INFLUENCE OF NEGATIVE INTENSITY GLIDES ON THE DISCRIMINATION OF SPEECH SEGMENT DURATION

Y. Nishinuma & S. Santi

Institut de Phonétique, Université de Provence
CNRS R.U.A. 261, 13621 Aix-en-Provence, France

ABSTRACT

The discrimination of duration was investigated using synthetic vowels containing negative intensity glides (0 dB, -6 dB, -12 dB, and -18 dB). Test stimulus durations ranged from 100 to 300 ms in steps of 20 ms. The standard stimulus was 200 ms in duration and had a stable intensity. Stimulus pairs were presented to 20 subjects (constant method) and their task was to state which vowel in the pair sounded longer (forced choice). Results indicate that a drop in intensity of more than 12 dB has a significant effect on the perception of duration, and thus on its discrimination.

1. INTRODUCTION

The prosodic analysis of speech, which consists of interpreting acoustic parameters such as duration, fundamental frequency, and intensity, is not an easy task. Two reasons for this are that (1) these factors are not independent in human perception and (2) they vary as the speech signal evolves in time (within a syllable, a word, a clause, etc.). It is known that the perception of pitch variations depends upon their duration [11]. Furthermore, the melodic contour of segments with negative and positive intensity glides are perceived differently [13].

However, we know little about the interaction between duration and intensity in speech. In particular, the influence of intensity variations on the ability to discriminate the duration of speech sounds has not been experimentally documented. This problem came up in our previous study, which investigated the differential threshold of syllable duration in a sentence context [8]. Duration discrimination was found to be significantly less accurate on the final syllable than on preceding syllables. The same tendency was observed in Klatt and Cooper's data [7], which show a higher threshold for fricatives at the end of sentences than in other locations. This led us to raise the question of whether a drop in intensity (-16 dB in our case) on the final syllable of a sentence would make it difficult to correctly perceive that syllable's duration. An experiment carried out to verify this hypothesis is reported below.

2. EXPERIMENT

Klatt's formant synthesizer was used to generate stimuli for the perception test [6]. The goal was to obtain speech-like stimuli which varied in both duration and intensity. Negative intensity glides were used to approximate the final syllable of declarative sentences.

The material was designed to be used in a psycho-acoustic test based on the constant method. The standard stimulus was the vowel /a/ with a duration of 200 ms (an average syllable length) and a stable intensity of 80 dB. The test stimuli were synthesized with durations ranging from 100 ms to 300 ms in 20 ms steps (for a total of 11 different durations). Four linear intensity glides were utilized: 0 dB, -6 dB, -12 dB, and -18 dB. The fundamental frequency contour was the same for all stimuli. A slight lowering of pitch from 140 Hz to 130 Hz made the stimuli sound natural. The standard stimulus was paired with each of the test stimuli. The two vowels in each pair were separated by a silent pause lasting 600 ms. The interval separating one pair from the next was three seconds. Both within-pair orders were used (standard-test, test-standard). Each pair occurred four times. Thus, the total number of pairs was 352 (11 durations x 4 intensity glides x 2 stimulus orders x 4 repeats). Stimuli were generated in random order by a computer and recorded on a digital audio-tape. A trial series of 22 pairs was added to the beginning of the test sequence. A short beep followed by a five second silence was inserted every 22 pairs.

The perception tests were carried out in a soundproof room. Twenty subjects were tested individually, each in a single trial lasting 20 minutes. The listening level of the standard stimulus leaving the headphones was set at approximately 70 dB SPL. The written instructions to the subjects were as follows: "You are going to listen to many pairs of vowels /a, a/. For each pair you have to ask yourself the following question: Which of the two vowels is longer, the first or the second? Even if the intensity changes, please judge only the duration." The subjects responded by checking the appropriate answers on a forced-choice answer sheet.

3. RESULTS AND DISCUSSION

An analysis of variance on the data yielded a significant difference between subjects ($F_{(19,351)} = 36.84$, $p < 0.001$). This means that each subject had his or her own strategy to carry out the task. In addition, as expected in this kind of psycho-acoustic test, the order in which the stimuli were presented affected the subjects' behavior ($F_{(1,351)} = 55.6$, $p < 0.001$). Variations in the intensity factor also produced significantly different scores ($F_{(3,351)} = 168.97$, $p < 0.001$). Note however that neither the repeat factor nor the response

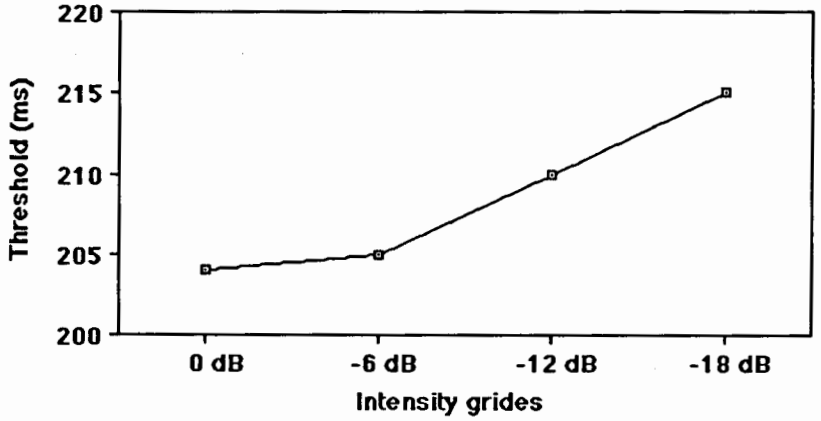

Figure 1. **Duration threshold as a function of intensity glide**

category (which member of the paircategory (which member of the pair was perceived as longer) had a statistically significant effect.

In order to compute the mean threshold (for the 20 subjects pooled) we interpolated the value for the duration at the 75% correct answer level by summing the four repeats per subject. This mean was calculated for the two stimulus presentation orders, two response categories, and four intensity glides. The averages of these values are shown in **Figure 1** and **Table 1.**

The threshold turned out to be proportional to the magnitude of the intensity glide. In other words, as the intensity glide became steeper, the detection of duration became less and less accurate. This effect on the discrimination of duration is clearly shown by the progressive increase in the means and standard deviations shown in Table 1. The difference between the stimulus with the steepest drop (-18 dB glide) and the stable intensity stimulus (no glide) exceeds 7%. Interestingly, the T-test on the data for the first two intensity glides (0 dB and -6 dB) did not yield any significant differences. Apparently, a 6 dB drop in intensity does not lead to difficulty in detecting the correct duration. In contrast, the stable vs. -12 dB difference ($t_{(19)} = 4.14$, $p < 0.001$) and the stable vs. -18 dB difference ($t_{(19)} = 5.36$, $p < 0.001$) were both highly significant. It is noteworthy that our results indirectly support those obtained by Rossi [13], who estimated the intensity glide threshold to be approximately 11 dB for a vowel lasting 200 ms. The observed change in the way intensity information is processed seems to depend on whether or not the intensity decreases beyond that critical value, although we do not know precisely where in our auditory system that change occurs.

This tendency is even more apparent if we consider stimulus presentation order. For the standard-test order, it can be hypothesized that subjects pay attention to the duration of the final syllable, which has a negative intensity glide in this experiment. The computed threshold values were 220 ms, 230 ms, 233 ms, and 235 ms, for 0 dB, -6 dB, -12 dB, and -18 dB, respectively. This indicates that when the penultimate syllable measures 200 ms and has a stable intensity, final syllables with intensity glides of 18 dB may have to be longer than 235 ms.

However, this hard and fast interpretation may need some qualification due to one peculiarity of this experiment. In comparison to the results published in psycho-acoustic studies using speech sounds, our threshold value at 200 ms is remarkably (even excessively) precise (1%, or 2 ms; cf. Figure 1 and Table 1). For a standard stimulus duration of about 200 ms, the reported threshold values fall between 8% and 30%. These experiments used several standard stimulus durations ranging from some ten milliseconds to several hundred milliseconds [1, 2, 3, 5, 7, 9, & 12]. In our experiment, all 352 stimulus pairs had a 200 ms vowel (the only standard duration used), with a level intensity in first or second position. This may have overexposed subjects to that particular duration, causing better performance. Therefore, the duration threshold defined here, (i.e. as a function of intensity glide) should be used in conjuction with those obtained under normal, stable intensity conditions.

Table 1. Duration threshold as a function of intensity glide

Intensity glides	Mean threshold	Standard deviation	N
0 dB	202	10	80
-6 dB	206	14	80
-12 dB	216	16	80
-18 dB	217	22	80

4. CONCLUSIONS

This study has provided some experimental evidence of how well we hear at the end of declarative sentences. The results of our perception tests demonstrated that the discrimination of duration may be significantly deteriorated by a progressive decrease in intensity of more than ten decibels. Our results may have some implications for the interpretation of prosodic data at the sentence level.

5. REFERENCES

[1] Bochner, J.H., Snell, K.B., & MacKenzie, D.J. (1988) "Duration discrimination of speech and tonal complex stimuli by normal hearing and hearing-impaired listeners", *JASA*, 84(2), 493-500.

[2] Bovet, P, & Rossi, M. (1977) "Etude comparée de la sensibilité différentielle à la durée avec un son pur et avec une voyelle", *Du temps biologique au temps psychologique*, PUF, Paris, pp.289-306.

[3] Eilers, R.E., Bull, D.H., Oller, D.K., & Lewis, D.C. (1984) "The discrimination of vowel duration by infants", *JASA*, 75(4), 1213-1218.

[4] Huggins, A.W.F. (1971) "Just noticeable differences for segment duration in natural speech", *JASA*, 51(4), 1270-1278.

[5] Huggins, A.W.F. (1971) "On the perception of temporal phenomena in speech", *JASA*, 51(4), 1279-1290.

[6] Klatt, D.H. (1980) "Software for a cascade/pararel formant synthesizer", JASA, 67(3), 971-995.

[7] Klatt, D.H., & Cooper, W.E. (1975) "Perception of segment duration in sentence contexts", in *Structure and Process in Speech Perception* (eds. A. Cohen & S.G. Nooteboom), Springer, New York, 69-89.

[8] Nishinuma, Y., (1990) "Discrimination of syllable duration in English and French short sentences", *JASA*, 87(Suppl 1), s72.

[9] Nooteboom, S.G. (1973) "The perceptual reality of some prosodic durations", *Journal of Phonetics*, 1(1), 25-46.

[10] Rochester, S. (1971) "Detection and duration discrimination of noise increments", *JASA*, 49(2), 1783-1794.

[11] Rossi, M. (1971) "Le seuil de glissando ou seuil de perception des variation tonales pour les sons de la parole", *Phonetica*, 23(1), 1-33.

[12] Rossi, M. (1972) "Le seuil différentiel de durée", *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*, (éd. A. Valdman), Mouton, The Hague, 435-450.

[13] Rossi, M. (1978) "The perception of non-repetitive intensity glides on vowels", *Journal of Phonetics*, 6, 9-18.