel), while the IPi stops never do (left panel). Instead, the IPi stops generally have less than 60% closure voicing.

Thus the stops in these two prosodic positions are well-distinguished as "fully voiced" vs. "partially voiced", and with just these cases, the rule could be considered categorical and phonological. In contrast, the APi stops vary continuously from one extreme to the other, though with a clear peak in the distribution at 100% voicing (representing the shorter APi stops). This continuous distribution indicates that voicing is not simply "optional" for APi stops (else there would be a bimodal distribution); there must be some additional factor besides phrasal position alone that determines how much voicing APi stops show. Jun says as much when she states that the lenis stop voicing rule is a "by-product of some other effect of prosodic position on the gestural amplitude and overlapping, thus producing a continuum of voicing (1995:250)." Until that factor is understood, we cannot say whether the gradient voicing of the APi stops reflects a phonetic rule.

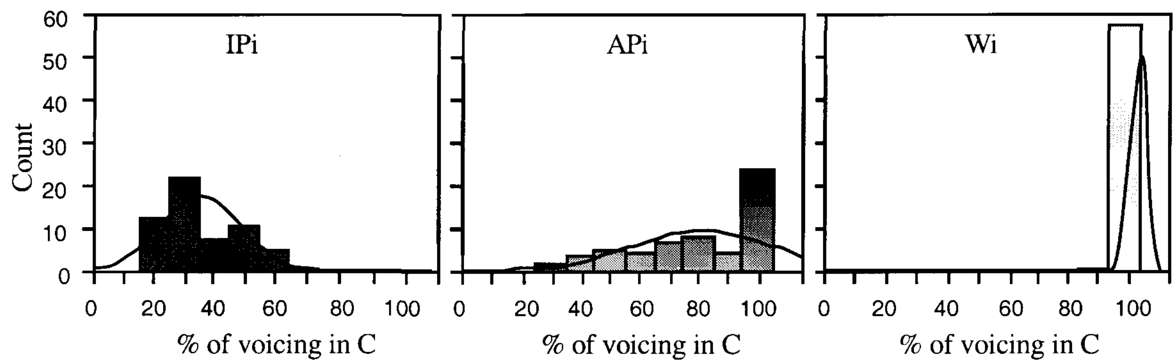


Figure 25. Distribution of %voicing measurements in different prosodic positions.

4.5. *Competition between Word-initial strengthening and Word-medial gemination*

Recall that in Korean, Word-medial tense stops are said to be lengthened (geminated), and perhaps articulated more forcefully, compared to Word-initial tense stops. Under some accounts, tense stops are underlyingly geminates of two plain stops (e.g., Martin, 1982; Yu, 1989; J. Jun, 1994; Han, 1996) but are obligatorily degeminated Word-initially (Han, 1996). In some discussions (e.g., J. Jun, 1994), aspirated stops are also considered to behave like phonological geminates Word-medially. Viewed as medial strengthening, such a process is at odds with the overall trend of cumulative initial strengthening that we have demonstrated in Korean stops. That is, medial strengthening of /t*/ and possibly /t^h/ is a more specific process that competes with the more general process of initial strengthening in Korean. How do our speakers resolve this competition?

In terms of linguopalatal contact, we saw (in Fig. 7 and Table IIIb) that two speakers had more contact for Word-medial than Word-initial /t^h/, and two speakers had more contact for Word-medial than Word-initial /t*/. In terms of <u>seal</u> duration, we saw (in Fig. 8b and Table IIIb) that all three speakers had longer Word-medial than Word-initial /t^h/, two speakers had longer Word-medial than Word-initial /t*/, and one speaker had longer Word-medial than Word-initial /t/. Thus, in our small sample at least, Word-medial lengthening and strengthening of /t^h/ and /t*/ often seem to override any general tendency to Word-initial strengthening of consonants. Still, sometimes speakers do follow the more general tendency, even for /t^h/ and /t*/. Thus the behavior of these stops in these domains is variable and inconsistent. It may be that in their competition, the two processes partially cancel each other out, so that the net medial lengthening and strengthening is small, or even nil.

A further point is that although medial lengthening is the general pattern in our sample, the amount of lengthening is by no means a doubling, as might be expected under "gemination", but is less than 20%. This is somewhat surprising since in Han (1996) Word-medial /t*/ was found to be nearly twice as long as Word-initial /t*/. However, in Han's study, the comparison probably involved more than simply Word-medial vs. Word-initial. The carrier sentence for Word-medial test consonants is "This is called____" in which the test word, a nonsense word, is usually produced emphatically, forming an AP by itself. On the other hand, Word-initial test consonants were located in real words and placed in an unemphatic position. Jun & Lee (1998) found that focused words were significantly longer than corresponding neutral words for three out of four speakers. So Han's durational difference is likely to be greater than ours. In our sample, the comparison is a minimal one.

4.6. *Inherent articulatory properties of stops*

In section 3.1.1 (Fig. 6), we observed that the test consonants varied in their extent of linguopalatal contact, in the decreasing order of /t*, t^h, t, n/. Furthermore, tense and aspirated stops can be grouped into a "strong" consonant type having greater linguopalatal contact, while nasal and plain stops can be grouped into a "weak" consonant type with lesser linguopalatal contact. As can be seen in Fig. 6, the two stronger consonants maintain their large contact area at the beginning of all prosodic domains, and therefore show smaller effects of prosodic position, while the two weaker consonants are overall more variable, and thus show larger effects of prosodic position, with their contact ranging from large in higher prosodic positions to small in lower prosodic positions. Most notably, the weaker consonants /t/ and /n/ show particularly large differences in contact between AP-initial and IP-initial positions compared to the stronger consonants /t*/ and /t^h/. The lesser contact in these lower prosodic positions has been previously noticed for /t/ as an oral lenition (cf. Shin, 1997); here we see that /n/ exhibits this same lenition.

As a result of the greater prosodic effect on /t/ and /n/, the differences in overall amounts of contact between the four consonants can be seen more clearly in the lower prosodic positions, where the stronger consonants remain strong while the weaker consonants show more weakening. In contrast, in higher prosodic positions where all the consonants are at their strongest, the differences between /t, n/ and /t*, t^h/ are less striking.

5. Conclusion

The hypotheses of this study were that Korean would show domain-initial strengthening in the articulation and acoustics of consonants, except perhaps for the tense stops. These hypotheses have been largely supported by the experiments. We have shown that Korean consonants /t, t^h, t*, n/ vary in their linguopalatal contact as a function of their position in a prosodic hierarchy. Consonants initial in higher phrasal prosodic domains are strengthened relative to consonants in lower domains. These results for Korean are particularly clear: not only do all four consonants studied here show this pattern, but there is more consistency across speakers than in previous studies of English and French consonant articulation. All speakers distinguish at least three prosodic levels by consonant contact. Thus the prosodically-conditioned variation in Korean VOT found by Jun is shown to occur for oral articulation as well.

A second clear result of this study is the strong correlation of linguopalatal contact with duration (both articulatory and acoustic). This relation was much less strong in the previous studies of English and French. Thus in Korean, unlike in these other languages, it seems likely that differences in articulatory contact directly result from, or directly give rise to, differences in duration. That is, in Korean "lengthening" and "strengthening" appear to be a single effect, whereas in the other languages they appear to be somewhat more independent effects. We have interpreted our data as suggesting that up to about 80 ms in duration, consonants show articulatory undershoot proportional to their durations; beyond this duration no undershoot is seen.

At the same time, this study shows that Korean has two different strengthening effects: one is the more general effect seen in other languages as well, namely that consonants are stronger in higher prosodic positions, with respect to oral, nasal, and glottal articulations; the second is an effect specific to Korean, namely that tense and aspirated consonants are stronger Word-medially than Word-initially in terms of their oral articulation. We have proposed that these two effects compete within the Word domain, with somewhat variable results across speakers.

A third result of this study is that several consonant acoustic dimensions also vary with prosodic position. Some of these (e.g., stop closure duration, VOT, Total Voiceless Interval, %voicing, vowel duration, nasal energy) are well-correlated with the linguopalatal contact measure, and some of these (e.g., stop closure duration), plausibly result from the variation in linguopalatal contact. Others of these (e.g., %voicing, Total Voiceless Interval, nasal energy) vary in ways that appear unrelated to linguopalatal contact (except for exhibiting the same overall prosodic pattern). Thus the prosodically-conditioned variation in Korean VOT found by Jun is shown to extend across additional prosodic domains. Any or all of these acoustic dimensions may allow listeners to hear the prosodic phrasing intended by the speaker; this remains to be established by perceptual testing.

Finally, this study shows that different stop consonants within a single language can vary in how they are affected by prosodic position. The Korean tense and aspirated stops are overall stronger, and show less effect of prosodic position. The Korean lenis and nasal stops are overall weaker, but show more effect of prosodic position.

Thus we have reinforced the point that there is no simple distinction between segmental and suprasegmental properties. In Korean, the manner (degree) and even the place of a stop consonant's primary oral constriction - defining properties of these segments - are in part determined by prosodic structure.

We would like to thank our subjects Namhee Lee and Jiyoung Yoon for participating in the experiment; Sun-Ah Jun for help throughout the project, especially with corpus design; and other members of the UCLA Phonetics Lab for comments on earlier presentations of the work. Preliminary versions of this study were presented at the 1998 Norfolk meeting of the ASA, and the 1998 International Conference on Korean Linguistics in Hawaii. This work was supported by grant SBR-9511118 from the NSF to the second author.

References

Abacus Concepts (1992) *StatView* (Abacus Concepts, Inc., Berkeley CA).

Barry, M. (1991) Temporal modelling of gestures in articulatory assimilation. *Proceedings of the XIIth International Congress of Phonetics Sciences, vol. 4,* 14-167.

Beckman, M. & Jun, S.-A. (1996) K-ToBI (Korean ToBI) Labeling Conventions. Unpublished manuscript. Ohio State University & UCLA.

Beckman, M. & Pierrehumbert, J. (1986) Intonational Structure in Japanese and English, Phonology Yearbook, **3**, 225-309.

Byrd, D. (1994) Articulatory Timing in English Consonant Sequences. Ph.D. Dissertation. University of California, Los Angeles. (Published in *UCLA Working Papers in Phonetics,* **86**.)

Byrd, D., Flemming, E., Mueller, C., & Tan, C-C. (1995) Using regions and indices in EPG data reduction. *Journal of Speech Hearing and Research,* **38**, 821-827.

Byrd, D. & Saltzman, E. (1998) Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, **26**(2), 173-200.

Cho, T. & Ladefoged, P. (1997) Variations and universals in VOT: evidence from 17 endangered languages. *UCLA Working Papers in Phonetics*, **95**, 18-40.

Cho, T. (1996) *Vowel correlates to consonant phonation: An acoustic-perceptual study of Korean obstruents.* The University of Texas at Arlington M.A. Thesis.

Cho, Y.-M. Y. (1990) Syntax and phrasing in Korean. In *The phonology-syntax connection* (S. Inkelas & D. Zec, editors), pp. 47-62. Chicago: Chicago University Press.

Cooper, A. (1991) Glottal gestures and aspiration in English. Ph.D. Dissertation, Yale University.

de Jong, K. (1989) Initial Tones and Prominence in Seoul Korean. *Working Papers in Linguistics,* **41**, Ohio State University.

Docherty, G. J. (1995) On a gestural account of lenis stop voicing in Korean: comments on Jun. In *Phonology and phonetic evidence: Papers in Laboratory Phonology IV* (B. Connell & A. Arvaniti, editors), pp. 254-264. Cambridge, UK: Cambridge University Press.

Fougeron, C. & Jun, S.-A. (1998) Rate effects on French intonation: Prosodic organization and phonetic realization. *Journal of Phonetics*, **26**, 45-69.

Fougeron, C. & Keating, P. A. (1996) Variations in velic and lingual articulation depending on prosodic position: Results for two French speakers. *UCLA Working Papers in Phonetics*, **92**, 88-96.

Fougeron, C. & Keating, P.A. (1997) Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America,* **101**(6), 3728-3740.

Fougeron, C. (1998) *Variations articulatoires en début de constituants prosodiques de différents niveaux en français.* Ph.D. dissertation, Université Paris III - Sorbonne Nouvelle.

Han, J.-I. (1996) *The Phonetics and Phonology of Tense and Plain Consonants in Korean.* Ph.D. Dissertation, Cornell University.

Hirst, D. & Di Cristo, A. (1984) French intonation: a parametric approach. *Die Neueren Sprachen,* **83**(5), 554-569.

Hsu, C.-S. & Jun, S.-A. (1998) Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic?. *UCLA Working Papers in Phonetics*, **96**, 69-89.

Jun, J. (1944) Metrical Weight Consistency in Korean Partial Reduplication. Phonology, **11**, 69-88.

Jun, S.-A. (1993) *The phonetics and phonology of Korean prosody*. Ph.D. dissertation, Ohio State University.

Jun, S.-A. (1995) Asymmetrical prosodic effects on the laryngeal gesture in Korean. In B. Connell & A. Arvaniti (eds.), In *Phonology and phonetic evidence: Papers in Laboratory Phonology IV* (B.Connell & A. Arvaniti, editors), pp. 235-253. Cambridge, UK: Cambridge University Press.

Jun, S.-A. (1998) The Accentual Phrase in the Korean prosodic hierarchy. *Phonology,* 15(2), 189-226.

Jun, S.-A., Beckman, M., & Lee, H. (1998). Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. *UCLA Working Papers in Phonetics,* **96**, 43-68

Jun, S.-A., & Lee, H. (1998). Phonetic and phonological markers of contrastive focus in Korean. *Proceedings of the 5th International Conference on Spoken Language Processing*, vol.4, pp. 1295-1298. Sydney, Australia.

Kang, O. (1992) Korean prosodic phonology. Ph.D. Dissertation, University of Washington.

Keating, P. (1984) Physiological effects on stop consonant voicing. *UCLA Working Papers in Phonetics*, **59**, 29-34.

Krakow, R. A., Bell-Berti, F., & Wang, Q. E. (1994) Supralaryngeal declination: evidence from the velum. In *Producing speech: A Festschrift for Katherine Safford Harris,* (F. Bell-Berti & L. Raphael, editors), pp. 333-353 (Chapter 23). Woodbury, NY: AIP Press.

Lee, S.-H. (1989) Intonational domains of the Seoul dialect of Korean. Paper presented at 117th meeting of ASA. Syracuse, NY.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge, Mass.: MIT Press.

Lindblom, B. (1963) Spectrographic study of vowel reduction. *Journal of the Acoustical Society of* America, **35**, 1773-1781.

Martin, S. (1982) Features, markedness, and order in Korean Phonology. In *Linguistics in the Morning Calm* (I. Yang, editor), pp. 601 - 618. Seoul: Hanshin.

Moon, S. J. & Lindblom, B. (1994) Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America,* **96**, 40-55.

Nespor, M. & Vogel, I. (1986) *Prosodic phonology.* Dordrecht: Foris Publications.

Pasdeloup, V. (1990) Modèle de règles rythmiques du français appliquées à la synthèse de la parole. Ph.D disseration, U. de Provence.

Pierrehumbert, J. & Talkin, D. (1992) Lenition of /h/ and glottal stop. In *Papers in Laboratory Phonology II: gesture, segment, prosody* (G. Docherty & D. R. Ladd, editors), pp. 90-117. Cambridge, UK: Cambridge University Press.

Selkirk, E. (1984) *Phonology and syntax: the relation between sound and structure*. MIT Press.

Selkirk, E. (1986) On derived domains in sentence phonology. *Phonology Yearbook*, **3**, 371-405.

Shattuck-Hufnagel, S. & Turk, A. (1996) A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Reseaarch*, **25**(2), 193-247

Shin, J. (1997) Consonantal production and coarticulation in Korean. Ph.D. Dissertation, University of London.

Silva, J.D. (1992) The phonetics and phonology of stop lenition in Korean. Ph.D. Dissertation, Cornell University.

Stevens, K. N. (1999). *Acoustic Phonetics.* Cambridge, Mass.: MIT Press.

Stevens, K.N., Keyser, S.J., & Kawasaki, H. (1986) Toward a phonetic and phonological theory of redundant features. In *Invariance and Variability in Speech Processes* (J.S Perkell & D.H. Klatt, editors), pp. 426-463. New Jersey: Lawrence Erlbaum Associates.

Yu, J. (1989) A study on Korean glottalized tense sounds and aspirated. *Hankul,* 203, 25-48. Seoul, Korea.

Appendix. Test sentences for /t, t^h, t*/

(a) Higher levels above Word for /t/ in /a#_a/. (# = a prosodic boundary)

U-Initial (Ui)	igosɨn patakk**a.** [$_U$ **ta**mbiga jəgisə nɛɾinda] this place seashore 'sweet-rain' here is-falling 'This place is the seashore. Sweet-rain falls down here.'
IP-Initial (IPi)	igosɨn patakk**a,** [$_{IP}$ **ta**mbiga nɛɾinɨn goʃida] this place seashore, 'sweet-rain' falls-Rel place 'This place the seashore, where the 'sweet-rain' falls down.'
AP-Initial (APi)	idiɾɨn modug**a** [$_{AP}$ **ta**mbiɾɨl] tʃoahanda these people all sweet-rain likes 'These people all like 'Sweet-rain'.'
Word-Initial (Wi)	idiɾɨn [$_{AP}$ patakk**a** ($_W$ **ta**mbiɾɨl)] tʃoahanda. these people seashore sweet-rain likes. 'These people like 'seashore sweet-rain'.'

(b) Word and Syllable levels for /t/ in /a#_a/. (# = a prosodic boundary)

Word-Initial (Wi)	idiɾɨn [$_{AP}$ kojesa ($_W$ **ta**ɾiɾɨl)] tʃabat*a. these people acrobat legs held 'These people held the legs of the acrobat.'
Syllable-Initial (Si)	idiɾɨn [$_{AP}$ koje ($_W$ sa($_S$ **ta**ɾiɾɨl))] tʃabat*a. these people circus ladder held 'These people held the circus ladder.'

(c) Higher levels above Word for /tʰ/ in /a#_a/. (# = a prosodic boundary)

U-Initial (Ui)	igosɨn patakk**a**. [U **tʰa**dʒani igosɨl tʃoahanda] this place seashore. Tarzan here like 'This place is the seashore. Tarzan likes it.'
IP-Initial (IPi)	igosɨn patakk**a**, [IP **tʰa**dʒanɨy kohjaŋida] this place seashore. Tarzan's hometown 'This place is the seashore, Tarzan's home town.'
AP-Initial (APi)	idɨrɨn modug**a** [AP **tʰa**dʒanɨl] tʃoahanda these people all Tarzan likes 'These people all like Tarzan.'
Word-Initial (Wi)	idɨrɨn [AP patakk**a** (W **tʰa**dʒanɨl)] tʃoahanda. These people seashore Tarzan likes 'These people like the seashore Tarzan.'

(d) Word and Syllable levels for /tʰ/ in /a#_a/. (# = a prosodic boundary)

Word-Initial (Wi)	idɨrɨn [AP pakhas**a** (W **tʰa**ŋsogɨl)] tɨljəda pwat*a. these people sergeant Park's bath tub looked in 'These people looked in Sergeant Park's bath tub'
Syllable-Initial (Si)	idɨrɨn [AP pakha (W s**a**(S **tʰa**ŋsogɨl))] tɨljəda pwat*a. these people mint candy looked in 'These people looked in the mint candy.'

(e) Higher levels above Word for /t*/ in /a#_a/. (# = a prosodic boundary)

U-Initial (Ui)	igosɨn patakk**a**. [U **t*a**kpuriga jəgisə sanda] this place seashore T*akpuli (nickname) here lives 'This place is the seashore. T*akpuli lives here.'
IP-Initial (IPi)	igosɨn patakk**a**, [IP **t* a**kpurɨy kohjaŋida.] this place seashore. T*akpuli 's hometown 'This place is the seashore, (which is) T*akpuli's hometown.'
AP-Initial (APi)	idɨrɨn modug**a** [AP **t*a**kpurilɨl tʃoahanda] these people all T*akpuli-Acc. likes 'These people all like T*akpuli.'
Word-Initial (Wi)	idɨrɨn [AP patakk**a** (W **t*a**kpurilɨl)] tʃoahanda. these people seashore T*akpuli likes 'These people like the seashore T*akpuli.'

(f) Word and Syllable levels for /t*/ in /ɛ #_a/. (# = a prosodic boundary)

Word-Initial (Wi)	kɨjədʒanɨn [AP tʃoŋib**ɛ** (W **t*a**ragagirɨl)] tʃoahanda. the woman-Top. sail boat running-after-Acc likes 'The woman likes running after the paper boat.'
Syllable-Initial (Si)	kɨjədʒanɨn [AP hjəndɛ (W b**ɛ**(S **t*a**ragirɨl)] tʃoahanda. the woman-Nom. modern 'a kind of dance' like 'The woman likes the modern 'pet*alaki'.'

Domain-initial articulatory strengthening in four languages

Patricia Keating, Taehong Cho, Cécile Fougeron[1],
and Chai-Shune Hsu[2]

1. Introduction

This paper is about one way in which prosody affects individual speech segments, with segmental phonetics showing a perhaps surprising sensitivity to higher-level linguistic structure. By *prosody* we mean the phrasal and tonal organization of speech. We will show that phonetic properties of individual segments depend on their *prosodic position,* or position in prosodic structure.

It is well-known that in a monosyllabic CVC word, the initial consonant can be pronounced differently than the final consonant, the initial consonant being longer and having greater articulatory magnitude (e.g. Byrd, 1994). Some interesting recent acoustic studies have extended this line of inquiry above the syllable and word level to phrasal levels. For example, at the LabPhonII conference, Pierrehumbert & Talkin (1992) presented a study in which they used acoustic measures of breathiness to show that /h/ is more consonant-like when it is phrase-initial than when it is phrase-medial. Similarly, VOT of /t/ is longer phrase-initially. This latter result was extended by Jun (1993), who compared the VOT of Korean /p^h/ in three positions: initial in a small phrase, initial in a word, medial in a word; VOT varied as shown in Figure 1.

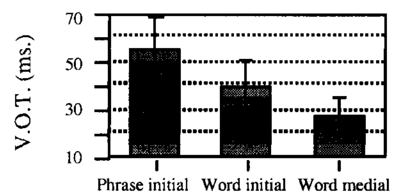


Figure 1. VOT of Korean /p^h/ as a function of prosodic position.
Summary of data from Jun 1993:235 (Figure 6.2).

Then Dilley, Shattuck-Hufnagel & Ostendorf (1996) showed that higher phrasal levels can also differ. They tabulated the presence of glottalization of vowel-initial words in a radio-news corpus, and found that the likelihood of glottalization depends on the prosodic position of the word. Glottalization is most likely at the beginning of an Intonational Phrase (a large phrase), next most likely at the beginning of an Intermediate Phrase (a smaller phrase), and least likely phrase-medially.

Articulatory studies that compare positions in phrases include Stone (1981), van Lieshout, Starkweather, Hulstijn & Peters (1995), Byrd, Kaun, Narayanan & Saltzman (1996), Gordon (1996), Hsu & Jun (1997), and Byrd & Saltzman (1998). In our own earlier work (Fougeron & Keating, 1997), we compared the articulation of /n/s in different prosodic positions. The speech materials consisted of arithmetic expressions as in (1).

[1] Lab. de Phonétique, ILPGA, Univ. Paris III and Lab. de Psycholinguistique Expérimentale, Univ Genève
[2] Now at Nuance Communications

(1) 89 x (89 + 89 + 89) = a lot

Reiterant speech was used, with most syllables replaced by the syllable /no/, as in (2).

(2)	89	times	(89	plus	89	plus	89)	= a lot
	nonono	no	(nonono	no	nonono	no	nonono)	= a lot

The prosodic organization of the test utterances was characterized by transcribing groupings of words into smaller phrases and larger phrases (using the ToBI conventions (Silverman, Beckman, Pitrelli, Ostendorf, Wightman, Price, Pierrehumbert & Hirschberg, 1992; Beckman & Elam, 1997)). Each reiterant syllable was then coded as initial, medial, or final in each of the prosodic domains Word, small Intermediate Phrase (or PP), large Intonational Phrase (or IP), and Utterance. /n/s which were not initial within a word were also coded as syllable-initial. The Utterance-initial /n/s were always and only at the beginning of the sentence, but otherwise there was no *unique* relation between prosodic position above the word and linear position in the sentence.

The relevant result here, shown in Figure X.2a, is that in general, /n/s which were initial in higher domains had more linguopalatal contact than /n/s which were initial only in lower domains. The effect of being in domain-initial position was generally cumulative. Each speaker showed a hierarchical pattern of peak contact, distinguishing at least three domains in this way. However, no speaker distinguished all the domains, and no distinction was reliable for all speakers. Speaker 1 distinguished IP, PP and W; Speaker 2 distinguished U, IP/PP, W and S, and Speaker 3 distinguished U, IP, PP/W, and S.

Articulatory duration (the duration of the stop consonant seal, from EPG data), not reported in Fougeron & Keating (1997) but shown in Figure X.2b, followed a similar pattern. Speakers 1 and 3 distinguished IP, PP, and W/S; Speaker 2 distinguished IP, PP, W, and S. (Note that the stop seal could not be isolated for Utterance-initial (post-pausal) tokens.) The within-speaker correlations between articulatory duration and linguopalatal contact for domain-initial tokens above the Word level were low to modest (r^2 from .09 to .27), but surprisingly, those between acoustic duration and contact were notably greater (r^2 from .3 to .55).

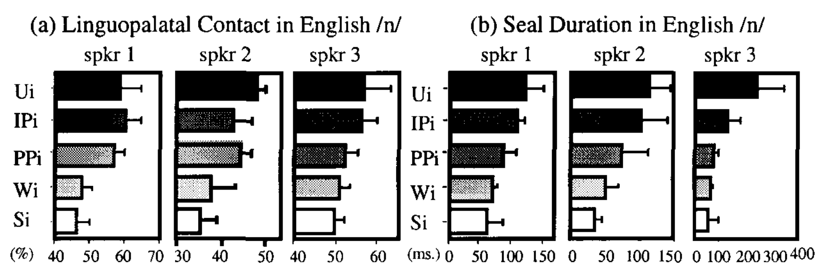


Figure 2. English EPG data by speaker for (a) Peak contact, based on Fougeron & Keating (1997). The horizontal bars show the %electrodes (of 96) contacted. (b) Articulatory seal duration. The horizontal bars show duration in ms. All graphs show values for consonants in initial position (indicated by small "i" in the axis labels) in the domain indicated (U for Utterance, IP for Intonational Phrase, PP for Phonological or Intermediate Phrase, W for Word, S for Syllable).

We called this pattern "domain-initial strengthening" because the lingual articulations appeared to be stronger for consonants at the beginning of each prosodic domain. Because in this corpus there were often three or more syllables in each domain, we could test specifically whether this resulted from weakening of all non-initial syllables (that is, the first syllable's consonant is different from all others), vs. final-syllable weakening (that is, the last syllable's consonant is different from all others). The results clearly showed the former. We also found no evidence for articulatory declination (global, utterance-level trends, e.g. Krakow, Bell-Berti & Wang, 1994). Therefore in the present study we will focus only on domain-initial consonants.

In this first study, however, as all test syllables were CVs, we could not look for domain-final weakening of pre-boundary coda consonants. We also could not clarify how domain-initial strengthening compares in magnitude with the word- or syllable-level effects that are the topic of many studies. A further study (Keating, 1997; Keating, Wright & Zhang, in prep.) compared word-initial consonants not only to sentence-initial consonants, but also to word-final ones. Four subjects produced test words (three-syllable /C...C/ nonsense words where C is one of /t d n l/) in different positions in a carrier sentence. (Other experimental factors in the study are ignored here, i.e. the data collapsed over them.) Peak linguopalatal contact of these consonants is shown in Figure 3: stop consonants /t d n/ have more contact at the beginning of the utterances than they do elsewhere, and they have more contact at the beginning of a word (and syllable) than at its end. For /l/, however, there is no effect of sentence-initial position, though the expected large effect of position in word/syllable is found. From these results it can be seen, first, that for stops the difference between U-initial vs. W-initial consonants is somewhat greater than the more familiar word-initial vs. word-final effect; and second, that the two kinds of effects must be independent since /l/ shows one but not the other.

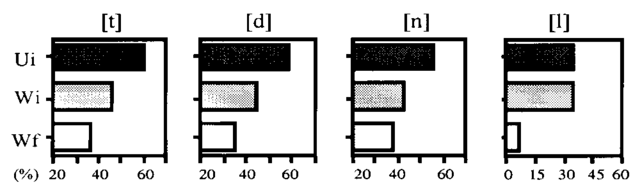


Figure 3: Peak EPG contact for four English consonants, averaged over four speakers, as a function of prosodic positions Utterance-initial, Word-initial, Word-final.

The present study follows up on our earlier results for English by comparing several languages. Not only do we want to know whether the results hold beyond English, but we want to know whether other prosodic differences among languages are reflected in any initial strengthening effect. Lehiste (1964) showed that languages differ in how they mark word boundaries. She proposed that this depends on a language's phonology; for example, a language with phonemic vowel length would not use vowel lengthening to mark boundaries. The same could be true for initial strengthening. Byrd *et al.* (1996) in their LabPhonV presentation found relatively little effect of phrasal position on spatial position of articulators in Tamil, though they did find effects on duration and timing. This suggests that spatial effects as in English might be language-particular. However, it is not clear that their Tamil corpus included a sufficient range of different prosodic domains. Therefore our study includes three languages and clear examples of larger and smaller phrasal domains.

Since English has such prominent lexical stress and nuclear pitch accent, it might be expected that its domain edges would be phonetically less marked than edges in languages with less prominent heads. The three languages studied here, French, Korean, and Taiwanese, allow such comparisons. Taiwanese is a lexical tone language, and thus, since it cannot use tones to mark domain heads, might be expected to show large edge-marking effects. On the other hand, Taiwanese tone sandhi is organized in a phrasal domain which does not seem to be prosodic (Hayes, 1990; Hsu & Jun, 1996), and for that reason prosodic domains might be expected to receive little phonetic marking. French and Korean differ from both English and Taiwanese in having neither lexical tone nor lexical stress. They are prosodically similar to each other; both have a small prosodic domain defined by phrasal tones. At the same time, it has been proposed (Jun, 1995, p.c.) that these two languages differ in terms of duration and amplitude variation within that phrase — Korean reinforces the beginning of the phrase but French the end. If this is so, we might expect French not to show domain-initial articulatory strengthening like Korean.

2. General Methods

2.1. Prosodic Domains

For each language, prosodic domains must be determined and defined. A schematic of a partial hierarchy of prosodic domains (mostly above the word level) is shown in Figure X.4. (For a thorough review of theories of prosodic hierarchies, see Shattuck-Hufnagel & Turk (1996).)

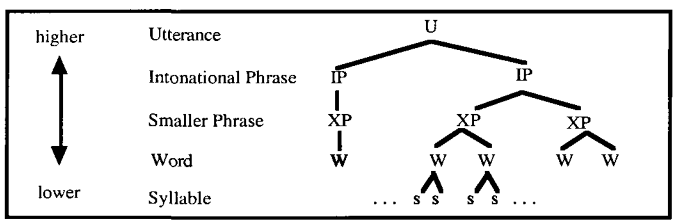


Figure 4. A partial Prosodic Hierarchy adopted in this study.
One or more instances of each level may appear under the level above it.

One domain that seems comparable across languages is the Intonational Phrase, or IP. An IP is marked by a complete intonational contour, and can be set off naturally by pauses. An IP can comprise a full sentence, but in our experiments it usually comprised a clause or topic phrase within a longer sentence (punctuated by a comma or semi-colon). For two of the languages, French and Korean, we also tested a possible higher domain, the Utterance, corresponding to the second of two sentences (punctuated by a period), and marked by a full pause, sometimes with a breath. Whether there is a systematic difference between Utterance and Intonational Phrase is somewhat controversial. Nespor & Vogel (1986) distinguished them on the basis of where some phonological rules apply. However, in terms of intonation and pausing, they need not be different; and Wightman, Shattuck-Hufnagel, Ostendorf & Price (1992) found no difference in their amounts of final lengthening. In our Korean experiment we instructed subjects not to pause within a sentence, so that the Utterance break is marked by a pause but the IP break usually is not. In our French experiment, which did not give explicit instructions, subjects were more likely to pause between IPs.

A phrasal domain smaller than the IP was also sought, corresponding to the Phonological or Intermediate Phrase studied for English in Fougeron & Keating (1997). Such a phrase would be marked by less than a complete intonational contour. In French and Korean the Accentual Phrase was chosen, as it is easy to transcribe from spoken utterances. An AP usually consists of a small number of content words, plus function words, with an associated phrasal tone pattern. Following the analysis of French prosody given by Jun & Fougeron (1995) and Fougeron & Jun (1998), the French AP has an underlying phrasal tone sequence LHLH. Following the analysis of Seoul Korean prosody given by Jun (1998), the Korean AP is also marked by an underlying phrasal tone sequence LHLH. For Taiwanese, however, no such prosodic domain, layered under the IP, could be identified in our initial pilot work, and therefore none was included.

Finally, initial and medial positions within a Word domain were included in each experiment. What counts as a Prosodic Word in a given language is controversial. In English our Word was fairly large by some prosodic standards, being lexically complex (e.g. "eighty-nine"), but nonetheless having only one primary lexical stress. In Taiwanese, a similarly large Word was used. In Korean, Words were single names or nouns, while for French, Words were parts of larger names. The Syllable-initial consonants were all Word-medial.

Table 1. shows the corpus for French /n/. The corpora for other /t/ and for the other languages are similar in design. In all studies, the prosodic position of test consonants was varied; by varying the text around the test syllable, the prosodic structure is varied, while the absolute position of the test syllable is kept the same. (Since it is possible that some language other than English might show articulatory declination, we control for this in all studies.)

Table 1. Corpus for French /n/

Positions	Test Consonant /n/ in /a_a/
Ui	Paul aime Tata. **N**adia les protège en secret. *'Paul loves Auntie. Nadia protects them in secret'*
IPi	La pauvre Tata, **N**adia et Paul n'arriveront que demain. *'Poor Auntie, Nadia and Paul won't arrive until tomorrow'*
APi	Tonton, Tata, **N**adia et Paul arriveront demain. *'Uncle, Auntie, Nadia and Paul will arrive tomorrow'*
Wi	Paul et Tata-**N**adia arriveront demain matin. *'Paul and Auntie Nadia will arrive tomorrow morning'*
Si	Tonton et A**n**abelle arriveront demain matin. *'Uncle and Anabelle will arrive tomorrow morning'*

2.2. Data Collection

The primary measure of strengthening reported here will be the maximum amount of contact between the tongue and the palatal surface, as recorded by electropalatography (EPG). The amount of contact is an index of tongue height at the point of contact, and thus a measure of the strength of an articulation. All studies used the Kay Elemetrics Palatometer. With the Palatometer, a talker wears an individual, custom-made pseudopalate that covers the surface of the hard palate and the inner surfaces of the upper teeth with 96 contact electrodes. For French, Korean, and Taiwanese speakers, the frontmost row of electrodes extends onto the back surface of the upper teeth, and two electrodes were placed at the middle of the front two incisors, so that at least some dental contact could be registered. This arrangement of electrodes is shown in Figure 5. The Palatometer records the pattern of tongue-pseudopalate contact every 10 ms. The audio signal was recorded with a head-mounted microphone, at 12.8 kHz, into the same data file.

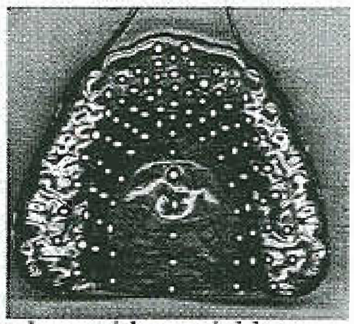


Figure 5. Scan of pseudo-palate, with special layout of the 96 contact electrodes.

2.3. Data measurement

Maximum linguopalatal contact was determined by calculating, for each data frame, the percentage of contacted electrodes over the 96 or some subset of electrodes. The maximum value in each test consonant was recorded as the peak contact for that token. Temporal measures were also made, including the number of frames showing a complete stop closure (articulatory seal duration), acoustic closure duration, and for voiceless stop /t/, acoustic VOT. Reliable differences were determined by ANOVA and *posthoc* tests at the .05 level of significance.

3. Methods and results for each language

3.1. French

3.1.1. Methods

Experiments on French have been reported in Fougeron & Keating (1996), and much additional data is included in Fougeron (1998). Two subjects participated in this study: one of the authors (female, Speaker 1) plus one other subject (male, Speaker 2). The test consonants reported on here were unaspirated /t/ and /n/. /n/ was in a /a_a/ context, /t/ in a /ɔ̃_ ɔ̃/ context.

There were 20 repetitions of each sentence, blocked by test sentence. Contact was measured over the whole palate, and in a small anterior region of 18 electrodes defined *posthoc*.

3.1.2. Results

EPG results are shown in Figures 6a-b. First, for the peak contact data, in Figure 6a, there was an effect of prosodic position for both speakers, with a generally cumulative increase of contact from lowest to highest domains. More distinctions are made for /n/: both speakers distinguish all domains except IP from Utterance. For /t/, not only is the distinction between Utterance and IP unclear (in fact, it is reliably reversed for one speaker), but also the distinction between Word-initial and Syllable-initial is not made. The reliable differences, then, are those between a large phrasal domain (IP, Utterance), a small one (AP), and something smaller (Word or Syllable). Detailed analysis of contact in the front region of the palate showed that the greater contact in higher prosodic positions was mainly located in the posterior part of that anterior region. This difference is seen in the sample tokens shown in Figure 7, along with other differences presumably reflecting the height of the tongue body.

The duration data show fewer distinctions. The duration of the articulatory closure or seal, in Figure 6b, shows a large difference between U/IP and the smaller domains. Which of the further, smaller, differences are reliable varies between the speakers. However, the overall lengthening pattern is cumulative like that for contact, and indeed the two measures are well-correlated (r^2 from .56 to .76). Acoustic duration of /n/ shows lengthening at beginnings of lower domains, but IP- and U-initial /n/s are very short. Acoustic closure duration of /t/, measured only for the lower domains because they involve no pause, patterns similarly to /n/ (and to articulatory duration, not surprisingly). For VOT of unaspirated /t/, shown in Figure 8a, there was little effect of prosodic position. The only difference found for both speakers was between Syllable-initial and IP-initial positions.

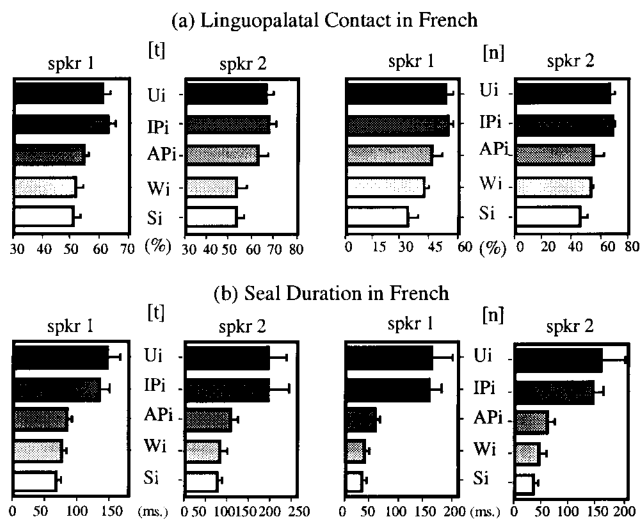


Figure 6. Data for French, displayed as in Figure 2. (a) Peak EPG contact for /t, n/; (b) Articulatory duration for /t, n/.

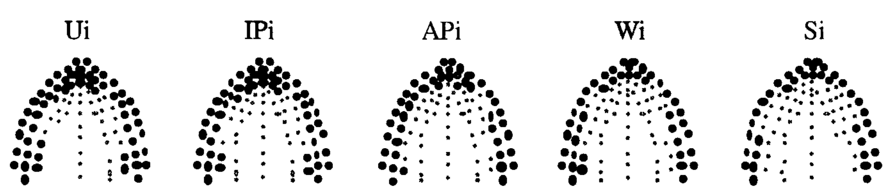


Figure 7. Sample French tokens for /n/ showing contact patterns across prosodic positions.

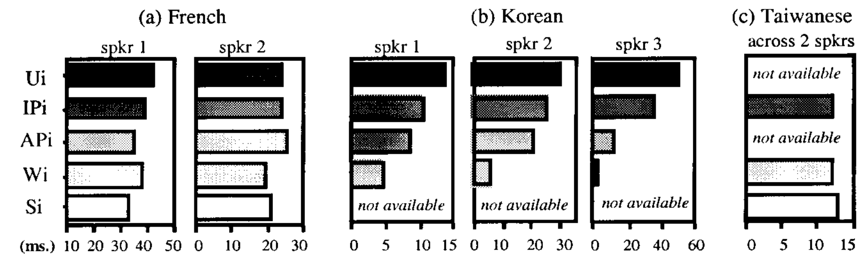


Figure 8. (a) French VOT, (b) Korean VOT & (c) Taiwanese VOT across prosodic positions.

3.2. Korean

3.2.1. Methods

Three subjects participated in this study, one of the authors (male, Speaker 2) and two others (one male, Speaker 1, and one female, Speaker 3). The complete study included test consonants /n t t^h t'/; here we report on only /t/ and /n/. All of the domains in Figure 4 were included. However, two corpora were used for each consonant, one for comparison of higher-level domains, another for word-level domains. Otherwise we could not construct meaningful and grammatical sentences. In the higher-level corpus, for domains Utterance, IP, AP, and Word, both /t/ and /n/ were in a /a_a/ context. In the lower-level corpus, for domains Word vs. Syllable, /n/ was in a /o_ɛ/ context and /t/ was in a /a_a/ context. Only Speakers 1 and 2 produced the lower-domain corpus for /t/. Speakers read 20 repetitions of each test item, blocked by test consonant.

Contact was measured over the whole palate, and also separately in front and back regions. Measurements taken from the data included those taken for French. Additional measures, and detailed comparisons of the four test consonants, are reported elsewhere (Cho, 1998; Cho & Keating, in preparation).

3.2.2. Results

EPG results are shown in Figures 9a-b. First, in the overall contact data, shown in Figure 9a, all prosodic levels are generally distinguished by all the speakers for both test consonants, except that Speaker 3 does not have more contact for AP-initial than for Word-initial for either consonant, and Speaker 1 does not differentiate W-initial from Syll-initial /t/. Figure 10 shows sample tokens. Here we can see that higher domains have more front contact, as well as more back contact. Statistical comparison of contact in front vs. back regions shows that the overall contact is most influenced by the front region (tongue blade) contact, though there is also a reliable small effect of prosodic position on back region (tongue dorsum) contact.

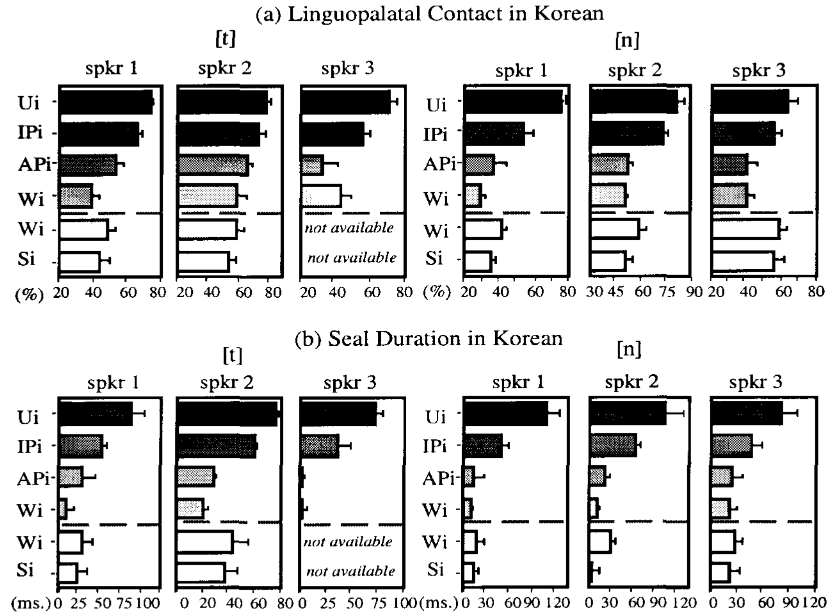


Figure 9. Data for Korean, displayed as in Figure 2. (a) Peak EPG contact for /t, n/; (b) Articulatory duration for /t, n/. Dashed horizontal line in each panel separates data from two different speech corpora; the two Word-initial conditions are not the same.

With articulatory seal duration, in Figure 9b, the phrasal domains are consistently distinguished by lengthening, but lower levels (AP vs. Word, Word vs Syllable) are generally not distinguished. Nonetheless, articulatory duration is well-correlated with peak contact, more so than in French (r^2 from .41 to .89). Acoustic duration is consistently cumulative when pooled across speakers, but the individual speaker data are not so consistent. Finally, unlike in French, VOT for /t/, shown in Figure 8b, distinguishes all four levels tested in Korean.

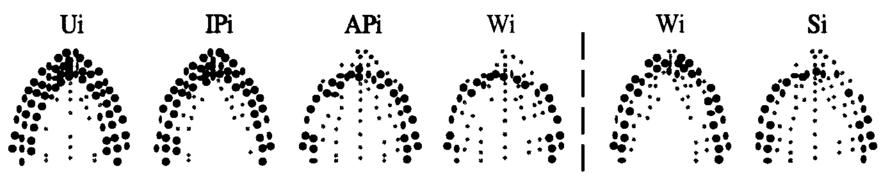


Figure 10. Sample Korean tokens for /n/ showing contact patterns across prosodic positions.

3.3. Taiwanese

3.3.1. Methods

Two subjects participated in a first study: one of the authors (female, Speaker 1) plus one other subject (male, Speaker 2). The test consonant was unaspirated /t/, in the syllable /ta/ with a surface mid-level tone, preceded by another /a/. As noted above, the prosodic domains tested in Taiwanese are fewer than in the other languages, in part because it was not clear if there is any kind of small phrase between Intonational Phrase and Word levels. A further study reported in

Hayashi et al. (1999, this volume) included an Utterance break between two sentences, and a small phrasal break associated with a heavy subject Noun Phrase.

The sentences were presented to the subjects written in Mandarin, to be translated by the speaker. There were 12 repetitions of each sentence, in varying orders. Ten repetitions of a control utterance (/pa-pa/, "Dad") were also collected, to determine the contact pattern of the vowel /a/. The very few electrodes contacted for /a/ by Speaker 1 were eliminated from analyses of the test consonant. Measurements taken from the data included the peak linguopalatal contact during /t/, the acoustic closure duration (which for /t/ is essentially the same as articulatory closure duration), and VOT.

3.3.2. Results

Results of the first study are shown in Figures 11a-b. There was an effect of prosodic position on peak contact for both speakers, but Speaker 2 distinguished all three positions in this way; Speaker 1 made no difference between Word and IP initial /t/. However, with acoustic (closure) duration, both speakers distinguished all three levels cumulatively. For VOT of these unaspirated stops, shown in Figure 8c, there was no effect of prosodic position.

In the follow-up study by Hayashi et al., more distinctions in contact were made by the speakers.

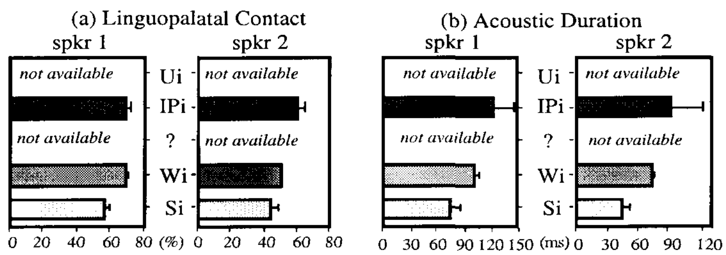


Figure 11. Data for Taiwanese /t/, displayed as in Figure 2. (a) Peak EPG contact; (b) articulatory duration.

4. Discussion

4.1. Domains

These experiments show clearly that there is phrasal/prosodic conditioning of articulation across languages — every subject makes at least one distinction (Word-internal vs. phrase-initial) and all speakers but one make at least one further distinction above the Word level, for every consonant studied. This conditioning generally affects both EPG contact, which reflects overall height of the tongue, and also duration, so the total effect is on contact-over-time. At the same time, the prosodic effects can be seen to differ across speakers and consonants within a language. It differs enough that we cannot say that any single prosodic hierarchy is exhibited by all languages and speakers, or that speakers are marking every level of a hierarchy.

In general, the distinction between two phrasal levels is robust, with most speakers distinguishing between a "high" phrasal domain and a "low" phrasal domain. In contrast, some other differences are not so robust. Most notably, Utterance is not consistently distinguished from Intonational Phrase. A phonetic distinction was found most clearly in Korean, where the difference between Utterance and IP was specifically linked to pausing. Thus our study supports Price, Ostendorf, Shattuck-Hufnagel & Fong (1991) in positing a break level “5” above the IP based on pausing.

Also in our results, Word-initial position is not consistently distinguished from both Syllable-initial and small Phrase-initial positions, and this is so whether our "words" are morphologically large (English, Taiwanese) or small (French, Korean).

The experiments presented here allow some comparisons of the relative sizes of different effects on linguopalatal contact. First, since French and Korean results are reported for two consonants, we can ask how the prosodic effect compares with the inherent segmental effect. In general, nasals have less contact than voiceless orals. It turns out that this difference is about the same in magnitude as the difference between pairs of prosodic positions. Compare, for example, French AP-initial /n/ for Speaker 1 with both Word-initial /n/ (prosodic comparison) and AP-initial /t/ (inherent segmental comparison) in Figure 7a. Another comparison is found in the two corpora for "higher" and "lower" domains in Korean. In the "higher" corpora, the Word is the lowest domain tested, while in the "lower" corpora it is the highest domain. The Word-initial consonants in the two corpora appear in different vowel contexts for /n/, which affect the contact location and extent. Again, this effect of vowel context turns out to be about the same as the difference between pairs of prosodic positions.

We have also presented data on articulatory and acoustic duration, and on VOT. In all of the languages, prosodic position affects consonant duration; sometimes this effect is more reliable than the effect on contact (e.g. in Taiwanese). However, only in Korean is VOT influenced by prosodic position.

4.2. Languages

Despite the various predictions made about possible language differences, the languages in this study show quite similar effects of prosodic position. In particular, French and Korean both distinguish IP from AP by linguopalatal contact. One intriguing difference between these two languages, though, concerns the strength of the correlations between initial consonant duration and contact: these are generally higher in Korean. Korean was also the only language to show an effect of prosodic position on the VOT (a temporal measure) of unaspirated stops. So there may well be a special pairing of temporal and spatial properties in domain-initial position in Korean.

On the other hand, the distinctions made in Taiwanese appear to be weaker than in the other languages. For example, Taiwanese is the only language in which any one of the speakers failed to distinguish IP-initial from Word-initial stops by linguopalatal contact. However, further study of Taiwanese Speaker 1, with a larger corpus more like that used for French and Korean, indicates very similar patterning to those languages. Therefore we hesitate to draw any general conclusions about either that language in particular, or tone languages in general.

In conclusion, we have shown that consonant articulation is subtly sensitive to a range of prosodic domains in several languages. Linguistic structure is relevant for even fine phonetic detail, and prosodic constituency can be marked by details of articulation as well as by the traditional prosodic parameters.

Notes

* This work was supported by NSF grant #SBR 95-11118. We also thank Kay Elemetrics for making the special pseudo-palates, the subjects who participated in the experiments (including Jiyoung Yoon, Namhee Lee, and Laurent Girard), and Lucy Vause and Wendy Hayashi for help with measurements and ms. preparation.

References

Beckman, M. & G. A. Elam. 1997. *Guidelines for ToBI Labelling. Version 3.* Unpulbished ms., Ohio State University.

Byrd, D. M. 1994. *Articulatory timing in English consonant sequences.* Ph.D. dissertation, UCLA, distributed as *UCLA Working Papers in Phonetics* 86: 1-196

Byrd, D. & E. Saltzman. 1998. Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, in press.

Byrd, D., A. Kaun, S. Narayanan, & E. Saltzman. 1996. Phrasal signatures in articulation. To appear in M. Broe & J. Pierrehumbert (eds.), *Laboratory Phonology V.* Cambridge: Cambridge University Press.

Cho, T. 1998. Domain-initial articulatory strengthening in the prosodic hierarchy in Korean: an EPG study. *Proceedings of the 11th International Conference on Korean Linguistics (ICKL)*: 363-372.

Cho, T. & P. Keating. In preparation. Articulatory and acoustic studies of domain-initial strengthening in Korean. Unpublished manuscript, UCLA.

Dilley, L., S. Shattuck-Hufnagel & M. Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24: 423-444.

Fougeron, C. 1998. *Variations articulatoires en début de constituants prosodiques de différents niveaux en français.* Ph.D. dissertation, Université Paris III - Sorbonne Nouvelle.

Fougeron, C. & S.-A. Jun. 1998. Rate effects on French intonation: Prosodic organization and phonetic realization. *Journal of Phonetics 26*: 45-69.

Fougeron, C. & P. Keating. 1996. Variations in velic and lingual articulation depending on prosodic position: Results for two French speakers. *UCLA Working Papers in Phonetics* 92: 88-96.

Fougeron, C. & P. Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101 (6): 3728-3740.

Gordon, M. 1996. The effect of stress and prosodic phrasing on duration, acoustic amplitude and air flow of nasals in Estonian. *UCLA Working Papers in Phonetics* 92: 151-159.

Hayashi, W., C.-S. Hsu & P. Keating. 1999. Domain-initial strengthening in Taiwanese: a follow-up study. *UCLA Working Papers in Phonetics* 97 (this volume).

Hayes, B. 1990. Precompiled phrasal phonology. In S. Inkelas & D. Zec (eds.), *The Phonology-Syntax Connection.* The University of Chicago Press and CSLI.

Hsu, C.-S. & S.-A. Jun. 1997. Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic?. Unpublished manuscript, UCLA.

Hsu, C.-S. & S. -A. Jun. 1996. Is Tone Sandhi Group part of the prosodic hierarchy in Taiwanese? Poster presented at the Third Joint Meeting of the Acoustical Societies of

America and Japan, Honolulu. Abstract published in *Journal of the Acoustical Society of America 100* (4) Pt. 2:2824.

Jun, S.-A. 1993. *The phonetics and phonology of Korean prosody*. Ph.D. dissertation, Ohio State University.

Jun, S.-A. 1995. Asymmetrical prosodic effects on the laryngeal gesture in Korean. In B. Connell & A. Arvaniti (eds.), *Phonology and phonetic evidence: Papers in Laboratory Phonology* IV. Cambridge: Cambridge University Press, 235- 253.

Jun, S.-A. 1998. The Accentual Phrase in the Korean prosodic hierarchy. *Phonology 15* (2), to appear.

Jun, S.-A. & C. Fougeron. 1995. The Accentual Phrase and the prosodic structure of French. *Proceedings of the 13th International Congress of Phonetic Sciences* 2: 722-725.

Keating, P. 1997. Word-initial versus word-final consonant articulation. Poster presented at 134th Meeting of the Acoustical Society of America, San Diego. Abstract published in *Journal of the Acoustical Society of America 102* (5) Pt. 2:3094.

Keating, P., R. Wright, & J. Zhang. In preparation. Word-level asymmetries in consonant articulation. Unpublished manuscript, UCLA.

Krakow, R. A., F. Bell-Berti, & Q. E. Wang. 1994. Supralaryngeal declination: evidence from the velum. In F. Bell-Berti & L. Raphael (eds.), *Producing speech: A Festschrift for Katherine Safford Harris*. AIP Press, Woodbury NY, Chap. 23, 333-353.

Lehiste, I. 1964. Juncture. *Proceedings of the 5th International Congress of Phonetic Sciences*, 172-200.

Nespor, M. & I. Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris Publications.

Pierrehumbert, J. & D. Talkin. 1992. Lenition of /h/ and glottal stop. In G. Docherty & D. R. Ladd (eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press, 90-117.

Price, P.J., M. Ostendorf, S. Shattuck-Hufnagel, & C. Fong. 1991. The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* 90 (6): 2956-2970.

Shattuck-Hufnagel, S. & A. Turk. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25 (2): 193-247.

Silverman, K., M. Beckman, J. Pitrelli, M. Ostendorf, C. Wightman, P. Price, J. Pierrehumbert, & J. Hirschberg. 1992. TOBI: A Standard for Labeling English Prosody. *Proceedings of the 1992 International Conference on Spoken Language Processing* 2: 867-870.

Stone, M. 1981. Evidence for a rhythm pattern in speech production: observations of jaw movement. *Journal of Phonetics* 9: 109-120.

van Lieshout, P.H., C.W. Starkweather, W. Hulstijn, & H.F.M. Peters. 1995. Effects of linguistic correlates of stuttering on emg activity in nonstuttering speakers. *Journal of Speech and Hearing Research* 38: 360-372.

Wightman, C.W., S. Shattuck-Hufnagel, M. Ostendorf, & P.J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America 92*: 1707-17.

Domain-initial strengthening in Taiwanese: a follow-up study

Wendy Hayashi[1], Chai-Shune Hsu[2] and Patricia Keating

Introduction

In a previous study (Hsu & Jun 1998; reported in Keating et al. 1998, this volume) we investigated whether Taiwanese consonants show domain-initial strengthening of consonant articulations of the sort we have found in other languages. Taiwanese is of interest in this regard because its prosodic domains are less well-established than in other languages. Most previous work on the prosody of Taiwanese (and related dialects, e.g. Chen 1987) has concerned the tone sandhi group (the domain in which tone sandhi takes place), but this domain is not strictly layered under the Intonational Phrase, and Hsu & Jun (1996) found no strengthening of consonants at the beginning of the tone sandhi domain. Therefore Hsu & Jun (1998) compared the articulation of the consonant /t/ when initial in only three prosodic domains: Syllable, Word, and Intonational Phrase. The measures of articulation were peak linguopalatal contact, Voice Onset Time, and acoustic closure duration. The result of the experiment was that, while both speakers distinguished all three domains by duration, and neither speaker distinguished them by VOT, only one of the two speakers distinguished all three by contact. The other speaker made no articulatory difference between Word-initial and Intonational-Phrase-initial /t/s. While this result was unexpected, Keating et al. were reluctant to conclude that Taiwanese behaves differently from other languages studied, because the Taiwanese study was the smallest of the set, and the corpus was not designed to be exactly parallel to those of the other studies. However, it would certainly be an interesting result if Taiwanese, a tone language, were in fact different from the other languages in this respect.

Therefore a second study was undertaken, designed to be more like those of French and Korean (this volume). To this end, additional prosodic domains were identified for Taiwanese, and an additional consonant was included.

Methods

Two subjects participated in the study, one female (Speaker 1, the second author) and one male (Speaker 2). These are the same subjects as in the original study.

The test domains began with unaspirated /t/ or /n/, followed by /a/, and were also preceded by /a/. The corpus for /t/ consisted of sentences containing real words, and is shown in Table 1. In some cases the target syllable was underlyingly /ta/; however, syllables closed by // become open syllables within the tone sandhi domain. The corpus for /n/ consisted of reiterant versions of the /t/ corpus, in which all syllables in the model sentences were instead pronounced as /na/. Both speakers were able to produce reiterant speech without difficulty.

The domains included in this study, along with the test sentences in the /t/ corpus, are listed in Table 1. The Syllable and Word domains differ from the previous study in that our test

[1] Department of Linguistics, UCLA
[2] Now at Nuance Communications

words are not reduplicated forms of the test syllables. The Intonational Phrase domain in our study is specifically defined as not set off by pause, but only by a large break. The Utterance domain in our study, as in Cho & Keating's Korean study (1999, this volume; also Keating et al. 1998 and this volume) is defined as an Intonational Phrase set off by pause; in Taiwanese it may also be marked by a small rise in pitch at the end. The other new domain in the present study is a small phrase, consisting of a heavy subject Noun Phrase. This domain is not tonally marked, but is characterized by a break greater than that between words.

Table 1. Corpus for /t/. Taiwanese forms are given in IPA plus punctuation. The test word is in boldface and the test syllable is underlined. Tones are marked only for the test syllable and the syllables on either side of it.

Domain	***Taiwanese***	***English meaning***
Utterance	wa u kʰuã-tiɤ papa^{55}. **ta^{23}ta^{55}** kʰai iaʔ be lai?	I can see Dad. Why isn't Tata here yet?
Intonational Phrase	wa kʰuã-tiɤ a! papa^{55}, **ta^{23}ta^{55}** kʰai iaʔ be lai?	I see it. Dad, why isn't Tata here yet?
Small Phrase	hit e laŋ e papa^{55} **ta^{31}-tiɤ^{31}** tsit-tsia katsuaʔ	That person's dad stepped on a cockroach.
Word	wa ka li kɔŋ, papa^{55} **ta^{31}-tiɤ^{31}** tsit-tsia katsuaʔ	Let me tell you, Dad stepped on a cockroach.
Syllable	wa kina kuã-tiɤ **ta^{33}ta^{33}** tsim^{53}a kɤ iŋ kiã.	Today I saw Auntie Tata and her child.

Speaker 1 read ten repetitions of each test sentence containing /t/ and six repetitions of the reiterant versions with /n/. Speaker 2 read fifteen repetitions of each test sentence containing /t/ and ten repetitions of the reiterant versions with /n/. Electropalatographic and audio signals were recorded with the Kay Elemetrics Palatometer, as in all our previous studies.

Peak EPG contact was measured over the whole palate: the percentage of electrodes contacted in each data frame was calculated, and the frame with the most contact was identified for each test consonant. Also, for each token, the number of data frames showing a complete stop occlusive seal was counted. This gives a measure of articulatory duration.

Two acoustic measures were also taken. The first was the VOT of /t/, as in previous studies. The second was a set of formant frequency measures: F1 and F2 were measured at the mid-point of the vowel following the test consonant, using LPC with a 25-ms window. Thus we could test whether the vowel of a domain-initial syllable shows any effect of its prosodic position.

Analysis of Variance was followed by Fisher's PLSD post-hoc tests in pairwise comparisons of the different prosodic positions, for each speaker and each consonant.

Results

Consonant Linguopalatal Contact

Taiwanese does in fact show phrasing effects like English, French, and Korean. First, prosodic position does affect the amount of contact, as seen in Figures 1-2. The overall effect of position on contact was highly significant for both speakers for both consonants. Generally, more overall contact was made by both speakers for both test consonants in the higher domains than in the lower domains. Differences in amount of contact across domains were more obvious for /t/ than for /n/, but most posthoc comparisons were significant at the .0001 level. Nonetheless, Speaker 1 distinguished only one pair of domains for /n/ (IPi vs. SPi, p < .0001), and did not distinguish Ui from IPi /t/s (p = .2721). Speaker 2 distinguished all four pairs of levels for /t/ and three for /n/, failing only to distinguish SPi vs Wi (p = .8157).

Figure 1. Peak EPG contact for /n/ when initial in five prosodic domains.

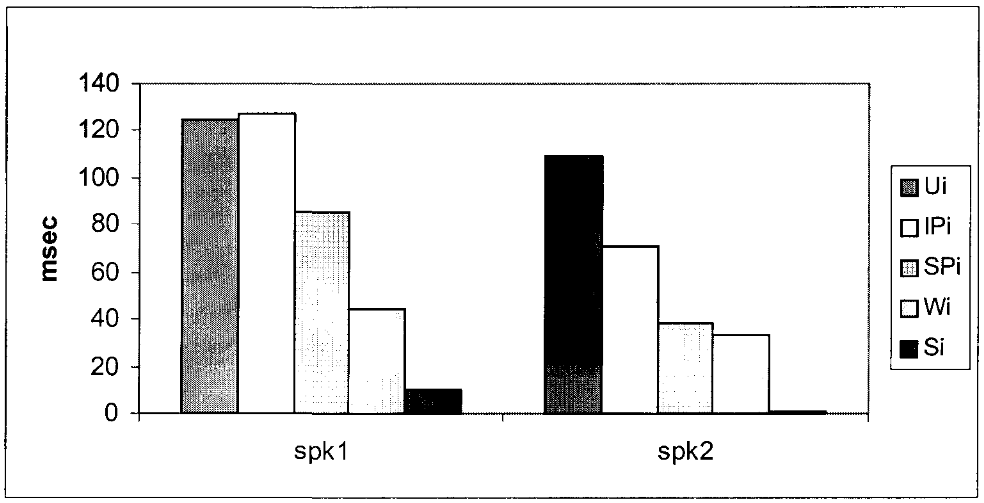
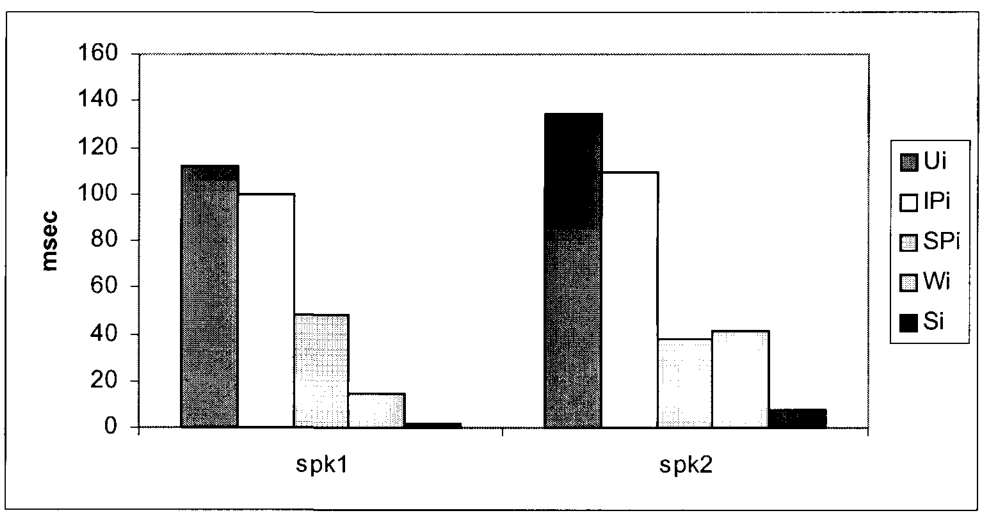


Figure 2. Peak EPG contact for /t/ when initial in five prosodic domains.Consonant Seal Duration



The articulatory seal durations, seen in Figures 3-4, show at least a two-way distinction, between lengthened higher prosodic domains (U-initial and IP-initial) and shortened lower domains (SP-initial, W-initial and S-initial). Both speakers make such a distinction for both test consonants. Whether there are additional reliable differences varies between the speakers, with the exception that both speakers distinguish between W-initial and S-initial for /t/. Speaker 1 failes to distinguish Ui from IPi for both consonants, while Speaker 2 fails to distinguish SPi from Wi for both consonants.

Figure 3. Duration of stop seal for /n/ when initial in five prosodic domains.

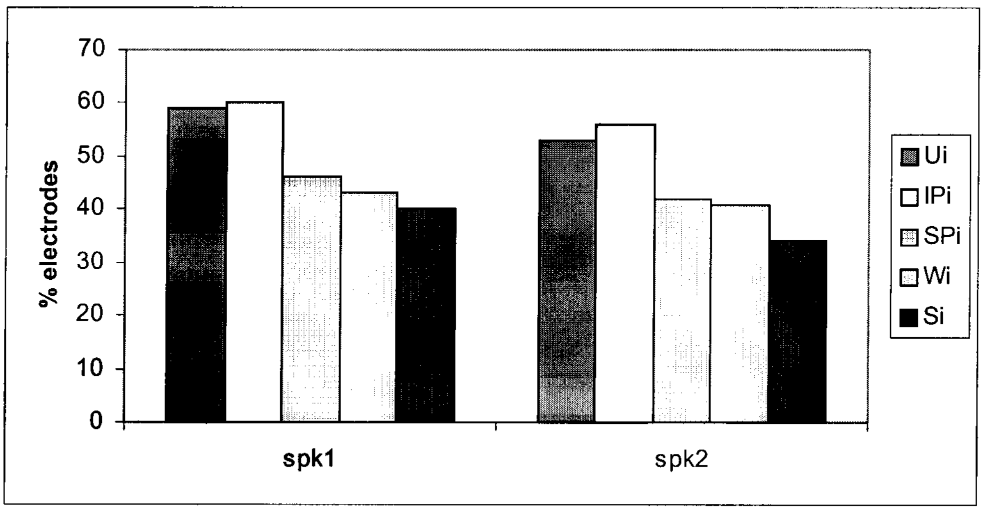
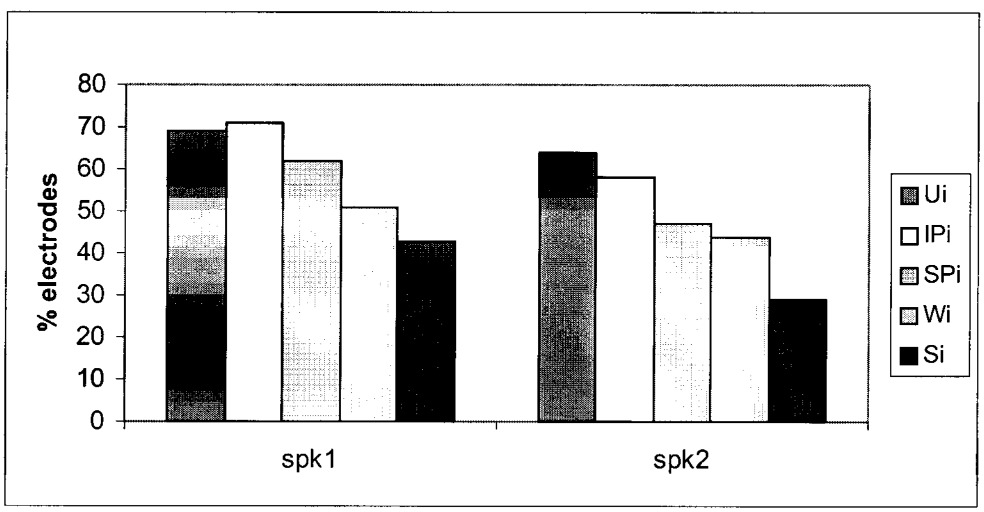


Figure 4. Duration of stop seal for /t/ when initial in five prosodic domains.



VOT of /t/

Similar to French, there was almost no reliable effect of prosodic position on VOT of the unaspirated /t/. Speaker 2 made the only distinction, which was between W-initial and S-initial.

Vowel formant frequencies

The formant data show no direct effect of prosodic position on F1 and F2 of both test consonants. There are no obvious patterns or tendencies in the formant data as with peak contact and seal duration. However, we do know that after a pause or utterance initially, the mouth tends to open slightly wider which causes the tongue the pull down and thus the F1 to increase. Speaker 2's F1 for /n/ in the U-initial position is significantly higher than the F1 values for the other domains, which shows that Speaker 2 has a more open mouth after a pause. This could explain why Speaker 2 generally makes less contact than Speaker 1.

Discussion

The results of this study are in accord with those of the studies of other languages. Linguopalatal contact varies systematically with prosodic position, with more contact in initial position of higher prosodic domains. Seal (closure) duration patterns similarly, and is furthermore strongly correlated with contact. However, neither VOT of the initial consonant, nor the formant frequencies of the following vowel, vary with prosodic position. The results for VOT agree with those from Korean and French, in that very-short-lag voiceless unaspirated stops seem not vary in VOT across prosodic positions.

This study resulted in more distinctions being made between initial positions in the different prosodic domains, than were found in the earlier study. The corpus in the present study was carefully between them as much as possible.designed to provide a larger set of prosodic domains, and to control the comparisons

Acknowledgment

This work was supported by NSF grant SBR-9511118 to P. Keating. We thank Moto Ueyama for help with ms. preparation, as well as our anonymous subject.

References

Chen, M. (1987). The syntax of Xiamen tone sandhi. Phonology Yearbook 4: 109-149.

Hsu, C.-S. and S.-A. Jun (1996). Is Tone Sandhi Group part of the prosodic hierarchy in Taiwanese? Poster presented at the Third Joint Meeting of the Acoustical Societies of America and Japan, Honolulu.

Hsu, C.-S. and S.-A. Jun (1998). Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic? UCLA WPP 96: 69-89.

Keating, P., T. Cho, C. Fougeron, and C.-S. Hsu (1998). Domain-initial articulatory strengthening in four languages. Paper presented at LabPhon6. Also in this volume.

Word-level asymmetries in consonant articulation

Patricia Keating, Richard Wright[1]
and Jie Zhang

I. Introduction

There has long been interest in whether and how syllable and word edges are marked phonetically, or, put conversely, how the position of a segment in its syllable or word affects its phonetic realization. For example, Pike (1947:162) noted:

> *"In many languages certain grammatical units - such as words - have as one of their characteristics the induction of sub-phonemic modification of some of the sounds."*

Pike went on to make a distinction between "modifiable" and other sounds. For "modifiable sounds", the grammatical boundary "becomes phonologically recognizable", that is, can be heard. For other sounds, "the boundary is not phonetically perceptible". Pike is surely correct that some sounds make boundaries easier to hear, and the focus of most research has been on such sounds. Regardless of whether such effects are perceptible, however, subtle boundary effects can be detected using acoustic and articulatory techniques of the phonetics laboratory.

Troubetzkoy (1949: Part II) also discussed the "delimitative function" in language: marking of morpheme, word, and sentence boundaries by "boundary signals". He considered these to be optional, though widespread ("it is possible to get along without them"). Languages range from those that rarely mark boundaries to those that mark them thoroughly; French is given as an example of the one extreme, and Tamil of the other. Troubetzkoy laid out several kinds of boundary signals. One division is between "positive" and "negative" signals: the presence of a sound or sound sequence may mark the presence, or the absence, of a boundary. Another, orthogonal, division is between phonemic vs. nonphonemic (allophonic) signals, e.g. whether a boundary is signaled by a distinctive feature; we can consider the nonphonemic signals to include any nondistinctive phonetic property. Yet another is between "individual" and "group" signals. The former arise from the presence of individual phonemes or allophones. Thus if a sound appears only adjacent to some boundary, it is thereby an individual, positive signal to the presence of that boundary. In contrast, (positive) "group" signals are seen when certain phoneme or allophone sequences can only occur across an intervening boundary, so that the combination itself serves as a boundary marker. For example, English sequences /θs, ðz, sθ, zð, tʃt, tʃs, ʃs, sʃ, dz/ "and very many others" are said to (positively) signal morpheme boundaries, while the German sequence /dl/ is a negative signal for word boundaries (that is, it only occurs word-medially). Troubetzkoy noted, however, that in languages with many loan words, such as English, phonemic boundary markings are statistically weakened and therefore less reliable. One might hypothesize that in a language with reliable phonemic boundary signals, marking by means of nonphonemic signals would be less important, while in a language with less reliable boundary signals, phonetic marking would be more favored.

Most subsequent research has been concerned with positive individual nonphonemic signals, i.e. allophones that occur adjacent to some boundary. For example, Lehiste (1960, 1961, 1964) conducted a classic set of experiments on allophonic cues to syllable, morpheme, and word

[1] Now in the Linguistics Department, University of Washington

boundaries. ("Internal open juncture" refers to a boundary inside an utterance.) She recorded spoken minimal pairs differing in juncture and analyzed them acoustically. Her studies showed that boundaries can be detected from "marginal allophones" and/or "the overall intensity and duration pattern" (1960:42). Initial allophones are longer than medials or finals; marginal allophones involve such phonetic properties as aspiration, flapping, or formant differences. However, the studies did not yield acoustic differences between different types of boundaries; morpheme boundaries were phonetically marked only if they were also "lexical word" boundaries.

Several studies have followed up on Lehiste's results on the acoustic duration of segments at a boundary (e.g. Klatt 1975). Some studies, like Lehiste's, have focused on the effect of changing the placement of a word boundary within a VC(C)V sequence. A recent example is Quené (1992), who showed that in Dutch, a word-initial C is longer than a word-final C, and that furthermore, lengthening of an initial consonant can serve as perceptual cue to Dutch word boundaries. Other studies have compared the acoustic durations of the initial vs. final Cs in a word. A recent example is Berkovits (1993), who showed that in Hebrew, the word-initial C is longer than the word-final C in phrase-medial position. However, she also found that in phrase-final position, pre-boundary lengthening results in a word-final C that is longer than the word-initial one.

In addition to acoustic duration, there are several other ways in which production of segments has been shown to vary according to the location of a word boundary. Many studies have compared the articulation of segments in word-initial vs. word-final position. These studies have spanned several articulators and several experimental techniques. They include: Rousselot (1901), Fromkin (1965), McGlone & Proffit (1967), McGlone et al. (1967), Fujimura and Sawashima (1971); Kohler and Hardcastle (1974), Benguerel (1977), Fujimura (1977), Hardcastle and Barry (1985), Vaissière (1988), Macchi (1988), Krakow (1989/1993), Byrd (1994/1996), Browman and Goldstein (1995). Fougeron (1998) summarizes the articulatory literature as showing that "In initial position, consonants have a glottal opening gesture that is longer and larger; vowels are preceded by a glottal stop or are glottalized. In initial position, consonants and vowels have greater muscular activity. At the level of the velum, initial oral and nasal consonants have a more elevated position of the velum. At the level of the tongue, consonants have a more elevated tongue position and greater pressure of the tongue against the palate. For vowels, the few recent spectral data suggest that vowels can have a more peripheral tongue position at the beginning of a Word in certain languages" [our translation]. Of the studies cited above, only Macchi (1988) provides a clear counterexample to these generalizations: in an analysis of X-ray microbeam recordings of short sentences and nonsense phrases with /p/ in different positions, she found no effect of position-in-word on lower lip movement, and effects on jaw height in opposite directions for her two speakers, though initial /p/ was longer than final /p/.

Although a word-level asymmetry thus seems fairly well established, the research to date raises several fundamental questions. One is the extent to which the *word* is really the relevant domain. In a CVC word, the first C is initial in both a syllable and a word, and the second C is final in both a syllable and a word. While some studies of CVCs describe the object of study as effects of position-in-word, and others as effects of position-in-syllable, in fact these effects are usually confounded. Moreover, as Gow et al. (1996) point out in their review, other effects are often confounded in studies of position-in-word: most notably, effects of lexical stress and phrasal position. Few studies explicitly compare position-in-word with position-in-syllable, at the same time controlling other variables. Krakow (1989), in a study of lip and velum movements in /m/, constructed materials which did compare these positions. Her results showed no difference between syllable-initial vs. word-initial /m/, but she did find differences between syllable-final and word-final /m/: the labial and the velic movements were longer in duration and greater in displacement word-finally. However, the prosody of the test items was not always tightly controlled. Fougeron and Keating (1997) compared articulation of word-initial with syllable-initial /n/s, and found that two of three speakers had reliably more linguopalatal contact in word-initial position, but they did not look at coda consonants.

Another question is the robustness of the effect across prosodic positions. As noted above, Berkovits (1993) found that in Hebrew, phrase final lengthening resulted in a word final consonant that was longer than a word initial one. If the same is true for English, then initial lengthening might be found only in some prosodic positions, and only there would a word-level asymmetry be seen.

Finally, as Pike suggested, some consonants seem more prone to exhibit boundary effects than others. Macchi's and Krakow's results described above suggest that effects are weak with labial consonant articulations. Browman and Goldstein (1995), in an Xray microbeam study of the peaks of consonant gestures (lip aperture, tongue tip height, tongue dorsum height) showed that boundary effects are much larger for /t/ than for /p/ or /k/; Byrd et al. (1996), in an EMA study of Tamil, found more significant effects of boundary on /n/ than /m/. A consonant that seems relatively unaffected by position is /s/. McGlone et al. (1967), measuring lingual pressure behind the incisors, found a difference in palatal pressure for /t d n l/ but not /s/, and Byrd (1994/1996), in an EPG study of linguopalatal contact, found a durational but not a spatial effect on /s/. However, no study has examined a large enough set of consonants to compare across manner and place.

Thus, although a word-level asymmetry in articulation is in one sense well-documented, at the same time the asymmetry is not always found and its nature is not clear. In this paper we investigate some factors that might seem to make the asymmetry less than universal. First, we consider a range of consonants, to see whether only some classes of consonants show the asymmetry. Second, we consider how the word-level asymmetry interacts with phrase-level effects on articulation. Third, we briefly consider whether a word-level asymmetry can be distinguished from a syllable-level asymmetry.

II. General Method Across Experiments

The three experiments were carried out in separate sessions. Additional studies on other topics, not reported here, were also included in these sessions. Except for the speech materials, which are described separately for each experiment below, the methods are essentially the same over the three sessions.

1. Electropalatography

The articulatory measure used in these studies is linguopalatal contact, that is, contact between the tongue and the palate, obtained by electropalatography (EPG). EPG data was collected with the Kay Palatometer, which uses custom-fitted pseudo-palates with 96 contact electrodes covering the entire hard palate and the inside surfaces of the molars. Contact information from these electrodes is sampled every 10 ms. The speech signal is recorded simultaneously into a computer file at a 12.8 kHz sampling rate.

2. Subjects

A total of six subjects served in these experiments, with five in the first, four in the second, and two in the third. All were linguists; one was the first author, the others were UCLA faculty, students or former students. All but two had participated in previous EPG studies (Byrd 1994 and/or Fougeron and Keating 1997), but all except the author were naive about the specific goals of the experiments.

3. Procedure

Test utterances were digitized direct to disk in Kay Elemetrics's CSL format. Subjects wore both their pseudo-palates and a head-mounted close-talking microphone. They were cued by the experimenter to read each test item one or more times from a printed list, and then paused while the experimenter saved the file to disk. Test items were generally repeated a total of 9 times; specific experiment details are given below.

Control items with labial consonants (instead of test consonants) were recorded in each experiment. These items revealed the contact patterns of the context vowels.

4. Data processing and analysis

The main measure used was peak (maximum) consonantal linguopalatal contact. Consonantal contact is defined with reference to those electrodes not contacted during the control (labial consonant) items. Thus the first step of analysis was to identify those electrodes contacted during the control consonants -- these are taken to be characteristic of the vowel's contact pattern. Those electrodes were then excluded from analysis of the consonants' contact patterns. The remaining electrodes may be considered to form a non-vowel region. Because different experiments used different test vowels, the non-vowel region had to be defined separately for each experiment.

The percentage of electrodes contacted in this region, over time, was computed for each file. The maximum contact during the test consonant(s) was then identified. Where other measures were made, they are described below, for the particular experiment.

III. Experiment I: consonant differences

1. Methods

The first experiment compares a variety of obstruent consonants (tʃ, dʒ, t, ʃ, ʒ, s, z, d, and k) in word-initial vs. word-final position in CVC test words. These are all the obstruent consonants of English (except /g/) for which linguopalatal contact can be registered for the consonantal constriction by our Kay pseudo-palates. The other English obstruents, which are labiodental or interdental, cannot be studied using EPG. The speech materials are shown in Table 1. Each CVC test word contains only one test C (either initial or final); the other C is always /b/. For the coronal test consonants, shown in Table 1A, the vowel in the test word is always /ʌ/ (a vowel which for our speakers of American English seems to involve little if any linguopalatal contact). Thus, for example, the test word for word-initial /t/ was *tub*, while the test word for word-final /t/ was *but*. As can be seen in Table 1, most of the resulting test words are real words, but a few are nonsense words. For the test consonant /k/, shown in Table 1B, a front vowel context was needed to produce a fronted velar with linguopalatal contact on the EPG pseudopalate. Therefore for these test items the vowel in the test word was /ɪ/. Thus the test words for /k/ are *kib* and *bic*. Unfortunately, there is no way to guarantee that all of the contact for a /k/ will take place on the pseudopalate. In our experiment, all subjects show a stop seal on the palate for all the tokens of /k/; however, only some tokens have all of this seal contact visible (i.e. in front of the backmost row of electrodes). Therefore, the results for velar consonants cannot be interpreted with as much certainty.

Table 1. Speech materials for Experiment 1. Materials for coronal and velar test consonants are shown separately. In the "carrier frame" column, CVC is a placeholder for the test words, and boldface indicates phrasal stress (nuclear accent).

Expt.	**test C**	**test words**	**carrier frame**
1A	t	bʌt, tʌb	"A CVC CVC **again**." (same CVC, equal stresses)
	d	bʌd,dʌb	
	s	bʌs, sʌb	
	z	bʌz, zʌb	
	ʃ	bʌʃ, ʃʌb	
	ʒ	bʌʒ, ʒʌb	
	tʃ	bʌtʃ, tʃʌb	
	dʒ	bʌdʒ, dʒʌb	
1B	k	kɪb, bɪk	as above

The test words were put in a carrier sentence "A CVC CVC ***again***.", with nuclear accent on *again* and near-equal stress, but no pitch accents, on the two CVC words. The two CVC words in a given utterance were the same, but only one of them was the test word in that utterance. For C in word-initial position, the second CVC was the test word (e.g. *tub tub*, while for C in word-final position, the first CVC was the test word (e.g. *but but*). This was done so that all the test utterances could be produced with the same prosody, while the segmental context remained as controlled as possible. In addition, following Byrd 1994, the CVC CVC design puts word-final Cs in an earlier position in the test utterance (CVC CVC) than the word-initial Cs (CVC CVC). Thus if there are any global, utterance-level trends in articulation (e.g. Krakow et al. 1994), these work against, not for, the hypothesis being tested. That is, if final C is reduced relative to initial C, it will not be because the final C came later in the utterance.

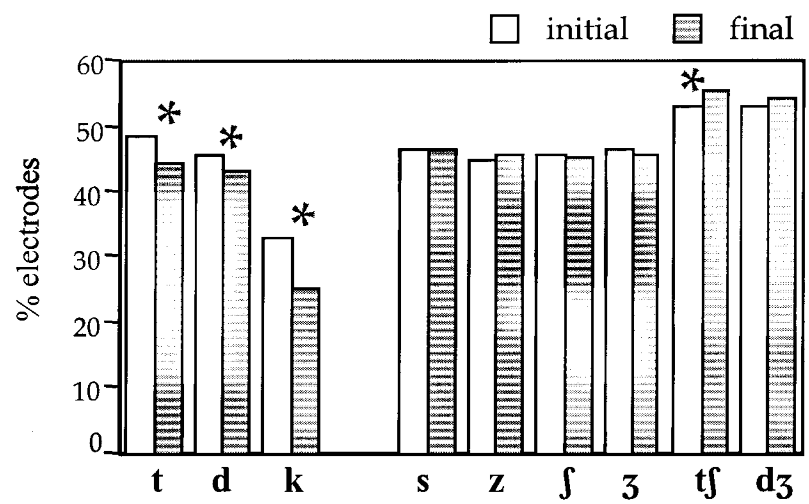
The test utterances were generally ordered so that all repetitions of an item were together, the initials for one consonant followed by the finals for that same consonant, and the consonants in the order listed above. However, one subject produced a different order, in which tokens were in blocks of three repetitions of each item at a time.

2. Results

Peak contact

Figure 1 shows the average (across speakers and repetitions) of the peak contact for each consonant in word-initial vs. word-final position. Note that for affricates, the peak contact was always found during the stop portion of the segment.

Figure 1. Peak EPG contact: word-initial vs. word-final stops and sibilants in Experiment 1. Data are pooled across subjects. Asterisks indicate significant differences within consonants.



This figure represents the interaction of factors Position and Consonant in a 2-way ANOVA, which is the comparison of interest for present purposes. This interaction was reliable at the .0004 level ($F(7) = 3.888$). Posthoc Fisher's PLSD tests showed that only stops /t/ and /d/ have significantly more contact in word-initial position than in word-final position, while affricate /tʃ/ has significantly more contact in word-final position than in word-initial position (for /t/, $p = .0004$, for /d/, $p = .0048$, for /tʃ/, $p = .0176$). Not surprisingly, then, in this analysis there was no main effect for Position ($F < 1$): across the nine consonants as a group, initial and final positions show the same amount of contact.

The velar stop /k/, analyzed separately, shows the same word position effect as /t/ and /d/: by paired t-test, initial /k/ has more contact than final /k/ ($t(46) = 8.041$, $p < .0001$). As noted above, not all the velar contact is captured in the EPG data. Thus it is possible that the initial /k/ shows more contact because it is articulated more forward on the palate (and thus more of its contact is on the pseudopalate). (A more fronted articulation would suggest more coarticulation with the following front vowel.)

It should be noted that when individual speaker data is analyzed there are some differences across subjects. Not all subjects show all of the effects summarized above (which hold for the subject group as a whole). For example, only 2 subjects make significantly more contanct for initial /t/ than for final /t/; only 3 subjects do so for /d/, and 4 for /tʃ/. Four subjects show a difference in contact for /dʒ/ according to its position, but they are evenly split as to the direction of that difference.

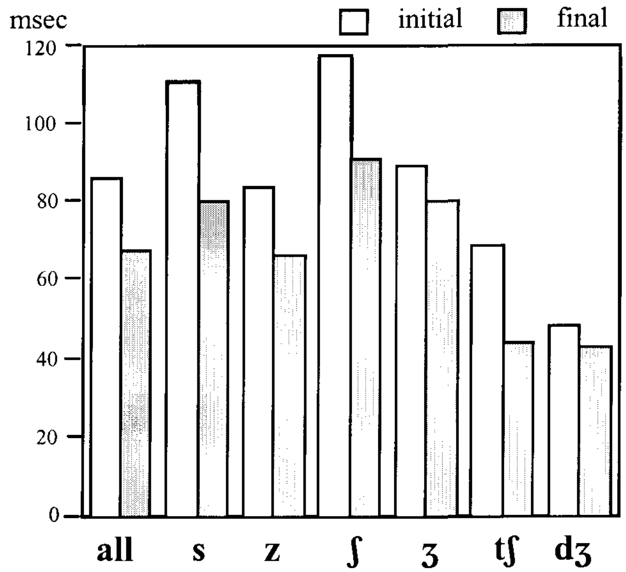
Segmental context

A follow-up analysis of these contact data addressed the following concern. Recall that the consonants are in $CVC_f\ C_iVC$ sequences. This controls for (indeed, works against) any effect of sequential declination; but it does put the test consonants into different segmental contexts. The final-C is preceded by a vowel, while the initial-C is preceded by a labial consonant. It is conceivable that a preceding vowel, with its open vocal tract, could pull open the consonant that follows it, more than a preceding labial consonant would. If that were so, then our asymmetry would not be due to position-in-word, but to the preceding segmental context. The corpus allows a check of this: the "non-test" word in each sentence provides a copy of the test consonant, but in the converse context. Thus in *a tub tub **again*** the first /t/ is an onset after a vowel, and this /t/ can be compared to our two test /t/s. Such a comparison was carried out for all nine consonants by ANOVA, but the relevant comparisons are for the four consonants already found to have a significant effect of Position-in-Word on contact, i.e. /t d k/, plus /tʃ/ in the reverse direction. All four of these consonants showed a significant main effect for the three-way comparison (for /t/, $F(2)=32.7$, $p<.0001$; for /d/, $F(2)=25.7$, $p<.0001$; for /tʃ/, $F(2)=5.87$, $p=.0036$; for /k/, $F(2)= 3.75$, $p=.026$). For these consonants, this ANOVA was followed by pairwise posthoc comparisons. For /t/ and /d/, contact is reliably the largest for the initial consonants in these non-test words (that is, after a vowel), with $p < .0001$ in each pairwise comparison. For /tʃ/, the non-test initial consonant has more contact than our test initial consonant ($p = .0014$), but the non-test initial and the final consonants are not reliably different. Finally, for /k/ the non-test initial consonant is not reliably different from either of the test consonants. These results suggest that our initial vs. final effect is generally due to position in word, not segmental context. (Why three onset consonants have more contact than our test onset consonants is a separate question, but given the clear negative results in Fougeron & Keating (1997), we doubt that it is due to declination. More likely, the first syllable tended to have a slightly greater stress than the second.)

Further analysis of fricatives

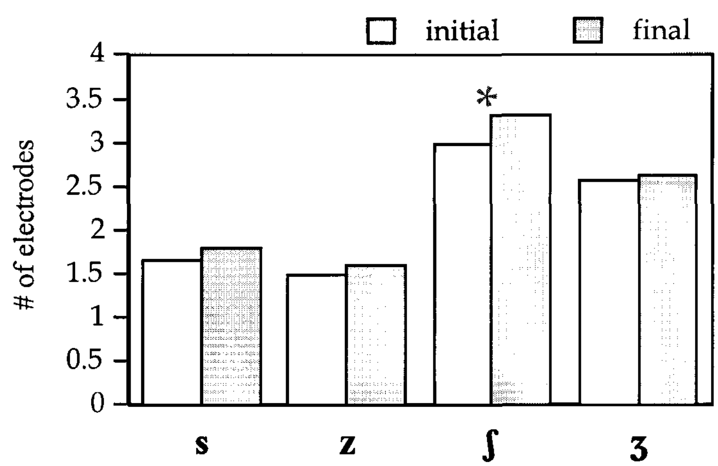
Because these analyses of peak contact during the consonant revealed no general effects of position-in-word on the four fricative consonants, two further analyses of these were performed. The first was of the acoustic duration, which was taken to include the interval of noise plus any silent gap next to the vowel (as is often seen after sibilants). For this analysis, the fricative components of the two affricates were also included (making a total of six consonants: /s z ʃ ʒ tʃ dʒ/. The results are shown in Figure 2. A 3-way ANOVA with factors Consonant, Position, and Speaker showed significant main effects for all of these factors; thus initial fricatives are longer than finals ($F(1) = 621.325$, $p < .0001$). All five speakers, when analyzed individually, show these main effects of Position and Consonant. However, the particular comparison of interest is the interaction between Consonant and Position, and this was also significant ($F(5) = 30.562$, $p < .0001$). Posthoc analysis shows that initial fricatives are longer than finals for all of these except (the fricative portion of) dʒ (for /s/, /z/, /ʃ/, and /tʃ/, $p < .0001$; for /ʒ/, $p =.0034$; for /dʒ/, $p = .0529$). The reason that there is only this marginal effect for /dʒ/ when speakers are combined, is that the speakers differ in how this consonant behaves: some have a difference in one direction, some in the other. For /ʒ/, one of the speakers differs from all the others in having longer final /ʒ/s, but for other consonants the individual speakers are quite similar.

Figure 2. Word-initial vs. word-final sibilant duration. Data are pooled across subjects. The first pair of bars shows the average effect for the six sibilants combined.



Because position does affect fricative duration, a finer analysis of the EPG record was performed for the four fricatives /s z ʃ ʒ/: the size of the fricative air channel was measured. The data frame with peak consonant contact was displayed for each token, and the fricative air channel was identified in that display. Length of the fricative channel and its width at its narrowest point were measured in #electrodes. Figure 3 shows the effect of position on the width of the fricative channel for each consonant. Across the five subjects, the fricative /ʃ/ does show an effect of Position, with longer but narrower channels when in initial position. No other fricative showed any effect.

Figure 3. Channel width in word-initial vs. word-final sibilant fricatives. Data are pooled across subjects. Asterisk shows significant difference.



IV. Experiment II: phrasal effect

1. Methods

Experiment II was a large experiment in which three experimental variables are of interest here. The speech materials are shown in Table 2. First, four different test consonants were used, all non-continuants: /t d n l/. Second, each test consonant occurs word-initially and word-finally, with each test word containing two test consonants (the initial and final consonants in the word). Third, each test word occurs either first or last in its sentence, which is a prosodic Intonational Phrase (IP). Thus the word-initial consonants occur IP-initially or IP-internally, and the word-final consonants occur IP-internally or IP-finally. However, because these sentences are set off by pauses, they are also Utterances (IPs with break index 5). In what follows, these will be called simply "utterances".

Table 2. Speech materials for Experiment 2. In the "test words" column, an acute accent indicates lexical stress, and underlining indicates the test consonants. In the "carrier frame" column, "word" is a placeholder for the test words, and boldface indicates phrasal stress (nuclear accent)

test C	**test words**	**carrier frame**
t	tɛ́bəbɛ̀t, tɛ̀bəbɛ́t	**Word** fed them. / One deaf **word.**
		Word fed **them.** / **One** deaf word.
d	dɛ́bəbɛ̀d, dɛ̀bəbɛ́d	
n	nɛ́bəbɛ̀n, nɛ̀bəbɛ́n	
l	lɛ́bəbɛ̀l, lɛ̀bəbɛ́l	

The experiment manipulated other factors, not part of the present report, in addition to the identity and position of the test consonant. These factors are the location of the lexical stress of the test word (each test consonant occurs in a syllable with primary lexical stress, or secondary lexical stress), and the location of the phrasal stress, or accent, of the test sentence (each test word occurs accented or unaccented). The effects of these factors will be reported elsewhere (Keating, Cho, and Wright in preparation).

Pilot work showed that the vowel /ɛ/ (unlike the vowel /ʌ/ used in the first experiment) has relatively similar contact patterns under different degrees of stress and accent, and therefore the test words contain that vowel.

Subjects were asked to accent only one word. If they do that sucessfully, then the entire sentence must be produced as a single Intermediate Phrase, as well as a single Intonational Phrase/utterance. For example, in the test sentence ***Tebabet** fed them.*, only the boldfaced testword **tebabet** was to be accented, with the following two words therefore post-nuclear and within the same Intermediate Phrase. However, post-hoc transcription of the test utterances, as actually produced by the subjects, showed that sometimes the prosody was different from requested. Of relevance for present purposes, one subject made a small phrasal break (an Intermediate Phrase, with a L- boundary tone and break index 3) after about half the test words in the ***Testword** fed them* utterances. As a result, about half of her word-final test consonants are also Intermediate Phrase-final. Thus in her data comparisons can be made between Word-final, Intermediate Phrase-final, and utterance-final consonants, though the numbers of tokens are unbalanced.

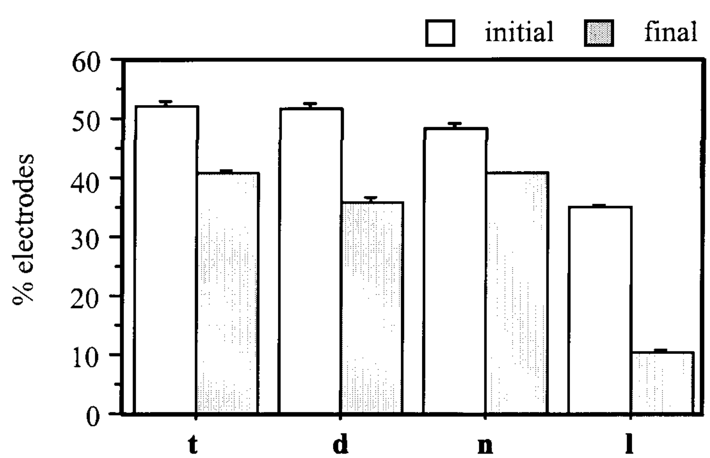
The test utterances were generally ordered so that all repetitions of an item were together, with the items for each test word in sentence-initial position followed by the sentence-final versions for that same word. However, one subject produced them in blocks of two repetitions of each item at a time. Four subjects (all of whom also participated in Expt. I) recorded the test materials.

Peak contact was measured for each test consonant as described for the previous experiment.

2. Results

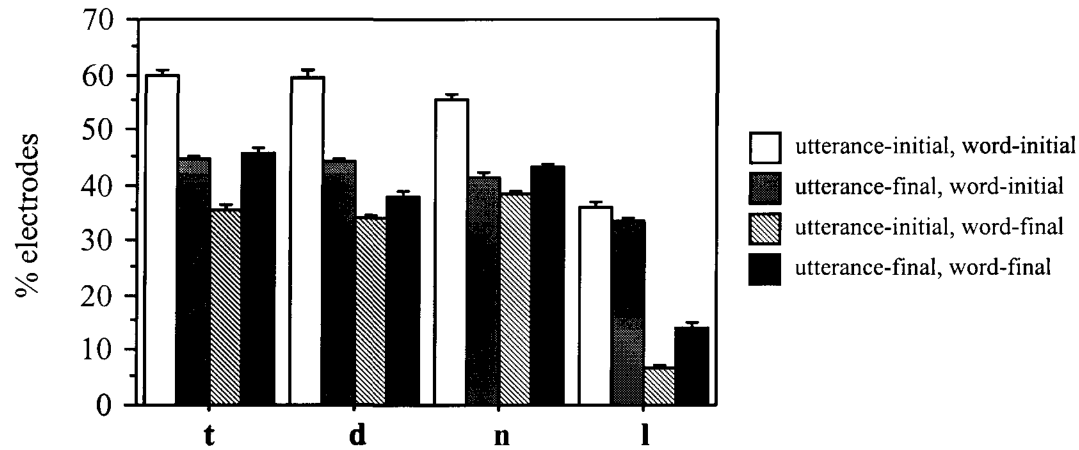
Statistical comparisons were performed by ANOVA. As can be seen in Figure 4, position-in-word affects contact, with initial consonants having more contact than finals ($F(1) = 1231$, $p < .0001$). These differences are reliable for all four test consonants, though it is weakest for /n/ and strongest for /l/ (for /t/, $F(1) = 141.9$, $p < .0001$; for /d/, $F(1) = 297.1$, $p < .0001$; for /n/, $F(1) = 83.9$, $p < .0001$; for /l/, $F(1) = 1539.5$, $p < .0001$).

Figure 4. Peak EPG contact in Experiment 2: word-initial vs. word-final /t d n l/. Data are pooled across position-in-utterance, subjects, and other prosodic variables. All four within-consonant comparisons are significantly different.



However, the main effect for position-in-word is not the best test for a word-level asymmetry, because it confounds the different positions in the utterance of the word-initial and word-final consonants. The interaction of Position-in-word and position-in-utterance is shown in Figure 5. When a word-initial consonant is also utterance-initial, then it is in absolute utterance-initial position; but when a word-final consonant is "utterance-initial", it is merely at the end of the first word in the utterance; conversely, when a word-final consonant is also utterance-final, then it is in absolute utterance-final position; but when a word-initial consonant is "utterance-final", it is merely at the beginning of the last word in the utterance. This confounding turns out to be important because consonants in absolute utterance-initial position have the most linguopalatal contact. Averaged across consonants and speakers, the most contact is seen for consonants which are utterance-initial and word-initial; the next-most contact is seen for consonants which are word-initial but utterance-medial; the next for consonants which are utterance-final and word-final; and the least contact is seen for consonants which are word-final but utterance-medial. All subjects show this pattern.

Figure 5. Peak EPG contact in Experiment 2: /t d n l/ showing interaction of position-in-word and position-in-utterance.



Deviations from this general pattern for individual consonants can also be seen in Figure 5. While for all consonants utterance-initial&word-initial consonants have the most contact, while utterance-medial&word-final consonants have the least, exactly how the other two categories pattern relative to these, and which differences are statistically significant, depend on the consonant.

In Expt. I all consonants were utterance-medial. Such a comparison can be made in the present experiment as well, if the word-initial consonants are taken from the utterance-final words and the word-final consonants are taken from the utterance-initial words. On average, this difference in contact is about 12 percentage points. This difference is statistically reliable by a post-hoc comparison ($p < .0001$). It is also statistically reliable for each consonant individually (for /t/, /d/, and /l/, $p < .0001$; for /n/, $p = .002$).

Another comparison picks out the utterance-edge consonants: word-initial&utterance-initial vs. word-final&utterance-final. Because the utterances were set off by pauses, the contact for the utterance-initial and utterance-final consonants is likely influenced by the pausing. In this comparison the word-level asymmetry is still clearly seen; on average it is larger, the difference in contact being about 17 percentage points. This difference is statistically reliable by a post-hoc comparison ($p < .0001$). It is also statistically reliable for each consonant individually (for all consonants, $p < .0001$). The asymmetry holds for all four subjects individually, but the size of the effect is variable, and for two subjects this effect is in fact smaller than the utterance-medial one.

The comparison which makes the word-level asymmetry greatest is the one in which the word-initial consonant is also utterance-initial, while the word-final consonant is utterance-medial (i.e. in the utterance-initial word). On average that difference in contact is 24 percentage points. This difference is statistically reliable by a post-hoc comparison ($p < .0001$). It is also statistically reliable for each consonant individually (for all consonants, $p < .0001$). Utterance-initial consonants generally have much more contact that other word-initial consonants, perhaps as noted above because they are post-pausal, utterance-initial. The effect in this comparison is to exaggerate the word-level asymmetry.

In contrast, the asymmetry is much weaker when we compare the utterance-final consonants to the word-initial, but utterance-medial, consonants. Utterance-final consonants, also adjacent to a pause, usually have somewhat more contact than other word-final consonants. This extra boost minimizes the advantage of the word-initial consonants. Across all the consonants, the difference is about 6 percentage points and is statistically reliable ($p < .0001$), but not all consonants show this effect. For /d/ the assymetry is small but is still reliable ($p < .0001$); for /t/ ($p = .255$) and /n/ ($p = .129$) there is no asymmetry (they have the same amount of contact in these two positions). However, /l/ maintains the more general pattern; this consonant shows very little effect of position in phrase ($p < .0001$).

Thus the word-level comparisons, however they are done, generally replicate (for /t/ and /d/) the results of Expt. I, and extend them to additional non-strident noncontinuants (/n/ and /l/). However, phrasal position also plays a role, either enhancing or minimizing the word-level asymmetry.

For the one subject who makes clear Intermediate Phrase (ip) breaks after some test words, a three-way comparison is possible between coda consonants in Word-final, ip-final, and utterance-final positions. One-way ANOVA followed by posthocs showed that all three positions are significantly different from one another (W vs ip, $p=.028$; W vs utterance, $p < .0001$). Oddly, although utterance-final have more contact than W-final, the ip-final consonants have the least. That is, utterance-final > W-final > ip-final. This result remains to be studied systematically.

The data in Figure 5 also allow us to answer another question about the word-level asymmetry, at least for these English speakers and these consonants. How does the magnitude of the word-level asymmetry compare with the magnitude of (utterance-level) initial strengthening? The magnitude of the word-level asymmetry is seen by comparing word-initial with word-final (both utterance-medial); the magnitude of initial strengthening above the word is seen by comparing word-initial with utterance-initial consonants. We have already seen that the first difference is, on average, about 12 percentage points. The second is on average also 12 percentage points. Viewed as ratios, the effects are also very similar in size (1:1.4 vs 1:1.3). That is, whatever boost a consonant gets from being at the beginning of a word, is about equal to what it gets from being at the beginning of an utterance.

V. Experiment III: word vs syllable

1. Methods

The goal of this experiment was to compare the coronal consonant /t/ in word-medial vs. word-edge positions. A set of minimal pairs was constructed in which one member of the pair contains a word-initial or word-final /t/, while the other member of the pair contains a matching syllable-initial or syllable-final /t/. These are shown in Table 3. The following are some design considerations for the corpus, beyond the usual considerations for all EPG experiments: first, the syllabic positions of the word-medial /t/ must be unambiguous; second, all other prosody (especially, the lexical stress and the phrase-level accenting) on the test syllables must be matched for the pairs of test items, third, the test syllables in the matched pairs should occur in similar linear locations.

Two subjects (the author, plus one who had not participated in any previous studies) were recorded in this experiment. They produced the test items, in a variety of orders, 10 times each. Peak EPG contact over all 96 electrodes was measured for each test consonant. A set of paired t-tests was done for each pair of test items, for each speaker separately.

Table 3. Speech materials for Experiment 3. The test words are underlined. Phrasal prosody is not shown here but was matched across pairs.

word-initial /t/	*syllable-initial /t/*
This is a prep-tile.	This is a reptile.
Say a tune again.	Say attune again.
Say (a) tapwater.	Say attack dog.
word-final /t/	*syllable-final /t/*
The cat must fear the sunshine.	An atmosphere of sunshine.
He bit Matt.	the bitmap
I saw a fat man.	I saw Batman.
I cut most pears.	Use utmost care.
The bout put James in the hospital.	The output changed in the summer (months).

2. Results and Discussion

The two subjects showed quite different results for the various pairs of test items. The first subject made no more contact for word-initial /t/ than for syllable-initial /t/. Two test pairs showed no difference ($p > .1$), while *attack* had more contact than *a tapwater* ($p < .02$). The second subject did make more contact for word-initial /t/ in two pairs ($p < .01$), but *attune* had more contact than *tune* ($p = .02$). Thus for none of the three initial-consonant pairs did the two subjects show the same result. The same was true for the five final-consonant pairs. The second subject showed no differences in contact for four of the five pairs, while in *bout put* had more contact than *output* ($p < .008$). In this pair, the syllable-final items are consistently completely reduced (showed no contact at all), while the word-final items have a bimodal distribution of reduced vs non-reduced contact. Note that although *utmost/cut most* also has /t/ followed by a labial, none of those completely reduced, so the effect is specific to the pair *bout put/output* for this speaker. The first subject, on the other hand, showed a reliable difference in contact for four of the five pairs, but the direction of the difference is variable. The one pair which for this subject shows no difference, *output/bout put*, is exactly the case where the other made a difference. This subject does have one pair of this sort, in which one word is consistently reduced and the other shows a bimodal distribution, but it is for the pair *utmost/cut most* ($p < .004$), and it is the word-final *cut* that is consistently reduced.

Because these two speakers behaved so very differently, we did not pursue this experiment with other speakers. It seems clear that no highly general result about Word vs. Syllable edges is likely to emerge from such a study. For neither of these two speakers do the two domains have any consistent differential effect on consonant contact.

VI. Discussion and Conclusions

These experiments show that, as expected, some lingual consonants have more linguopalatal contact in word-initial, compared to word-final, position. However, the experiments have also clarified how this word-level asymmetry effect interacts with other influences on articulation.

First, not all lingual consonants show the effect. Of the set of obstruents tested in Experiment I, only /t/, /d/, and /k/ did; Experiment II added /n/ and /l/ to this set. These consonants share the property of being non-sibilant non-continuants. Most notably, there is no effect of position-in-word on the peak EPG contact of sibilant fricative articulations for most subjects. This result accords with and extends results of previous studies. Furthermore, for the sibilant affricate /tS/, the effect of position is in the opposite direction: finals have more contact. Nonetheless, fricatives, like the other consonants, are generally acoustically longer in initial position, and subtler effects are seen on the contact pattern for one of them, /S/, namely the shape of the fricative channel, which is narrower and longer in initial position.

Second, it was seen in Experiment II that the word-level asymmetry can be countered by an independent effect of position in phrase. When a word-final consonant is also final in an utterance, that consonant has more contact than otherwise, and for /t/ and /n/ this increase equalizes the word-initial and word-final contact. This increase in contact in utterance-final position accords with Vaissière (1988), who observed that for velum articulation, final position in a large phrase is a strong position. However, these phrasal effects are seen only for the stops, and are weak at best for /l/.

One point to be made from this result is that whether or not the word-level asymmetry will be found in a study of CVCs would seem likely to depend in part on the carrier phrase used. Thus, the asymmetry will be seen if no carrier phrase is used, because that puts the test consonants at the edges of the utterance; similarly if carrier phrases like "Say (word) again" or "(Word) is it" are used the asymmetry will be seen, because they put the word-final consonant in utterance-medial position; but if a carrier phrase like "Say (word)" is used, the asymmetry would probably not be seen, because only the final consonant is at the edge of the utterance.

Another point to be made from this result is that the word-level asymmetry must be independent of the phrasal effect. Thus we saw that the consonant /l/ shows the word-level asymmetry quite strongly, but not the phrasal effect. Furthermore, for those consonants that do show both effects, the experiments allow us to compare the size of the word-level asymmetry with the size of domain-initial strengthening above the word. Comparison of word-initial consonants with utterance-initial vs. word-final shows that the word-level asymmetry is about equal in size to the phrasal strengthening effect.

Finally, Experiment III showed no clear effects of position in word vs. position in syllable. Comparisons of initial positions and of final positions for two speakers showed no systematic or consistent differences. This result is somewhat surprising for the comparison of word-initial and syllable-initial consonants, since Fougeron and Keating (1997) did find more contact in word-initial position, but only one test "word" was used in their study, /nonono/. One of our subjects did show this effect, but only for two of the three word pairs tested. Thus we cannot tease apart the contributions of word and syllable to the word-level asymmetry. This result is somewhat different from that of Krakow (1989), who did find some reduction (both spatial and temporal) in syllable-final, compared to word-final, /m/; but is in accord with her failure to find any systematic differences between word-initial and syllable-initial /m/. It is quite possible that there are stronger effects of lexicalization and lexical frequency that dominate for each pair of words and each speaker, but that are not seen in the nonsense word data.

In sum, the word-level asymmetry in consonant articulation, by which word-initial consonants are stronger in articulation than word-final consonants, is robust for certain consonants, the non-sibilant non-continuants /t/ /d/ /k/ /n/ and /l/, but is reflected in greater acoustic duration for other word-initial consonants. However, the asymmetry can be lost in phrase-final position, where it is countered by phrase-final strengthening of coda consonants.

Acknowledgments

This work was supported by NSF grant SBR-9511118 to P. Keating. We would like to thank our subjects, including Matt Gordon and Bruce Hayes; Taehong Cho who helped with data analysis, and Cecile Fougeron for related collaborations.

References

Benguerel, A.-P. (1977). Velar coarticulation in French: a fiberscopic study. *J. Phonetics 5*: 145-158.

Berkovits, R. (1993). Utterance-final lengthening and the duration of final-stop closures. *J. Phonetics 21*: 479-489.

Browman, C. and L. Goldstein (1995). Gestural syllable position effects in American English. In F. Bell-Berti and L. J. Raphael (eds.) *Producing Speech: A Festschrift for Katherine Safford Harris*, 19-34. Woodbury, NY: American Institute of Physics Press.

Byrd, D. (1994). *Articulatory timing in English consonant sequences.* Ph.D. dissertation, UCLA (distributed as *UCLA WPP 86*)

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *J. Phonetics 24*: 209-244.

Byrd, D., A. Kaun, S. Narayanan, and E. Saltzman (1996). Phrasal signatures in articulation. To appear in *Papers in Laboratory Phonology 5.*

Fougeron, C. (1998). *Variations articulatoires en début de constituants prosodiques de différents niveaux en français*. Ph.D. thesis, Université Paris III - Sorbonne Nouvelle.

Fougeron, C. and P. Keating (1997). Articulatory strengthening at edges of prosodic domains. *JASA 106*: 3728-3740.

Fromkin, V. A. (1965). *Some phonetic specifications of linguistic units: an electromyographic investigation*. Ph.D. dissertation, UCLA (distributed as *UCLA WPP 3*)

Fujimura, O. (1977). Recent findings on articulatory processes. In Carré, Descout and Wajskop (eds.) *Articulatory modeling and phonetics*, 115-126. Grenoble: Groupement des Acousticiens de Langue Française.

Fujimura, O. and M. Sawashima (1971). Consonant sequences and laryngeal control. *Annual Bulletin of research Institute of Logopedics and Phoniatrics, University of Tokyo, 5*: 1-13.

Gow, D. W., J. Melvold, and S. Manuel (1996). How word onsets drive lexical access and segmentation: evidence from acoustics, phonology and processing. *Proceedings of ICSLP 96* (CD-ROM). Philadelphia.

Hardcastle, W. and W. Barry (1985). Articulatory and perceptual factors in /l/ vocalization in English. *Working Papers of the Phonetics Laboratory, U. of Reading, 5*: 31-44.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in connected discourse. *J. Phonetics 3*: 129-140.

Kohler, K. and W. Hardcastle (1974). The instability of final alveolars in English and German: proposal for an instrumental investigation. *Speech Communication Seminar 2*: 95-98. Stockholm, Sweden.

Krakow, R. (1989). *The Articulatory Organization of Syllables: A Kinematic Analysis of Labial and Velic Gestures*. Ph.D. dissertation, Yale University.

Krakow, R. (1993). Nonsegmental influences on velum movement patterns: syllables, sentences, stress, and speaking rate. In M. K. Huffman and R. A. Krakow (eds.) *Nasals, nasalization and the velum,* 87-116. San Diego: Academic Press.

Krakow, R. A., F. Bell-Berti, and Q. E. Wang (1994). Supralaryngeal declination: evidence from the velum. In F. Bell-Berti & L. Raphael (eds.) *Producing speech: A Festschrift for Katherine Safford Harris,* 333-353 (Chapter 23). Woodbury, NY: AIP Press.

Lehiste, I. (1960). An acoustic-phonetic study of internal open juncture. Supplement to *Phonetica 5*.

Lehiste, I. (1961). Acoustic studies of boundary signals. *Proceedings of the 4th ICPhS*, 178-187., Helsinki, Finland.

Lehiste, I. (1964). Juncture. *Proceedings of the 5th ICPhS*, 172-200. Munster, Germany.

Macchi. M. (1988). Labial coarticulation pattern associated with segmental features and syllable structure in English. *Phonetica 45*: 109-121.

McGlone, R. and W. R. Proffit (1967). Lingual pressure associated with speaker consistency and syllable variations. *Phonetica 17*: 178-183.

McGlone, R., W. R. Proffit, and R. L. Christiansen (1967). Lingual pressure associated with alveolar consonants. *J. Sp. and Hear. Res. 10*: 606-615.

Oller, D. K. (1973). The effect of position in utterance on speechsegment duration in English. *JASA 54*: 1235-1247.

Pike, K. L. (1947). Grammatical prerequisites to phonemic analysis. *Word 3*: 155-172

Quené, H. (1992). Durational cues for word segmentation in Dutch. *J. Phonetics 20*: 331-350.

Rousselot, P. J. (1901). Principes de phonétique expérimentale. Paris: Welter.

Troubetzkoy, N. S. (1971). *Principles of Phonology*. Translated from *Grundzüge der Phonologie*, 3rd edition, 1962, by C. A. M. Baltaxe. Berkeley and Los Angeles: University of California Press.

Vaissière, J. (1988). Prediction of velum movement from phonological specifications. *Phonetica 45*: 122-139.

An experimental study of vowel duration in phrase-final contexts in Japanese[1]

Motoko Ueyama

1. Introduction

One of the main effects of prosodic structure[2] on segmental duration in English and other languages is the well known phenomenon of *phrase-final lengthening*. This effect is usually interpreted as the lengthening of segments in the final position of prosodic constituents. The lengthening of segments in group-final position was observed for English in early studies such as Oller's (1973) experimental work on positional effects on segmental duration and Klatt's (1975) statistical investigation of variance in segmental duration. Later studies discovered that in several languages phrase-final lengthening reflects the strength of prosodic boundaries, i.e. segments tend to be lengthened to a greater degree in the final position of larger prosodic units (e.g. for English, Cooper & Paccia-Cooper 1980, Wightman et al. 1992; for French, Hirst & Di Cristo 1984, Pasdeloup 1990, Fletcher 1991, Fougeron & Jun 1998; for Korean, Jun 1995, Cho & Keating 1999; for Dutch, Cambier-Langeveld 1997).

Probably, the most comprehensive study on phrase-final lengthening was conducted by Wightman et al. (1992) in American English. Wightman and his collaborators investigated a large pool of speech data in order to answer a specific question: How many phrasal levels are distinguished by the lengthening of segments in the vicinity of phrase boundaries? As a working hypothesis, the following 7 break sizes were set:

0 = no prosodic break
1 = prosodic word boundary
2 = accentual phrase (AP) or minor tone group boundary
3 = intermediate phrase (ip) boundary
4 = intonational phrase (IP) boundary
5 = breath/pause within sentence
6 = sentence boundary

Three trained listeners individually labeled the breaks based on juncture judgments and tonal transcriptions. The duration of segments in phrase-final position was measured. The resulting values were normalized with respect to the overall speaking rate of the sentence and expressed in z-scores. The results showed that levels 1-4 were significantly differentiated while 0 vs. 1, 4 vs. 5, and 5 vs. 6 were not.

In Japanese, most evidence on the existence of the prosodic hierarchy comes from the studies of intonational patterns (Beckman & Pierrehumbert 1986; Pierrehumbert & Beckman 1988). However, it has rarely been investigated how the intonationally defined prosodic hierarchy reflects degrees of phrase-final lengthening. Boundary effects have been mainly investigated in studies

[1]I would like to thank Marco Baroni, Taehong Cho, Sun-Ah Jun, Patricia Keaing and Jennifer Venditti for helpful comments.

[2] Prosodic structure largely reflects syntactic structure, although the two are not isomorphic, as shown in Ferreira (1993)'s study.

whose goal was to model segmental duration in Japanese speech synthesis. These studies are based on multiple regression analyses of large-sized speech corpora. The deviations in duration from the overall mean duration are computed and the correlation with segmental duration of each postulated independent factor is calculated. Some earlier studies such as Sagisaka & Tohkura (1984) and Takeda et al. (1989) showed the existence of breath group-final lengthening and sentence-final shortening.

The existence of two more prosodic levels was later tested for vowels in Kaiki et al. (1990) and for consonants in Kaiki & Sagisaka (1992). In these studies, durational patterns in the final positions of four prosodic units were analyzed: sentence, breath group, accent group[3], and word. The results showed significant effects of sentence-final shortening and phrase-final lengthening for breath and accent groups, but neither lengthening nor shortening effect at the word level. Another important finding was that different parts of speech had a significant, although small, effect. Vowel duration is overall longer in content words than functional morphemes; however, some functional morphemes such as the coordinate conjunctive or topicalizer tend to be lengthened.

The findings of Kaiki et al. (1990) and Sagisaka & Kaiki (1992) are questioned by Campbell (1992). Campbell pointed a potential pitfall of statistical analyses based on large-scaled corpora:

> Sentence-final shortening and noun lengthening have been reported in the literature for Japanese... To what extent are these effect statistical artifacts of the particular phone and word distributions of a particular corpus? Unfortunately, until more and bigger corpora become available this will be difficult to answer, but even then there is the danger that language-specific peculiarities will be manifest in the data, that lead to generalisations beyond their scope. (p.418)

The speech corpus used by Kaiki et al. (1990) and Kaiki & Sagisaka (1992) consist of sentences from written texts such as newspapers and magazines. The distribution of parts of speech probably differs remarkably between written and spoken discourse, and such difference heavily affects the endings of prosodic units. For example, in written texts noun phrases (NP) obligatorily end with enclitic case particles and verbs often end with tense or aspect markers. However, these characteristics are not necessarily observed in spoken Japanese, where common case particles are dropped and additional verbal particles are attached to verb endings for pragmatic effects. Therefore, it is not safe to draw conclusions about boundary effects in Japanese by solely analyzing speech corpora based on written texts. To make this point, Campbell (1992) re-analyzed the same corpus used by Kaiki et al. (1992) and Kaiki & Sagisaka (1992) by normalizing duration values in z-scores. The results showed durational variations of /a/ depending on the voicing of onset stops. In general, the duration of /a/ was significantly shorter when it followed the voiceless stops, especially /t/, than when it followed the voiced stops. He also compared /a/ duration in the syllable /ta/ in sentence-final vs. non-sentence final contexts. The results showed some degree of shortening in the sentence-final position, but the difference due to sentence-final position was significantly lower than the phonetically-motivated difference due to the voicing distinction. The syllable or mora /ta/ commonly occurs in written texts as the past-tense marker *-ta*. Campbell points that "the shortening reported for sentence-final vowels results primarily from the

---

[3]The accent group was determined on the basis of tonal transcriptions and accent types. The accent group and the breath group were not explicitly defined. Here we assume that the accent group and the breath group in Kaiki et al. (1990) and Kaiki & Sagisaka (1992) correspond to the accentual phrase (AP) and the (sentence-medial) intonational phrase (IP), which is followed by a pause, of the present study.

predominance of the /t/-/a/ combination in this position in Japanese sentences, rather than any context-specific shortening effect operating from situational features alone (p. 416)". Also, the effects of types of parts of speech observed by Kaiki and his colleagues are questioned. Campbell's results showed that "segments aren't necessarily longer in nouns, but that nouns have more longer segments (p. 411)" such as geminates. Campbell also added that particles were more likely to undergo phrase-final lengthening effect. This point is clearly demonstrated by comparing the duration of the vowel /a/ in the syllable /wa/: its duration is longer in the topicalizer *-wa* than in other textual contexts. A similar asymmetry was shown for the syllable /ga/ in the comparison between the case particle *-ga* and other contexts .

2. Purpose of the Study

In light of Campbell's warnings, we need to re-investigate boundary effects on segmental duration in Japanese within different textual contexts from the ones that are most common in the corpora used by Kaiki and his colleagues. Remember that Kaiki et al. (1990) found three significant effects on vowel duration: sentence-final shortening and accent group-final and breath group-final lengthening. Will the same effects still be observed if prosodic units end with items other than functional morphemes common in written texts such as verbal (e.g. the past-tense marker-*ta*) or nominal morphemes (e.g. the topicalizer *-wa*)? How many prosodic levels are distinguished by the phrase-final effects? The main purpose of this study is to attempt to answer these questions on the basis of the analysis of controlled laboratory speech.

For the purpose of the production experiment presented here, we assumed the hierarchical levels of phrasal boundaries shown in Table 1, following the proposals of Pierrehumbert & Beckman (1988) and Venditti (1995). Note that the accentual phrase (AP) and the intonational phrase (IP) defined here correspond to the accent group and the breath group in the study of Kaiki et al. (1990), respectively.

Table 1: Four boundary levels in the present study

	Boundary levels	Labeling Criteria
AP =	accentual phrase (AP) boundary	the presence of L% boundary tone and phrasal H- tone on the second mora of AP
ip =	intermediate phrase (ip) boundary	F0 resetting before the next ip; no post-boundary pause
IP =	intonational phrase (IP)	F0 resetting before the next IP; the presence of a post-boundary pause and a boundary tone; delimited by an orthographic comma
S =	boundary between two independent sentences (S)	F0 resetting before the next sentence; presence of a post-boundary pause and a boundary tone; delimited by an orthographic period

We tested the following two hypotheses:

Boundary effects at different prosodic levels

Hypothesis 1: On the basis of what has been found for English in Wightman et al. 1992, we expect to find that three prosodic levels will trigger phrase-final lengthening effects in Japanese regardless of types of part of speech: AP < ip < IP = S.

Earlier research reports sentence-final shortening effects in Japanese. Campbell (1992) hypothesized that the sentence-final shortening, which was observed by Kaiki et al. (1990), was due to two main factors: the significant shortening effect of the onset plosive /t/ on the vowel /a/ within the same syllable /ta/, and the predominance of the past-tense marker *-ta* in written-text-based corpora. Taking this into account, we also predict that there will be no durational difference between a sentence-final vowel and a vowel in the final position of a sentence-medial IP, if the sentence and the sentence-medial IP end with the same syllable/mora.

Vowel duration in noun- vs. particle-ending conditions

Hypothesis 2 The earlier literature (e.g. Beckman 1992 for a review) has been discussing the tendency in Japanese to keep mora duration equal due to a general mora-timing constraint on temporal organization. We assume that this tendency is stronger in content words, where there is a need to preserve lexical identity. Taking this into account, we expect less durational elasticity in phrase-final position when the phrase ends with a content word (a noun in this study). Thus, a smaller overall variance in the duration of /a/ will be observed in the noun-ending paradigm than in the particle-ending paradigm.

3. Experiment

3.1. Corpus

The duration of the vowel /a/ in the final position of the words *obab<u>a</u>* # and *obaba-g<u>a</u>* # was compared. *Obaba* is a noun meaning 'grandma'; *obaba-ga* is the concatenation of the noun *obaba* and the case particle-*ga*. Each word was positioned at the end of a boundary between two phrases belonging to the same prosodic level. For each word the duration of the phrase-final vowel /a/ was compared for the four different boundaries (i.e. AP *vs.* ip *vs.* IP *vs.* Sentence).

Test sentences were designed in order to check how the duration of phrase-final /a/ differs depending on the types of boundary: AP, ip, IP and S. We tried to keep the sentences as similar as possible with respect to the other factors that have been reported to affect segmental duration in several languages (see Kaiki et al. 1990 and Kaiki & Sagisaka 1992 for Japanese): a) mora count from the beginning of the AP containing the target vowel /a/ up to that vowel (3 ~ 4 μ); b) mora count from the beginning of the utterance up to the targeted vowel /a/ (11 ~ 12 μ); c) utterance size indicated by the total mora count of the utterance (24 ~ 26 μ); d) type of segment that follows the target vowel /a/, (/m/ in this experiment).

The sentences used in the experiment are listed in Table 2. For each of the four boundary levels, two comparable sentences were constructed. The two sets are referred to as the particle- and noun-ending paradigms, respectively.

Table 2. Sentences used in the experiment. Duration of the vowel /a/ was analyzed in the underlined positions. The numbers of moras are indicated for each word; the total mora count of each utterance is also shown. An apostrophe marks a lexically accented mora.

a. before AP boundary

[[3 5 [4(3)]$_{AP}$ =12 (11) [5]$_{AP}$]ip [5 5] ip Total 26(25)
Too-i sinrui-no **[oba'ba-ga]**$_{AP}$ majimeda'-to Yu'richan-ga itte'ta.
Too-i sinrui-no **[oba'ba**-ø]$_{AP}$ majimeda'-to Yu'richan-ga itte'ta.
far relative-GEN grandma-SUBJ be serious-COMP (name)-NOM said
'Yurichan said that Grandma, a far relative, is serious."

b. before ip boundary

[[3 5 [4(3)]$_{AP}$ =12 (11)]ip [2 4 6= 12] Total 24(23)
[[uchi-ni tomatte-ru **[oba'ba-ga]**$_{AP}$]$_{ip}$ [ma'da gu'uguu nete-ru-rasi'i]$_{ip}$.
[[uchi-ni tomatte-ru **[oba'ba**-ø]$_{AP}$]$_{ip}$ [ma'da gu'uguu nete-ru-rasi'i]$_{ip}$.
house-LOC stay-PRES grandma-SUBJ only soundly sleep-PRES-I've heard
'Grandma who is staying at our place is still sleeping soundly, I've heard'

c. before IP boundary

[3 5 4(3)] = 12 (11) [7]=7 [7] = 7 Total 26(25)
[[[uchi-ni yatteku'ru **[oba'ba-ga]**$_{AP}$]$_{ip}$]$_{IP}$, [matagiki-da'kedo]$_{IP}$, [ijiwa'ru-rasii]$_{IP}$.
[[[uchi-ni yatteku'ru **[oba'ba**-ø]$_{AP}$]$_{ip}$]$_{IP}$, [matagiki-da'kedo]$_{IP}$, [ijiwa'ru-rasii]$_{IP}$.
house-LOC coming grandma-SUBJ hearsay-although nasty-I've heard
'The grandma coming to our place, although this is second-hand information, is nasty, I've heard.'

d. before sentence boundary

[8 4(3)]IP = 12 (11) [5 7]IP = 12 Total 24(23)
[[[koma'tteru-rasii **[oba'ba-ga]**AP]ip]IP]S. [ma'ri-san-mo soo-itte'-ta-yo]IP.
[[[koma'tteru-rasii **[oba'ba-ø**]AP]ip]IP]S. [ma'ri-san-mo soo-itte'-ta-yo]IP.
in trouble-I've heard grandma-SUBJ (name)-HON-also so-say-PAST-SFP
'Grandma is in trouble. Mari also said so.'

In spoken Japanese it is possible to end a phrase with a particle-free noun, since it is acceptable to drop an enclitic particle. This makes it possible to check boundary effects on segmental duration in the final position of noun-ending phrases. Also, a sentence can end with a NP because word order is relatively flexible. This allows us to compare the duration of phrase-final /a/ in *obaba* vs. *obaba-ga*, which can be both located in sentence-final position.

3.2. Procedure and Measurements

The prepared sentence sets were pseudo-randomized, and two female native speakers of Tokyo Japanese, CT and HY, read each sentence 11 times. The speakers were asked to pretend Their speech performances were tape-recorded in the sound attenuated room of UCLA Phonetics Laboratory.

The tokens of the first and last repetitions were not analyzed. The speech data were digitized at a 10kHz sampling rate, and the duration of the phrase-final vowel /a/ was measured using Pitchworks. For all measurements, waveform and wide-band spectrograms were used. F0 tracks were also examined in order to control tokens for phrasing variation. Tokens were not analyzed if a) there were hesitations or disfluencies; b) phrasing patterns were exceptional.

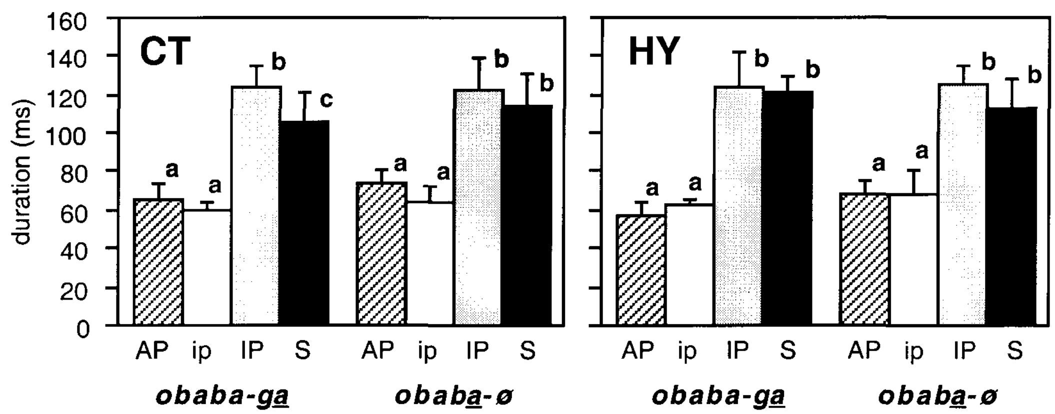
4. Results and Discussions

The measured values were statistically analyzed, testing the presence of the two types of effects on /a/ duration: a) boundary type effects (AP, ip, IP and sentence); b) morphological ending types (with vs. without *-ga,* the enclitic case particle).

4.1. Boundary effects at different prosodic levels

The means and standard deviations of /a/ duration of all relevant contexts are presented in Figure 1, where the four different prosodic levels (AP, ip, IP and S) are compared in the particle- and noun-ending conditions (i.e. *obaba-ga* vs. *obaba*):

Figure 1. Mean durations and standard deviations of phrase-final /a/ as a function of prosodic levels in the particle- and noun-ending conditions.



ANOVA results show that boundary strength significantly affects the duration of the phrase-final vowel /a/ ($p < 0.0001$ for both CT and HY). Scheffe's posthoc tests were additionally conducted in order to find how the four different prosodic levels are distinguished by phrase-final vowel durations within the particle- and the noun-ending paradigms ($\alpha = 0.05$). The grouping patterns of the four prosodic levels are indicated for each paradigm by letters above the bars of the graphs (i.e. a, b, c). In HY's speech, three general patterns are commonly observed within each morphological ending paradigm: a) AP and ip form one group, IP and S form another group; b) the mean difference between AP and ip is not statistically significant; c) the IP-final mean tends to be greater than the S-final mean, but the difference between the two means is not statistically significant. The patterns are identical for CT's data except that the IP-final mean duration is significantly greater than the S-final mean duration in the particle-ending condition (i.e. *obaba-ga*).

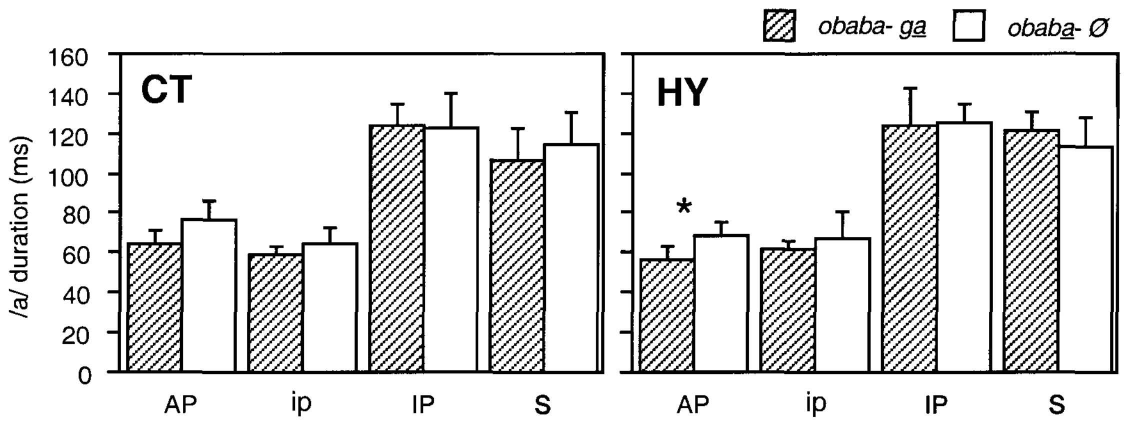
The first part of Hypothesis 1 states that the following hierarchy should be reflected in different degrees of phrase-final lengthening: AP < ip < IP = S. The results of the experiment show that duration means of the phrase-final vowel /a/ fall into two major groups at least, i.e. AP ≈ ip < IP ≈ S. Furthermore, between the two higher boundary levels, there tends to be a lesser degree of vowel lengthening in the sentence-final position than in the IP-final position. Thus, the pattern observed in our data can be summarized as AP ≈ ip < IP > S[4]. This pattern is different from the predicted pattern, AP < ip < IP = S. Thus, the first part of Hypothesis 1 is not supported.

The second part of Hypothesis 1 states that there should be no difference between the sentence-final and IP-final position. This statement is supported by the results of the experiment, since there is no significant mean difference between the two boundary types. However, we did find, as a general trend, that sentence-final vowels tend to be shorter than vowels in the final position of the sentence-medial IP. Kaiki et al. (1990) found a similar form of sentence-final shortening, and Campbell (1992) pointed that this might be due to the frequent occurrence of the past-tense marker *-ta* in the sentence-final position in their speech corpus. Since this marker is not included in the sentences of this experiment, the lesser degree of sentence-final lengthening found here cannot be explained by Campbell's conjecture.

4.2. Effects of phrase-final morpheme types: noun- vs. particle

For each boundary level, the mean durations of final /a/ are compared in the particle vs. noun final conditions (i.e. *obaba-g<u>a</u>* vs. *obab<u>a</u>*, respectively) in Figure 2. This comparison makes it possible to test whether there is a significant difference in phrase-final vowel duration depending on the morpheme types in phrase-final position, even though the comparison of the two ending conditions is not entirely valid due to the different onset consonants of the final syllables (/ga/ vs. /ba/).

Figure 2. The mean durations and standard deviations of phrase-final /a/ as function of phrase-final morpheme types within each boundary level. An asterisk indicates that a difference between the means of the particle- and noun- ending conditions (i.e. *obaba-g<u>a</u>* vs. *obab<u>a</u>*) is statistically significant according to the results of Scheffe's posthoc tests ($\alpha = 0.05$).



[4]A pattern similar to the one observed in our data (i.e. AP ≈ ip < IP > S) is found in the study of Venditti and van Santan (1998). They compared the mean durations of approximately 400 phrase-final short vowels which were corrected for the effects of factors other than the boundary type with a multpile regression method. Since I came to know the existence of their work after completing the present study, I did not discuss their paper here.

Hypothesis 2 predicts that a smaller overall variance in the duration of /a/ will be observed in the noun-ending condition than particle-ending condition due to the smaller durational elasticity of segments in content words for preserving lexical identity (in Japanese where short and long vowels are phonemic). The comparison of the standard deviations in the particle- vs. noun-ending conditions within each boundary level does not show any clear pattern for either CT or HY. Thus, Hypothesis 2 is not supported.

Figure 2 also shows a general trend for both speakers: /a/ durations are longer in *obaba* than in *obaba-ga*. In other words, phrase-final /a/ tends to be lengthened to a greater degree in the noun-ending condition than in the particle-ending condition. This pattern is followed in three out of four prosodic boundaries by both speakers: AP, ip and S boundaries in CT's speech; AP, ip and IP boundaries in HY's speech. However, this difference is only statistically significant in HY's speech noun-final /a/, which is longer than particle-final /a/ at the AP boundary, according to the results of Scheffe's posthoc tests.

Even though not statistically confirmed, there is a trend for longer vowel durations in the noun-ending condition in the data of both speakers. The trend may be caused by the different numbers of moras in the noun- and particle-ending paradigms. In a phonetic study of *tanka,* a Japanese meter, Homma (1991) found that the average duration of moras was slightly shorter in 7-mora lines than in 5-mora lines. This suggests that speakers speed up slightly to produce the 7-mora lines. In the experimental design of the present study, the particle-ending phrase has one extra mora/syllable due to the presence of the case marker *-ga*. Thus, it is possible that the durations of segments including phrase-final /a/ are overall shorter in the particle-ending paradigm, since the particle-ending phrase is produced at slightly faster rate than the noun-ending phrase. By taking all these elements into account, we conjecture that the observed trend for longer vowel duration in the noun-ending condition is not the expression of a systematic pattern.

Conclusion

The present study showed that in Japanese /a/ duration is significantly shorter before smaller phrase boundaries (AP, ip) than larger boundaries (IP and S). Even though the amount of the collected data is limited, we believe that this finding is valid because of the consistency of observed patterns across the two speakers. The data suggest that in Japanese an all-or-nothing distinction is made between smaller breaks and larger breaks in terms of magnitudes of vowel durations. At this point, we cannot tell whether AP- or ip-final vowels are lengthened or not, since we did not measure a reference vowel, such as a phrase-medial vowel. However, it is safe to conclude that the lower levels of the prosodic hierarchy are not affected by phrase-final lengthening as much as the higher levels in Japanese.

References

Cambier-Langeveld, T. (1997). The domain of final lengthening in the production of Dutch. In J. Coerts and H. de Hoop (eds.), *Linguistics in the Netherlands 1997.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Campbell, N. (1992). Segmental elasticity and timing in Japanese speech. In Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (eds.), *Speech Perception, Production and Linguistic Structure.* Tokyo: Ohmsha.

Cho, T. & P. Keating. (1999). Articulatory and acoustic studies of domain-initial strengthening in Korean. *This volume.*

Cooper, W. F. & J. Paccia-Cooper. (1980). *Syntax and Speech.* Cambridge, mass.: Harvard U. Press.

Ferreira, F. (1993). Creation of Prosody During Sentence Production. *Psychological Review* **100(2)**, 233-253.

Fletcher, J. (1991). Rhythm & final lengthening in French. *Journal of Phonetics* **19**, 193-212.

Fougeron, C. & S.-A. Jun. (1998). Rate effects on French intonation: Prosodic organization and phonetic realization. *Journal of Phonetics* **26**, 45-69.

Hirst, D. & Di Cristo, A. (1984). French intonation: a parametric approach. *Die Neueren Sprachen,* **83(5)**, 554-569.

Homma, Y. (1991). The rhythm of *tanka,* short Japanese poems: Read in prose style and contest style. *Proceedings of the XIIth International Congress of Phonetic Sciences* **2**, 314-317, Aix-en-Provence, France.

Jun, S.-A. (1995). Asymmetrical prosodic effects on the laryngeal gesture in Korean. In B. Connell & A. Arvaniti (eds.), *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV.* Cambridge, UK: Cambridge University Press.

Kaiki, N. & Y. Sagisaka. (1992). The control of segmental duration in speech synthesis using statistical methods. In Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (eds.), *Speech Perception, Production and Linguistic Structure.* Tokyo: Ohmsha.

Kaiki, N., Takeda, K., Sagisaka, Y., Katagiri, S., Umeda, T. and Kuwabara, H. (1990). A large-scale Japanese speech data base. *Proceedings of the International Conference on Spoken Language Processing,* **1(5)**, 17-20.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* **3**, 129-140.

Oller , D. K. (1973). The effect of position in utterance on speech segment duration in English, *JASA (54),* 1235-1247.

Pasdeloup, V. (1990). *Modèle de règles rythmiques du français appliquèes á la synthèse de la parole.* Ph.D. dissertation. U. de Procence.

Pierrehumbert, J. & M. E. Beckman. (1988). *Japanese Tone Structure.* Cambridge, MA: MIT Press.

Takeda, K., Sagisaka, Y., and Kuwabara, H. (1989). On sentence-level factors governing segmental duration in Japanese. *Journal of the Acoustical Society of America* **86**, 2081-2087.

Venditti, J. J. (1995). Japanese ToBI labeling guidelines.

Venditti, J. J. & J. P. H. van Santan. (1998). Modeling segmental durations for Japanese text-to-speech synthesis (CD-rom). *Proceedings of the 3rd ESCA TTS Workshop,* Jenolan Caves, Australia.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* **92**, 1707-17.