PERCEPTUAL OPTIMIZATION OF SYLLABLE DURATION IN SHORT FRENCH SENTENCES

Yukihiro Nishinuma and Danielle Duez
Institut de Phonétique, CNRS: UA 261
29, avenue Robert Schuman, 13621 Aix-en-Provence, France

Abstract

Two experiments were conducted in order to study the perceptual discrimination of syllable duration within sentences. Two natural French sentences each consisting of 8 syllables were re-synthesized with 15% increases and decreases in syllable duration up to ±60%. The modified stimuli were presented to subjects whose task was to judge how natural they sounded. The same stimuli were then used to determine whether the modifications in duration were perceptible. The results provide evidence of the fact that judgment of the rhythmic quality of sentences (a linguistic task) and temporal discrimination (a psycho-acoustic task) both have the same functional basis. The detection threshold for unstressed syllables was found to be lower than that of stressed syllables due to linguistic familiarity.

INTRODUCTION

Analysis of the temporal structure of French sentences by manipulation of syllable duration has provided evidence of a certain number of perceptually important facts (1). Among other things, the initial syllable has been shown to have certain particularities, and unstressed syllables to be relatively stable. Experimentation is currently being expanded to cover words with more syllables and sentences with a syntactically balanced structure.

EXPERIMENT I

Stimuli were synthesized while varying only syllable duration in two naturally-spoken sentences: "Mes beaux bébés restaient là-bas" (literally, "My beautiful babies were staying over there") and "La sœur de Guy mangeait son pain" (Guy's sister was eating his bread), pronounced by a woman. Both sentences had eight syllables made up of a four-syllable noun phrase and a four-syllable verb phrase of similar syntactic structure, as can be seen in figure 1.

![Figure 1. Syntactic structure of the test sentences](image1)

The initial syllable duration (figure 2) of one of the syllables was modified in each test sentence. The new durations ranged from -60% to +60% of the initial duration, in steps of 15%.

![Figure 2. Initial duration of the test sentences](image2)
Each stimulus sentence was derived from the original test sentence by modifying the duration of one syllable by a given percentage. For each target syllable, a series of stimuli were generated in random order, with a between-stimulus interval of 2 seconds and a between-series interval of 5 seconds. Each series of stimuli was presented once to ten subjects in an anechoic room, and their task was to judge the rhythmic quality of the sentence. Three possible judgments were proposed on the answer sheet: natural, acceptable, and unacceptable. The raw data was analyzed statistically, after exclusion of the practice series and the dummy stimuli. Most of the sentences judged to be natural were those in which the syllable had been modified the least, that is, the ones in the middle of the range (-30%, -15%, 0%, +15%, and +30%), as can be seen in figure 4. Each point on the curve represents the percentage of subjects who judged the sentence to be natural.

A Kruskal-Wallis one-way analysis of variance indicated that the difference between the two sentences was non-significant. ($\chi^2(9)=12.999$, $p=0.441$).

Figure 5 shows that the percentage of "natural" answers also varied significantly according to which syllable had been modified, as indicated by the analysis of variance ($\chi^2(7)=51.79$, $p<.001$). More "natural" answers were obtained for duration variations on the first syllable than for variations on other syllables. Initial syllables can undergo considerable lengthening and shortening.

Several hypotheses might be set forth to explain this phenomenon. The first syllable may be relatively long for phono-stylistic reasons such as hesitation, emphatic stress, or stress for contrast, although under normal circumstances, initial syllables are short. This may have caused subjects to judge temporal modifications on this syllable favorably. Another hypothesis is that since speech rate is not actually set until the second syllable, there may be a high degree of perceptual flexibility on the first syllable. The problem could also be discussed in terms of short-term memory capacity. By the time the subject arrives at the end of the sentence, short-term may memory overflow, removing the initial syllable from the buffer. In this case, judgments pertaining to this syllable would become random.

It should also be noted that the scores obtained on syllables without word stress in the "Mes beaux ..." sentence were lower than for the
stressed syllables in those same words. This was true for "bébés" (syllables 3 & 4 in the sentence), "restaient" (syllables 5 & 6), and "là-bas" (syllables 7 & 8) (stressed syllables are shown in boldface). The same phenomenon occurred in the "La soeur..." sentence on the ward "mangeait" (syllables 5 & 6). The Mann-Whitney test confirmed this stress effect (U(1) = 193752, p < .005, N = 1296).

In general, in like phonetic contexts, unstressed syllables are physically shorter than stressed syllables. The difference in the scores obtained for these two types of syllables suggests that the acceptability threshold for unstressed syllables is lower.

**EXPERIMENT II**

In order to test the above hypotheses, a second experiment was conducted using the same stimuli. Here the subjects were asked to state whether or not the modifications in syllable duration were perceptible. They were to answer with a plus sign (+) on the answer sheet if they perceived lengthening, a minus sign (-) if they perceived shortening, and an equal sign (=) if they heard no modification. The stimuli were presented to ten subjects who had previously been asked to perform the linguistic task used in Experiment I for familiarization. After a ten-minute break, the actual experiment was begun.

Answers were the coded numerically (+, - , and = becoming 0, 1, and 2, respectively). A Kruskal-Wallis, one-way analysis of variance on the mean answer given on each syllable yielded a significant difference between subjects ($\chi^2(9) = 28.999$, $p < .001$ and $\chi^2(9) = 35.87$, $p < .001$, $N = 72$). A significant difference between sentences was also found ($F(1,140) = 6.283$, $p < .013$).

On the other hand, the analysis of variance indicated that the difference between stressed and unstressed syllables was non-significant ($\chi^2(7) = .071$, $p = .442$ and $\chi^2(7) = 1.103$, $p = .899$, $N = 72$).

Moreover, memory for syllable duration seems to have deteriorated for syllables near the beginning of the sentence, as shown by the notable effect on the first syllable (see figure 6).

duration threshold, then the threshold is located at the point where duration was modified by approximately 25% (see Table 1). This estimate of the duration threshold is broader than that defined in other studies [2-5]. Since the target stimuli were "drowned" in a medium whose duration was approximately 1.5 seconds, a significant mask effect most likely occurred, leading to the less accurate threshold.

The 25% threshold value is quite close to the one found at the 66.6% "natural" answer level in Experiment I. This means that the linguistic task and the psycho-acoustic task seem to have been performed on the basis of same duration differential. The greater acuteness observed for unstressed syllables in Experiment I must therefore be learned, being reinforced by intensive use of the language.

<table>
<thead>
<tr>
<th></th>
<th>Expérience I</th>
<th>Expérience II</th>
</tr>
</thead>
<tbody>
<tr>
<td>sentence &quot;La soeur&quot;</td>
<td>30.0%</td>
<td>25.3%</td>
</tr>
<tr>
<td>sentence &quot;Mes beaux&quot;</td>
<td>22.8%</td>
<td>24.4%</td>
</tr>
</tbody>
</table>

Table 1. Estimated differential duration threshold
CONCLUSIONS

These two experiments (psycholinguistic and psycho-acoustic) revealed the following:

1. A ±25% variation in duration may be the threshold at which natural sentences are distinguished from unnatural sentences. This threshold is the differential duration threshold defined in the psycho-acoustic task in Experiment II.

2. In the first experiment, the value of the duration threshold for unstressed syllables is significantly smaller, which means that unstressed syllables do not withstand much temporal variation. In the psycho-acoustic experiment, however, no syllable effect was found, which proves that the effect is acquired linguistically.

3. Initial syllables were granted a great deal of perceptual flexibility in both experiments, where mediocre scores were generally obtained near the beginning of sentences. Further experimentation is necessary to determine the effects of short-term memory and masking.