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 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, 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, 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 of 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 (eurhythm) [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 syntactic 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 Phrase5, 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

---

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].
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).

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:


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 Xp is obligatorily made up of one or several constituents at the level immediately below (Xp-1) and never includes other Xp constituents. Thus, an Xp-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. 

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 mêlодique" 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

---

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.
an (optional) cue to the left boundary of a constituent ("initial rise," in Vaissière 1975). Section 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 them³.

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 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

\textit{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 \textit{contenant} and \textit{contenu}, between sound and structure [Saussure 1915, Grammont 1933].

**II. 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 \textit{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 & Kruckenber 1991, Sluijter 1995].

I will present different types of prominence (or stress levels) with French and English examples. 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
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>English</td>
</tr>
<tr>
<td>Lexical Stress</td>
<td>Foot</td>
</tr>
<tr>
<td>Pitch Accent</td>
<td>?</td>
</tr>
<tr>
<td>Nuclear Pitch Accent</td>
<td>Intermediate Int. Phrase</td>
</tr>
<tr>
<td>Initial Accent</td>
<td>Ø</td>
</tr>
</tbody>
</table>

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 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.

c- Initial stress in French: 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.

d - Focal or emphatic stress: 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 (export

---

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.

vs. *deport*). 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].
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 –– E 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 E/ 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 /E/. 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 Intonation-Phrase-final stressed syllables (before a pause) and Intonation-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 (intragestural differences), whereas the difference between Intonation-Phrase-final and non-final stressed syllables is due to a difference in the phasing of the closing and opening gestures (intergestural 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 [pA]. 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 /ʊ/.

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 /ʌ/ and /ɪ/ but not in non-low vowels /u i E/ 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 speaker 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 /U/ 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.

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

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.

Preliminary note on the distinction between “onset” and “coda”: not only a positional difference, but also a difference in the combination of segments

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.

**Initial consonants:** 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.

**Initial vowels:** 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.

Spatial variations in consonants: 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/ [Brownman & 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 possible to 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 å A/ (for /i/, higher F2 and lower F1; for /u/, lower F2 and F1; for /å 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 *hair*, *pêage*, *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 [Benguereel 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. Vaisière (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).

Variations in initial position:
+ displacement
+ duration
+ tongue pressure

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 /mæ/ 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 /tʰ/ 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 /kʰ/ (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 /pʰ/ 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) diapophonia (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: \[ \text{IPi}>\text{PPi}>(\text{Wi, Si}) \quad \text{(with } \text{Ui}=(\text{IPi, PPi}) \]
Speaker 2: \[ \text{Ui}>(\text{IPi, PPi})>\text{Wi}>	ext{Si} \]
Speaker 3: \[ (\text{Ui, IPi})>(\text{PPi, Wi})>\text{Si} \]

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.

g. 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 (CV₁V₂N). 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
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 a 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.

**B4. 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, Quéné 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-finall) and medial syllables (which are short), initial syllables have an average duration (which is lengthened by the consonant only). In Dutch, Quéné (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 velo trace 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 articulatorily 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:

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Word-initial [#C]</td>
<td>1. Syllable-initial before C [.VC]</td>
</tr>
<tr>
<td>2. Syllable-initial after C [C.C]</td>
<td>Syllable-medial between C [.CVC.]</td>
</tr>
<tr>
<td>3. Word-final [C#]</td>
<td>2. Syllable-medial (without C) [.V]</td>
</tr>
<tr>
<td>5. Syllable-final before C [C.C]</td>
<td></td>
</tr>
</tbody>
</table>

(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: [ʼfaktum] becomes [ʼfactu]. The final consonant in the stressed syllable is now in the weakest position. It is altered, then absorbed into the vowel: [ʼfaktu] becomes [façtʼ] then [faʃt] then [fajt]. /t/ thus becomes final and is dropped, since it is in a weak position, resulting in the French word [fɛjt]. 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. *Six hommes* [sizʊm]).
- 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:

*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

---

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.
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 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.

B.2. What do these changes consist of?

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 feature 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 &
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 "parade fad foot parole" 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. "repeat fad foot repair") [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:

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

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

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

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

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

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 (Génervier, 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."

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

---

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.
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 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.

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

\[C1.c.\] Contrasts described with reference to "articulatory force"

\[\alpha\]. 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:

<table>
<thead>
<tr>
<th>Force Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>stops &gt; affricates &gt; continuants &gt; vowels</td>
</tr>
<tr>
<td>voiceless &gt; voiced &gt; nasals and liquids</td>
</tr>
<tr>
<td>palatal &gt; non-palatal</td>
</tr>
</tbody>
</table>

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

\[\beta\]. 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:

---

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

<table>
<thead>
<tr>
<th></th>
<th>Jaw</th>
<th>Tongue</th>
<th>Lips</th>
<th>Velum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stops</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>oral</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>nasal</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>lateral</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>Fricatives</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>Vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>open</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
<td>↑</td>
</tr>
<tr>
<td>close</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
<td>↑</td>
</tr>
<tr>
<td>nasal</td>
<td></td>
<td></td>
<td>↑</td>
<td></td>
</tr>
</tbody>
</table>

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.

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\textsuperscript{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 \textit{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\textsuperscript{18}. This position is shared by Fujimura (1990) who notes that, even though the physiological mechanism of velum control in speech is not perfectly

\textsuperscript{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).

\textsuperscript{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.
understood (p. 201),

*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 Fougeron (1998).

<table>
<thead>
<tr>
<th></th>
<th>tongue</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>velum</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Stops</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>oral</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
<td>/t k/</td>
</tr>
<tr>
<td>nasal</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td>/n/</td>
</tr>
<tr>
<td>lateral</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>/l/</td>
</tr>
<tr>
<td>Fricatives</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td>↑</td>
<td></td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td>/s/</td>
</tr>
<tr>
<td>Vowels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>open</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>close</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
<td></td>
<td></td>
<td>=</td>
<td></td>
<td></td>
<td>↑</td>
<td>/i/</td>
</tr>
</tbody>
</table>

| nasal | ↓ | ↑ | ↑ | /ã/ |

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 Society 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: Journées d'Études sur la Parole
JSR: Journal of Speech and Hearing Research
LSA: Linguistic Society of America meeting
MIT QPR: MIT Quarterly Progress Report


Language and Speech, 36 (2, 3), 197-212.


