

# THE ROLE OF RHYTHMIC GROUPS IN THE SEGMENTATION OF CONTINUOUS FRENCH SPEECH\*

*Delphine Dahan*

Max-Planck Institut für Psycholinguistik  
Nijmegen, The Netherlands

## ABSTRACT

French is characterized by the presence of a final stress at the end of rhythmic groups. Lexical processing could be facilitated for words whose right boundary also corresponds to the rhythmic-group boundary. Sentences were constructed with a target syllable at various positions relative to word and rhythmic-group boundaries. These sentences were presented to French listeners (experiment 1) and to Dutch listeners (experiment 2), whose task was to detect a target syllable. RT analysis suggests an influence of the syllable position within the rhythmic group, independently of its position in the word. The closer to the right edge of the rhythmic group the syllable was, the faster its detection. Although this pattern was quite similar for French and Dutch listeners, some differences suggest that French listeners, but not Dutch listeners, have taken advantage of specificities of French rhythm. Universal and language-specific procedures in speech perception are discussed.

## 1. INTRODUCTION

Prosody, via pauses and variations in duration and pitch, delimits constituents larger than the word. This function may be of use to listeners in the segmentation of continuous speech. In French continuous speech, words are grouped in 'rhythmic groups', marked by a final stress on the last syllable of this group. The final stress is characterized by lengthening of the syllable and by a slight F<sub>0</sub> movement (e.g. Delattre, 1966). The occurrence of final stresses contributes to the rhythmic structure of French (Wenk & Wioland, 1982; Bailly, 1989). There would be an obvious advantage for the listener to use this prosodic segmentation, since the rhythmic group delimits a unit immediately superior, and sometime equivalent to the word. The right-boundary of this unit, acoustically marked by the final stress, always corresponds to a word boundary. This could facilitate segmentation, and thereby speed up lexical access.

In the present study, listeners detected target syllables either in initial, medial, or final position within the word, and either in initial, antepenultimate, penultimate, or final position within the rhythmic group. We hypothesized that a syllable is detected faster when it occurs closer to the right edge of a rhythmic group.

Previous studies showed that phonemes in word-medial position are detected faster than in word-initial position, presumably because listeners use the lexical information they have already got to anticipate the incoming phonemes (Frauenfelder & Segui, 1989; Frauenfelder, Segui, & Dijkstra, 1990). We therefore expected the position in the word to influence syllable detection, independently or in interaction with the possible rhythmic-group effect.

## 2. EXPERIMENT 1: FRENCH LISTENERS

### 2.1. Method

**Material.** Three target syllables were selected (/bi/, /te/, and /pi/). For each syllable, we selected several polysyllabic words which contain the syllable either in initial, medial, or final position, and these words were embedded in sentences. The position of the target syllable relative to the rhythmic-group (RG) boundaries was varied, so the syllable occurred either in initial, antepenultimate, penultimate, or final position. The rhythmic groups were defined following the criteria proposed by Verluycen (1982) in his description of the accentual group in French: A rhythmic group is composed of a lexical word (e.g. noun, verb, adjective, adverb) plus its clitics (e.g. *avec une habilité*); moreover, if a noun is followed by a monosyllabic adjective, those two rhythmic groups are reconstructed into one rhythmic group (e.g. *une toupie rouge*). In the sentences, we avoided the target-bearing word occurring in the last rhythmic group before a major prosodic break. The semantic predictability of the target-bearing word in the sentence was minimized and controlled by pre-testing the sentences with native-speakers of French.

Complete counter-balancing of syllable position within the word and within the rhythmic group was not possible (for example, no syllable can be both word-initial and RG-final). Table 1 indicates the seven crossed modalities used. For each crossed modality and each target-syllable (/bi/, /te/, or /pi/), we constructed from 3 to 5 different sentences. There were 24 sentences for each syllable, 72 experimental sentences in all. In addition, 30 filler sentences were constructed, ten for each syllable. They were negative sentences, i.e. did not include an occurrence of the syllable target.

The material was recorded in a sound-damped booth by a female native-speaker of French. It was then redigitized and examined with a speech editor. Timing pulses were added to trigger the recording of subject responses. In addition, the duration and the mean F0 of the target syllables were measured. The sentences were recorded onto the left channel of a DAT tape, the timing pulses were recorded onto the right channel of the tape, so they were inaudible to the subjects.

**Subjects and Procedure.** Thirty six Psychology students at the University Paris V were tested individually in a quiet room in the Laboratory of Experimental Psychology. They heard the tapes over headphones. The materials were presented in three blocks, each with a different target syllable. The order of block presentation was varied across subjects. Before each block, a recorded sentence instructed listeners to push the button as soon as they detected the target syllable, as well to understand the meaning of each sentence. Timing and data collection were controlled by a microcomputer. A recognition test was administered at the end of the experiment, comprising twenty written sentences, half of which had been heard by the subjects exactly, while the other half were constructed by putting together phrases from more than one sentence. This test aimed to check that subjects attended to the content of the sentences.

**Analysis.** Because Word- and RG-position factors were not fully crossed, we analyzed the global effect of each factor, as well as specific comparisons consisting of analyzing the effect of one factor while holding the other factor constant. We excluded from the analyses any reaction time (RT) superior or inferior to two standard deviations from the mean, individually for each subject. Those cases were added to the non-responses and false alarms, hence indicating the miss rate. Analyses of variance were conducted separately across subjects (F1) and across items (F2) as random factors. Acoustic features of the target syllables were also analyzed, using items as random factor.

## 2.2 Results

Three listeners were excluded on the basis of their recognition test performance. The hit rates clearly indicated that the syllable /te/ was much harder to detect than the two others (75%, compared to 97% for /pi/ and 93% for /bi/). The high phonological/phonetic variability of the syllable /te/ might have disrupted its detection (e.g. “terrible” can be pronounced [teriblð] or [teriblʒ]). This syllable is therefore excluded from the subsequent analyses. Table 1 presents RTs to syllable detection. As a general remark, many analyses were found to be non-significant by items, while they were highly significant by subjects. This indicates the high variability across sentences, certainly due to the fact that different words were used for each condition. We will come back to this point later.

	RG	initial	antepenult.	penult.	final
	initial	440.9	423.9	348.3	
WORD	medial		477.7	471.3	
	final			448.1	382.8

**Table 1:** Experiment 1. Mean RT (in msec) as a function of the syllable position in the word (initial, medial, final) and in the RG (initial, antepenultimate, penultimate, final).

**Word effect.** Syllable detection varies as a function of position in the word (global word effect is significant,  $F(2,64)=23.4$ ,  $p<.001$ ,  $F(2,45)=3.14$ ,  $p=.05$ ). With a constant penultimate position within the RG, RTs are faster to syllable in word-initial position than in word-medial or -final position ( $F(2,64)=30.36$ ,  $p<.001$ ;  $F(2,45)=4.08$ ,  $p<.05$ ). No significant difference was found between word-medial and word-final positions. When the syllable is in antepenultimate position within the RG, RTs are faster to word-initial syllable than to word-medial syllable (423.9 vs. 477.7 sec,  $F(1,64)=9.33$ ,  $p<.005$ ;  $F(2,1,45)=1.25$ ,  $p>.10$ ).

**Rhythmic-group effect.** Syllable detection was found to vary as a function of position relative to the RG boundaries (global RG effect,  $F(3,96)=9.99$ ,  $p<.001$ ;  $F(2,3,44)=1.08$ , NS). Specifically, a word-initial syllable is easier to detect when it occurs in penultimate position (348.3 msec) than in antepenultimate position (423.9 msec) or in initial position (440.9 msec), ( $F(2,96)=34.6$ ,  $p<.001$ ;  $F(2,44)=3.74$ ,  $p<.05$ ). The advantage for the late position within the RG is confirmed by faster RTs for a word-final syllable in final RG position (382.8 msec) than in penultimate RG position (448.1 msec) ( $F(1,96)=17.16$ ,  $p<.001$ ;  $F(2,1,44)=1.86$ ,  $p=.06$ ). No significant difference between a word-medial syllable in antepenultimate (477.7 msec) or penultimate position (471.3 msec) was found.

In summary, a word-initial syllable is faster to detect than a word-medial or word-final syllable. In addition to this word-position effect, an influence of rhythmic-group position seems to be superimposed, consisting of a facilitation for syllables near the right edge of the RG. In particular, a word-final syllable is easier to detect if it is also group-final.

**Acoustic analysis of the syllables.** No significant variation of syllable duration as a function of position within the word or within the RG was found. Only mean F0 seems to have varied as a function of syllable position within the RG: A word-final syllable is higher in final position in the RG than in penultimate position (273.3 vs. 238.5 Hz,  $F(1,44)=8.87$ ,  $p<.005$ ), which suggests the presence of a final rise. Moreover, an initial-word syllable is higher in initial or antepenultimate position in the RG (respectively 268.8 and 273.5 Hz) than in penultimate position in the RG (242.2 Hz), ( $F(2,44)=4.11$ ,  $p<.05$ ). Nevertheless, the correlations between RTs and syllable duration and F0 values were not significant (respectively  $r[46]=-.08$  and  $r[46]=-.178$ ).

## 2.3 Discussion

The word effect goes against previous findings on phoneme-detection, which have shown that the later in the word the target occurs, the faster its detection. These findings were concerned with consonant detection. Our results showed a different effect for syllable detection: When the syllable occurs in word-initial position, it is detected faster than in medial or final position. This replicates van Ooijen's (1994) findings on vowel detection. The nature of vowel and syllable processing during speech perception, clearly different from consonant processing, is still an open issue.

The effect of syllable position within the rhythmic group suggests that the prosodic structure attracts listener's attention towards the right edge of the rhythmic constituent. However, the variability between sentences, responsible for the weak F2 values, indicates that other factors than the prosodic structure influenced the syllable detections, such as the use of different words for each prosodic condition; those different effects cannot be disentangled easily. Nevertheless, the material contained two words ("habits" whose target was 'bi', "tapis", whose target was 'pi') which were used in two different RG conditions. In one condition, the syllable target was final in the RG, in the other, it was penultimate, i.e. the noun was followed by a monosyllabic adjective. A specific analysis was carried out on these two items. RTs were significantly faster when the target was final in the group, than when it was penultimate (for "habits", 335.3 vs. 527 msec,  $F(1,30)=26.0$ ,  $p<.001$ ; for "tapis", 381.5 vs. 470.1 msec,  $F(1,27)=4.35$ ,  $p<.05^1$ ). For the same target-bearing word, syllable detection is faster if the syllable coincides with the right edge of the prosodic group. This result, although more suggestive than conclusive, strengthens the hypothesis of an influence of the rhythmic group in speech processing.

One could argue that the differences in syllable-detection latencies might be due to degree of local perceptual prominences. Although only little evidence of differences across syllables was found from acoustic analyses, this argument cannot be completely ruled out. Measurements of syllable length and mean F0 do not fully account for perceptual prominence. A way to test whether the facilitation is only due to local acoustic features is to present the same material to non-speakers of French; their syllable detection must be based on local acoustic features, rather than on appreciation of global prosodic structure based on language experience

## 3. EXPERIMENT 2: DUTCH LISTENERS

### 3.1. Method

Thirty Dutch subjects were tested at the Max-Planck Institute. They had no or almost no knowledge of French. Materials, procedure, and analyses were as in Experiment 1.

### 3.2. Results

The syllable /te/ again appeared to be much harder to detect than the two other syllables (the hit rate was 59.3%, compared to 86.1% for /bi/ and 85.3% for /pi/), and was thus excluded from the analyses. Table 2 presents RTs to syllable detection. As in Experiment 1, the F2 values were usually found non-significant.

	RG	initial	antepenult.	penult.	final
	initial	505.0	592.7	466.8	
WORD	medial		659.7	551.1	
	final			660.7	533.6

**Table 2:** Experiment 2. Mean RT (in msec) as a function of the syllable position in the word (initial, medial, final) and in the RG (initial, antepenultimate, penultimate, final).

**Word effect.** Syllable detection varies as a function of position in the word (global word effect being significant,  $F(2,58)=23.4$ ,  $p<.001$ ;  $F(2,45)=2.59$ ,  $p=.09$ ). More specifically, when the syllable is in penultimate position within the RG, RTs are faster to syllable in word-initial position (466.8 msec) than in word-medial position (551.1 msec), and RTs are faster in word-medial (551.1 msec) than in word-final position (660.7 msec) ( $F(1,2,58)=40.07$ ,  $p<.001$ ;  $F(2,45)=4.43$ ,  $p<.05$ ). The same effect was found when the syllable is in antepenultimate position within the RG: RTs are faster to word-initial syllables than to word-medial syllables (592.7 vs. 659.7 msec,  $F(1,58)=9.25$ ,  $p<.005$ ;  $F(1,45)=1.02$ , NS).

**Rhythmic group effect.** Syllable detection was found to vary as a function of position relative to the RG boundaries (global RG effect,  $F(3,87)=15.93$ ,  $p<.001$ ;  $F(2,44)=2.01$ ,  $p>.10$ ). The specific comparisons showed that, for a word-initial syllable, RTs are faster when the syllable occurs in penultimate position (466.8 msec) than in antepenultimate position (592.7 msec) or in initial position (505 msec), ( $F(1,2,87)=16.7$ ,  $p<.001$ ;  $F(2,44)=2.1$ ,  $p>.10$ ), the difference between initial and penultimate RG positions being not significant. The same advantage for late position within the RG was found for a constant word-medial position of the syllable: RTs are faster to penultimate syllables than to antepenultimate syllables (551.1 vs. 659.7 msec,  $F(1,87)=22.23$ ,  $p<.001$ ;  $F(1,44)=2.8$ ,  $p=.10$ ). Finally, for a word-final syllable, RTs are faster for final than for penultimate syllables in the RG (533.6 vs. 660.7 msec,  $F(1,87)=29.4$ ,  $p<.001$ ;  $F(1,44)=3.7$ ,  $p=.06$ ).

### 3.3. Discussion

The pattern of results for Dutch listeners is quite similar to the pattern found for French listeners: A word-initial syllable is detected faster than a word-medial or word-final syllable. Independently of this word-position effect, a consistent effect of syllable position within the RG was found: The closer to the right edge the syllable is, the faster its detection. The effects observed for French listeners may be only due to local acoustic prominence of the syllable: The syllable would be more perceptually prominent at the beginning of a word and near the end of a RG. However, more than a purely acoustic effect seems to be involved. The correlations between RTs and syllable duration and F0 values were not significant for Dutch listeners (respectively  $r[46] = -.04$  and  $r[46] = -.064$ ). Furthermore, and by contrast with what was found for French listeners, the analysis conducted on the two items used in two different RG-positions (“habits” and “tapis”) failed to show any consistent advantage for the final position compared to the penultimate position (for “habits”, 476.5 vs. 548.8 msec,  $F(1,26)=2.9$ ,  $p>.05$ ; for “tapis”, 757.2 vs. 658.2 msec<sup>2</sup>). Finally, only a weak positive correlation between RTs for French and Dutch listeners was found, marginally significant ( $r[46]=.241$ ,  $p=.095$ ). Those findings suggest that syllable detection by French listeners is sensitive to specific factors that Dutch listeners are not sensitive to.

### 4. CONCLUSION

French is characterized by a prosodic constituent, the rhythmic group, which is phonetically marked by a final stress on its last syllable. We hypothesized that rhythmic-group structure is used in the on-line processing of sentences, the rhythmic-group boundaries indicating some word boundaries, and hence facilitating lexical processing. French listeners showed a consistent effect of syllable position within the RG: The closer to the right edge the syllable is, the faster its detection. Dutch listeners performed similarly to French listeners. However, some detections might have been controlled by language-specific procedures. We suggest that, for some sentences, the right boundary of the rhythmic group also corresponded to a higher prosodic constituent, comparable to an intonational phrase (see e.g. Beckman & Pierrehumbert, 1986; Gussenhoven, 1988; Hirst, 1988). This prosodic constituent is not specific to French, but has been found in many languages, including Dutch. The facilitation for Dutch listeners to detect a late syllable in the rhythmic group could reflect their sensitivity to intonational phrases, rather than to the rhythmic group *per se*. French listeners, by contrast, revealed a rather consistent influence of the rhythmic group in their syllable detection. As said earlier, this influence needs to be confirmed by using a set of sentences in which the same target-bearing words would occur in different position in the rhythmic group

This study suggests that listeners do not process speech linearly. Prosodic structure, via high-level constituents such as rhythmic group or intonational phrase, seems to lead listeners’ attention towards the right edge of those constituents. This phenomenon would facilitate the processing of sentences, since these constituents delimit linguistic entities.

### 5. REFERENCES

1. Baily, G., “Integration of rhythmic and syntactic constraints in a model of generation of French prosody”, *Speech Communication*, Vol. 8, 1989, 137-146.
2. Beckman, M.E., and Pierrehumbert, J.B., “Intonational structure in Japanese and English”, *Phonology Yearbook*, Vol. 3, 1986, 255-309.
3. Delattre, P. *Studies in French and comparative phonetics*,. Mouton, La Haye 1966.
4. Frauenfelder, U.H., and Segui, J., “Phoneme monitoring and lexical processing: evidence for associative context effects”, *Memory & Cognition*, Vol. 17, 1989, 134-140.
5. Frauenfelder, U.H., Segui, J., and Dijkstra, T., “Lexical effects in phonemic processing: facilitatory or inhibitory?”, *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 16, 1990, 77-91.
6. Gussenhoven, C., “Adequacy in intonation analysis: the case of Dutch”, in H. van der Hulst and N. Smith (eds.), *Autosegmental studies on pitch accent* (pp 95-121), Foris Publications, Dordrecht, 1988.
7. Hirst, D., “Tonal units as constituents of prosodic structure: the evidence from English and French intonation”, in H. van der Hulst and N. Smith (eds.), *Autosegmental studies on pitch accent* (pp 151-165), Foris Publications, Dordrecht, 1988.
8. Ooijen, B. van, *The processing of vowels and consonants*, Unpublished Ph. Dissertation, University of Leiden, 1994.
9. Verluoyten, S. *Recherches sur la prosodie et la métrique du français*, Unpublished Ph. Dissertation, University of Antwerpen, 1982.
10. Wenk, B.J., and Wioland, F., “Is French really syllable-timed?”, *Journal of Phonetics*, Vol. 10, 1982, 193-216.

\* Research supported by Fyssen Foundation, Paris. I thank the Laboratory of Experimental Psychology, in Paris, for help in testing French listeners. I am grateful to Anne Cutler for her help and useful suggestions. Mailing address: Max-Planck Institut für Psycholinguistik, Postbus 310, NL-6500 AH Nijmegen. email: delphine.dahan@mpi.nl

<sup>1</sup>Several listeners (one for “habits”, 5 for “tapis”) had to be excluded from the analysis, because they had missed the target for one of the conditions.

<sup>2</sup>Five subjects were excluded from the “habits” analysis; concerning the item “tapis”, no analysis of variance was carried out because of too many missing data.