



**HAL**  
open science

## Further note on French prosody

Jacqueline Vaissière

► **To cite this version:**

Jacqueline Vaissière. Further note on French prosody. Quarterly Progress Report, Research Laboratory of Electronics, Massachusetts Institute of Technology, 1975, 115, pp.251-262. halshs-00363974

**HAL Id: halshs-00363974**

**<https://shs.hal.science/halshs-00363974>**

Submitted on 24 Feb 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## 5.2 FURTHER NOTE ON FRENCH PROSODY

### 1. Introduction

This report continues our analysis<sup>1</sup> of fundamental frequency ( $F_0$ ) contours in French. The previous study has been extended by giving a modified version of the paragraph (Reading Material) that appeared in the appendix of the previous report and by adding a list of new isolated sentences to the speech that has been analyzed. The new corpus serves as an appendix to this report. Three more speakers have been used as subjects, and two of our previous speakers also read the new corpus.

The previous report<sup>1</sup> made the following points.

1. The speaker decomposes the sentences by inserting pauses into sense groups (each sense group being composed of closely related words), and the number of sense groups in a sentence tends to decrease as the rate of elocution increases.

2. The intonation patterns of words in final position of a sense group are determined by the position of that sense group in the sentence (intonation of the words, a falling or a rising, depends on whether the sense group is in final position in the sentence). We noted that words that are not in final position in the sense group are pronounced with certain types of intonation patterns depending on the speaker's judgment of the closeness of the relation between the successive words.

Our analysis of the actual shape of the  $F_0$  contours for each word within a sense group leads us to conclude that  $F_0$  contours of any word can be schematized visually by one of the four patterns represented in Fig. XVII-1 which represents  $F_0$  contours found for words containing three or more syllables. The patterns for shorter words (monosyllabic and disyllabic) often degenerate into simpler contours.

We shall now describe the four patterns in detail and then discuss the factors that influence the speaker to choose a particular pattern when he pronounces a word embedded in a sentence.

### 2. Contour Patterns

The schematized  $F_0$  contours of the words are described in terms of certain attributes, such as Rising, Lowering or Peak, which were used by Shinji Maeda in his study of characterization of  $F_0$  contours for American English.<sup>2</sup>

#### a. Pattern P4 (Ri+F)

The pattern P4, shown in Fig. XVII-1a, is characterized by an initial rise (Ri) followed by a falling contour (F). This pattern was found for the final word at the end

(XVII. SPEECH COMMUNICATION)

of a sentence, or for words in nonfinal position in a sense group. The falling contour always reaches a lower  $F_0$  value in sentence final position than in nonfinal position

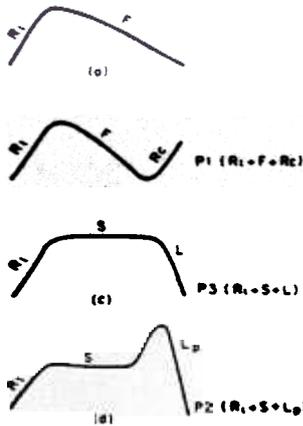


Fig. XVII-1.

Schematized  $F_0$  contours for the long words embedded in a sentence. The six attributes are as follows:

- |                       |                                   |
|-----------------------|-----------------------------------|
| Ri: Initial Rise      | Lp: Lowering associated with Peak |
| Rc: Continuation Rise | S: Sustained                      |
| L: Lowering           | F: Fall                           |

within a sentence. The initial rise (Ri) occurs at the onset of the word (more details on Ri will be given later). The falling F immediately follows the rise Ri, and continues until the end of the word. The extension of falling to the onset of the following word seems to depend on various factors such as the rate of elocution and the habits of individual speakers.

b. Pattern P1 (Ri+F+Rc)

The pattern P1, shown in Fig. XVII-1b, was found at the end of a sense group which is followed by a pause; only one of the nine speakers used P1 for words inside a sense group, followed by a very short pause. Pattern P1 is primarily characterized by a rising contour at the end of the word. The  $F_0$  contour is raised at the beginning of the word (Ri) and then falls (F) as observed in the previous pattern P4, but is terminated by the continuation rise Rc (following the terminology of Delattre<sup>3</sup>) during the last-sounded syllable of the word. In disyllabic words, the pattern P1 can be realized fully as Ri+F+Rc, but it often degenerates into a simpler contour so that the important attribute for this pattern, Rc, is still conserved. Therefore the schematic pattern for P1 in short words may be described as

$$P1 = \left\{ \begin{matrix} Ri \\ \emptyset \end{matrix} \right\} + \left\{ \begin{matrix} S \\ F \end{matrix} \right\} + Rc$$

where one of the elements in { } has to be chosen, and  $\emptyset$  indicates a deletion.

## c. Pattern 3 (Ri+S+L) and Pattern 2 (Ri+S+Lp)

The patterns P3 and P2 (Fig. XVII-1c and 1d) were found for the nonfinal words in the sense groups. They are characterized by a sustained  $F_0$  contour (S) which indicates a plateau during the intermediate syllables of the word. During the plateau, the  $F_0$  values may fall gradually along the syllables or stay approximately level (depending on the speaker). In both patterns P3 and P2 the  $F_0$  rises at the onset of the word, and is then sustained. For pattern P3, the  $F_0$  then falls rapidly at the end of the word (during the last phoneme or during the whole last syllable) and the lowering contour (L) is connected to that of the following word. In pattern P2 the initial rise Ri and the sustaining S are similar to Ri and S in P3, but the lowering L in the last syllable is preceded by a sharp rise. The successive rise and fall during the last syllable of the word indicate a peak (p) in the  $F_0$  contour. The rise occurs at the onset of the last syllable, and involves generally only the first phoneme of that syllable.

## 3. Correspondence between the Attributes and the Phonemes

Each attribute corresponds to a certain portion (sequence of successive phonemes) of the word. We have already described the parts of the word corresponding to the attributes L, Lp, and Rc. We shall now describe the portion taken by Ri (the sequence of phonemes taken by the attributes F and S are automatically determined if the portions taken by the other four attributes are known). The portion corresponding to the initial rise Ri depends on the phonemic category at the onset of the word. Three separate cases must be considered: in case 1 the initial phoneme is a consonant other than /l/, /m/ or /n/; in case 2 it is one of the consonants /l/, /m/ or /n/; in case 3 the initial phoneme is a vowel. In these three cases we define the portion of the rise Ri as the segment between the end of the preceding word and the phoneme in which the rising Ri is connected to F (in P1 and P4) or to S (in P3 and P2). This definition can be adopted for the rise Ri that occurs during the unvoiced consonants in case 1, where the  $F_0$  contour is interrupted; for the initial voiced stops and fricatives and the consonant /r/, the  $F_0$  contour often contains a valley (micromelody), but we recognize the valley as a segmental effect, and hence ignore it.

In case 1, where the initial phoneme is a consonant other than /l/, /m/, /n/, Ri occurs during that consonant, and occasionally (for some speakers) the rise Ri is extended to the following vowel. See, for example, the  $F_0$  contour of the word "la compatibilité" in Fig. XVII-2a, the word "la parasitologie" in Fig. XVII-3, or the word "la confédération" in Fig. XVII-4a. (In the three examples rise Ri occurs during the

(XVII. SPEECH COMMUNICATION)

voiceless consonant; other examples were given in our last report.<sup>4)</sup> In case 2, where the initial phoneme is the consonant /l/, /m/ or /n/, the rise Ri occurs at least during that initial consonant and the following vowel. See, for example, the word "manifestations" in Fig. XVII-4b [and also the words "météo" and "Massachusetts" in Figs. XVI-22 and XVI-26 (Quarterly Progress Report No. 114, pages 214 and 219)]. Often the rise Ri is extended to the following sonorant segment, if any. See, for example, the  $F_0$  contour of the word "manifestations" in Fig. XVII-5 pronounced by a different speaker from the one of Fig. XVII-4. In that case the rise is extended to the two first syllables. When the word begins with a vowel (case 3), the rise Ri occurs at least during the first sonorant segment of the word (we call a sonorant segment a sequence of phonemes composed of a vowel and the consonants /l/, /m/, /n/). Compare, for example, the  $F_0$  contours in the words "la compatibilité" and "l'incompatibilité" in Fig. XVII-2;

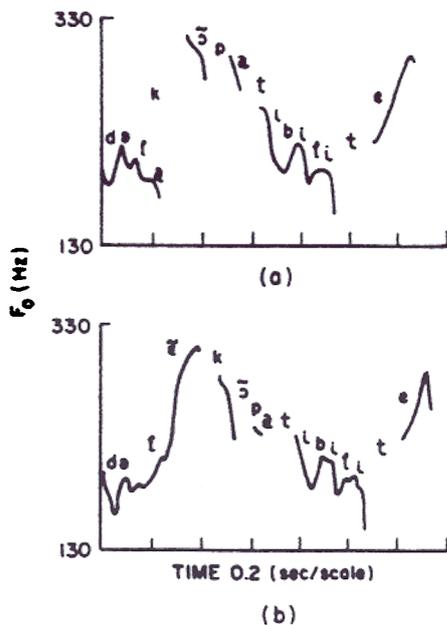
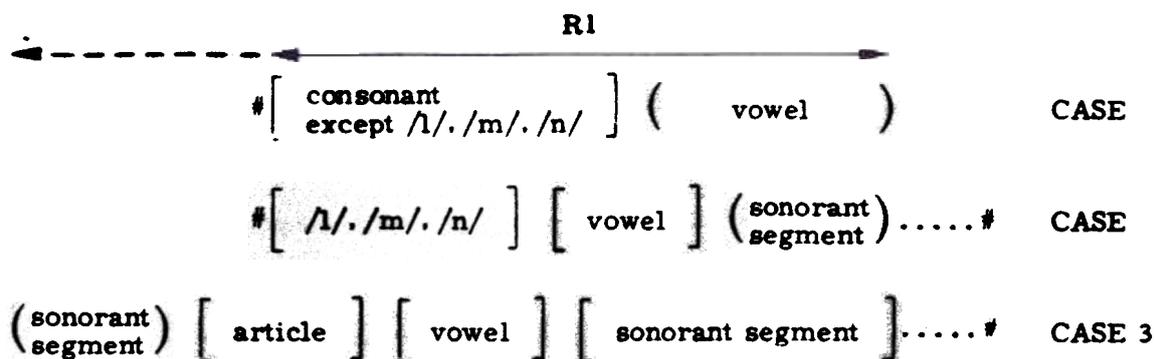


Fig. XVII-2.

$F_0$  contours of the words "de la compatibilité" and "de l'incompatibilité" embedded in a sentence.

in the second word the portion of Ri includes the first syllable. The  $F_0$  contour of the word "l'anesthésiologie" in Fig. XVII-3 shows the extension of Ri to the second (sonorant) syllable of the word. In case 3 the rise Ri may also be extended to the sonorant segment of the preceding function word, particularly at the beginning of a sense group. See, for example, the extension of the rise Ri to the preceding article "une" in the sense group "est une institution mixte" in Fig. XVII-7a and XVII-7b.

We may summarize the preceding results by the following diagram: brackets indicate the portion taken by Ri, parentheses indicate an optional extension of Ri, and (#) indicates the lexical word boundaries.



Note that the portion of R1 is closely related to the duration of the phonemes. For words located in a similar position in a sentence and pronounced with one of the four patterns (P1, P2, P3 or P4), the angle of the slope of  $F_0$  rising during R1 tends to stay constant, independent of the identity of the phonemes, and the duration of the rising portion tends to be kept fixed, regardless of the number of phonemes underlying the rising. Figure XVII-3 shows the superimposed four  $F_0$  contours of the words "la parasitologie," "la radiologie," "l'anesthésiologie," and "la cardiologie" spoken by one speaker (the four words compose the subject phrase in a sentence). We note that the overall contours

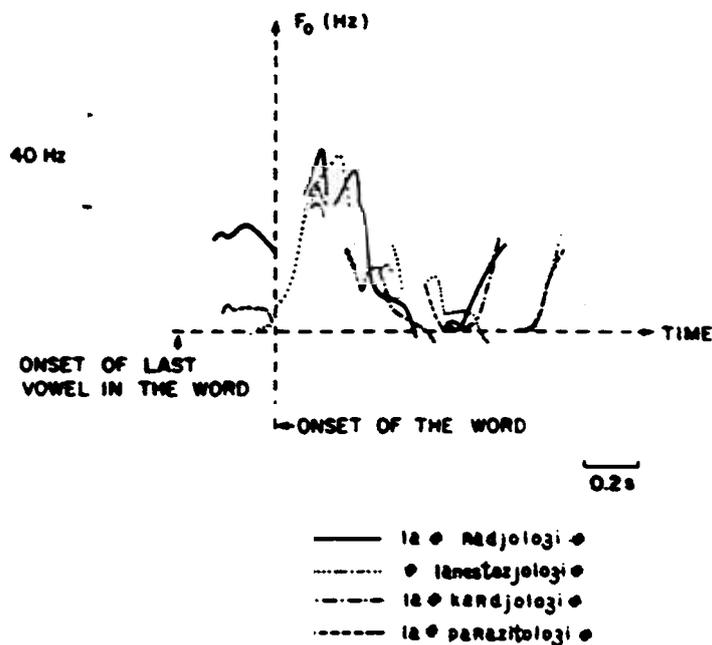


Fig. XVII-3. Superimposed  $F_0$  contours of the four words "la radiologie," "l'anesthésiologie," "la cardiologie," and "la parasitologie." The  $F_0$  contours of each word are shifted along the frequency scale to line up the  $F_0$  values at the onset of the last vowel (/i/) in the words.

194

## (XVII. SPEECH COMMUNICATION)

for the four words are quite similar. We also observe that if the distribution of the portions for the attributes  $R_i$  and  $F$  differs from one word to another,  $R_i$  depends essentially on the phonemic context,  $R_c$  depends on the habit of the speaker (in Fig. XVII-3, the speaker raises  $F_0$  only during the last vowel), and the fall  $F$  takes the remaining portion. In Fig. XVII-4,  $F$  takes two syllables in the words radiologie and cardiologie, three syllables in anesthésiologie and four syllables in parasitologie.

### 4. Nuclear Stress Rule within the Sense Group in French

What factors influence the speaker when he chooses one of the three patterns P2, P3, or P4 for the words inside the sense groups? The most common case is sense groups composed of three lexical words. Suppose that  $P_x$ ,  $P_y$ , and  $P_z$  represent the patterns attributed to the first, second, and last lexical words. In nonfinal sense groups we have found all possible combinations for the pattern of the first word,  $P_x$ , and the pattern for the second word,  $P_y$ . The combinations of the pattern for the first and second words can be separated into two cases: in case 1 the speakers repeat the same pattern (in that case,  $P_x = P_y$ ), and in case 2 they choose two different patterns for  $P_x$  and for  $P_y$  (i. e.,  $P_x \neq P_y$ ).

In the first case, the speakers generally choose their most frequent pattern for  $P_x$  and  $P_y$ , either P2, P3, or P4. For example, the subject noun phrase "L'Institut de Technologie du Massachusetts..." has been pronounced with the following descriptions depending on the speakers: P4P4P1 by two speakers, P3P3P1 by two speakers [see an example in Fig. XVI-26d (Quarterly Progress Report No. 114, page 219)], and P2P2P1 by three speakers (see Fig. XVI-26a in Quarterly Progress Report No. 114, page 219). In this case the distribution of the patterns does not give information about the internal syntactic structure of the sense group. We call a combination where  $P_x = P_y$  a prosodic parallel structure. Prosodic parallel structure has been found frequently for sense groups representing a single meaning (such as an institution or an organization) or a noun phrase [such as the noun phrase subject "un retour offensif de l'hiver" represented in Fig. XVI-23 (Quarterly Progress Report No. 114, page 216, with the description P4P4P1)].

In the second case, however, the speaker used two different patterns for  $P_x$  and  $P_y$  ( $P_x \neq P_y$ ). Figure XVII-4 illustrates the complete  $F_0$  pattern for the clause "La confédération générale du travail a organisé des manifestations importantes..."; the clause is divided by a pause into two sense groups, the first one equivalent to the subject noun phrase (represented in Fig. XVII-4a) and the second to the predicate (represented in Fig. XVII-4b). The first sense group may be described by the sequence of patterns P4P2P1, and the second by the sequence P2P3P1. The noun-phrase subject has been spoken generally with a parallel prosodic structure (by 6 speakers) and also with the

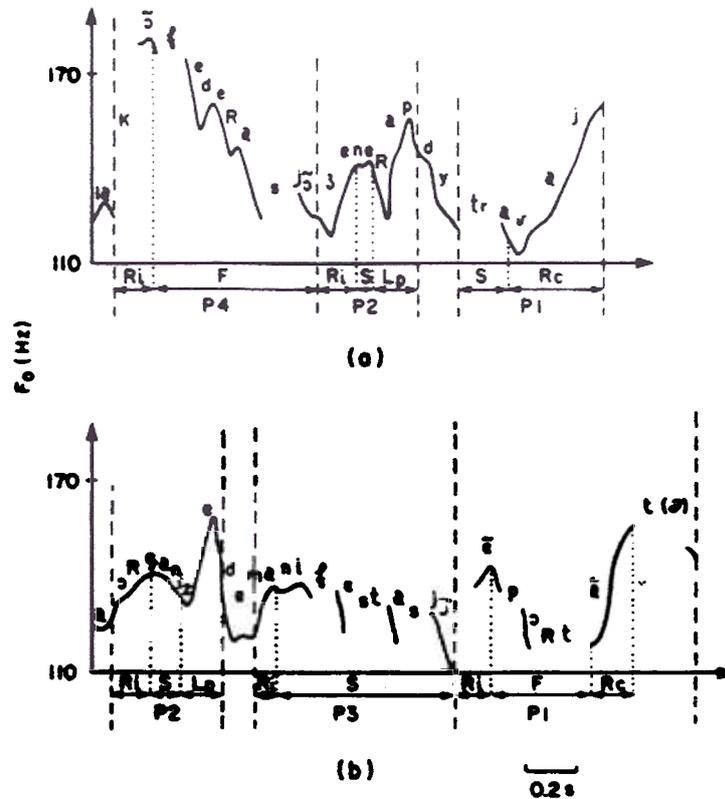


Fig. XVII-4.  $F_0$  contour of the clause "La confédération générale du travail a organisé des manifestations importantes ..." spoken by speaker BD.

description P3P2P1. We found one parallel prosodic (P2P2P1) for the predicate for one speaker; the other speakers used either P2P3P1 (Fig. XVII-4b) or P2P4P1 or P3P4P1 (Fig. XVII-5). In case of a syntactic left-branched structure (such as the noun-phrase subject in Fig. XVII-4a) the speakers realized it with a parallel prosodic structure ( $P_x = P_y$ ), or with a left-branched prosodic structure, such as P4P3P1, P4P2P1 or P3P2P1. This may be described as  $P_x P_y P_1$  where ( $x > y$ ). Similarly, the speakers realized a syntactic right-branched structure such as the predicate represented in Figs. XVII-4 and XVII-5 with a parallel structure or with a right-branched prosodic structure such as P3P4P1, P2P4P1 or P2P3P1. This may be described as  $P_x P_y P_1$  where ( $x < y$ ). Figure XVII-6a shows the organization of the pattern in nonfinal sense groups composed of 3 lexical words.

There is an interesting analogy between the stress pattern for a phrase derived by Chomsky's Nuclear Stress Rule<sup>5</sup> in English and the sequence of the patterns  $P_x P_y P_z$  that we have described. In the case of a left-branched structure, the Nuclear Stress Rule for English assigns the degrees of prominence 3, 2, and 1 to the lexically stressed

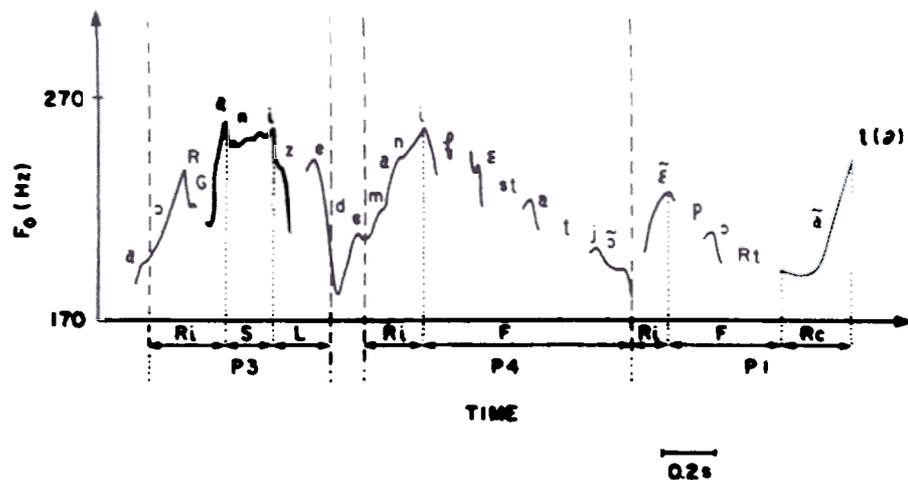


Fig. XVII-5.  $F_0$  contour of the predicate "a organisé des manifestations importantes ..." spoken by speaker MP.

SYNTAX	PROSODY	
	Nonfinal Sense Group (Z=1)	Final Sense Group (Z=4)
Left-branched Structure	Left-branched or Parallel Prosodic Structure	
<p>(Chomsky-Halle Rule 321)</p>	$x \geq y > z$ x and y = 2, 3 or 4	$z > x \geq y$ y = 2 or 1 (insertion of pause) x = 2, 3 or 4
Right-branched Structure	Right-branched or Parallel Prosodic Structure	
<p>(Chomsky-Halle Rule 231)</p>	$y \geq x > z$ x and y = 2, 3 or 4	$x \leq y < z$ y = 2 x = 2 or 1 (insertion of pause)

Fig. XVII-6. Organization of the  $F_0$  patterns in nonfinal and final sense groups composed of three lexical words.

syllables of the first, second, and third words, respectively; it assigns the degree of prominence 2, 3, and 1 in case of a right-branched structure. In all cases, number 1 designates a higher degree of stress. A rule similar to the Nuclear Stress Rule would also be valid for French, but it would not assign degrees of prominence: it would assign certain  $F_0$  patterns or a possibility of certain combinations of patterns to the words. The more stressed syllables in English correspond to the more rising syllables in French: P1 where the last syllable has a rising intonation, P2 where the last syllable has consecutively a rising and a falling contour, P3 where the last syllable is sustaining and falling, and P4 where the whole word has a falling intonation.

When the sense group is sentence-final, there are less attested possible combinations (see Fig. XVII-6). First, the last word has a falling intonation ( $P_z = P_4$ ). And, as we have found previously,<sup>1</sup> the final fall starts from the last syllable of the penultimate lexical word, which has pattern P2. There are a few exceptions where the penultimate word has pattern P3, and there is one case of pattern P4. Consequently, speakers only exercise choice in the pattern of the first word, which can be P4, P3, or P2. In case of a left-branched structure, the speaker inserts generally a small pause after the second word and P2 becomes P1, or pronounces the sense group with a parallel prosodic structure, P2P2P4. In case of a right-branched structure, the most attested combination of patterns is also P2P2P4. The opposition between right- and left-branched structure is neutralized in most cases, because of the constraints on  $P_y$  and  $P_z$ . To disambiguate phrases like "un marchand de tapis persans" and "un marchand de tapis persan" (in the first phrase, the carpets are Persian, and in the second, the seller is Persian), most speakers inserted a pause in the second case and the most attested distributions of patterns for the two cases are P2P2P4 and P2P1P4. The first case when heard in isolation is still ambiguous. When the sense group is composed of only two words, the first word may have one of the three patterns, P2, P3, or P4 in a nonfinal sense group, but generally only P2 in a final sense group. Figure XVII-7 illustrates the sense group "est une institution mixte" pronounced by three

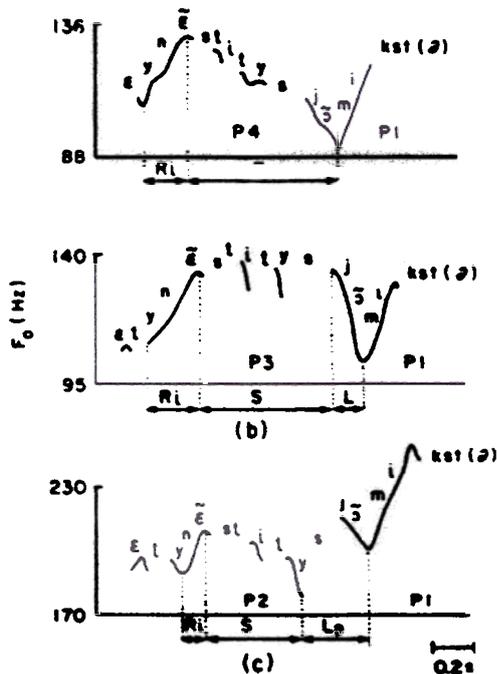


Fig. XVII-7.

$F_0$  contour of the words "est une institution mixte" spoken by three speakers. Each speaker uses a different pattern (P4, P3 or P2) for the word "institution."

## (XVII. SPEECH COMMUNICATION)

different speakers, each of whom chooses a different pattern for the first word: P4 for the first speaker, P3 for the second, and P2 for the third. When they speak rapidly P4 is generally preferred by speakers to the two other patterns, probably because it requires less effort for this particular example.

### 5. Conclusion

A phonetic unit such as the sense group that we have described is found not only in French but also in English. Shinji Maeda<sup>2</sup> claimed that American-English sentences are demarcated into smaller units, which he called "phonetic groups." He composed a certain type of  $F_0$  pattern corresponding to each group. The contents of the sense group and the phonetic group in French and in English are quite similar and correspond to linguistic constituents such as noun phrases or predicates. The patterns differ radically, however, in French and in English. In English, the  $F_0$  pattern for a phonetic group indicates roughly a "hat pattern." The contour is raised at the first primary syllable in the first content word and is lowered at the end of the primary stressed syllable in the final word, so that the lexical stresses are well represented by a rapid rising and a rapid falling of  $F_0$  values. In French, on the other hand, the sense groups are distinguished by certain combinations of  $F_0$  patterns for each word inside the sense group. This difference may arise because in English sense groups there are lexical stresses constraining the  $F_0$  pattern, whereas in French such stresses are absent and hence  $F_0$  patterns are applied to words as a whole.

## Appendix

### Reading Material 2

#### 1 Modified Version of the Previous Paragraph:

L'institut de technologie est une institution privée et mixte, dont les centres d'intérêts sont l'architecture, les sciences pures et les sciences de l'ingénieur. Il a apporté sa contribution aux progrès technologiques des dernières années, et il continue sa participation dans les techniques les plus récentes. La gamme de ses recherches est très étendue, et elle s'étend de l'électronique à la biologie, en passant par les sciences nucléaires, la linguistique et l'économétrie. Les étudiants peuvent suivre des cours très divers et participer à des recherches très variées, en profitant de l'association de l'institut avec les universités des environs. C'est à la fin de la seconde guerre mondiale que fut construit le laboratoire de recherche en électronique. Une centaine de professeurs, encadrant trois cent cinquante étudiants y conduisent des recherches.

## 2. Isolated Sentences

The following phrases are embedded in two carrier sentences so that each phrase forms a final and a nonfinal sense group.

- Un professeur de sociologie américain
- Un professeur de sociologie américaine
- Un marchand de tapis persan
- Un marchand de tapis persans
- Un spécialiste de la géographie de l'Amérique
- Un professeur de géographie d'Amérique.

### References

1. Jacqueline Vaissière, "On French Prosody," Quarterly Progress Report No. 114, Research Laboratory of Electronics, M. I. T., July 15, 1974, pp. 212-223.
2. Shinji Maeda, "A Characterization of Fundamental Frequency Contours of Speech," Quarterly Progress Report No. 114, Research Laboratory of Electronics, M. I. T., July 15, 1974, pp. 193-211.
3. P. Delattre, "La leçon d'intonation de Simone de Beauvoir," French Rev. 35, 59-67 (1961).
4. Quarterly Progress Report No. 114, pp. 212-223, the words "consonnes" and "finales" in Figs. XVI-24 and XVI-25, and the word "technologie" in Fig. XVI-26.
5. N. Chomsky and M. Halle, The Sound Pattern of English (Harper and Row Publishers, Inc., New York, 1968).

## 7 PRINCIPES DE COMPARAISON PROSODIQUE ENTRE LES LANGUES

### 7.1 LANGUAGE-INDEPENDANT PROSODIC FEATURES (ARTICLE)

#### 5.1 Introduction

The purpose of this contribution is to investigate the similarities in form and function of prosody among diverse languages. All speakers, regardless of their specific language, are equipped with the same production and perception apparatus, and consequently have the same capabilities and must face the same physiological constraints. Such similarities should be reflected in the acoustic production of any speaker. The first specific aim of this contribution is to review a number of striking acoustic similarities in the suprasegmental aspects of neutral sentences in different languages, together with possible physiological explanations for them.

Since the global function of speech (i.e. communication) is language independent, the type of information conveyed by the speech signal should not vary greatly among languages. The second purpose here is thus to compare the linguistic functions assigned in several languages to similar suprasegmental phenomena, such as lengthening, fundamental frequency rises and falls, and intensity peaks.

Although the use of suprasegmental variation for paralinguistic functions (such as the expression of emotion) is also of prime importance in everyday conversation, the scope of this contribution is restricted to the strictly linguistic functions of prosody.

#### 5.2 Language-Independent Similarities

##### 5.2.1 Pauses

Acoustic analysis of speech production shows that speakers insert a large number of pauses while talking. Studies have indicated that the ratio of articulated sequences to total speaking time does not differ significantly from one language to another: at least, Grosjean and Deschamps [1975] have found no difference for such a ratio between native speakers of English and of French in radio interviews (about 83% for both languages).