

PERCEPTION OF HESITATIONS IN SPONTANEOUS FRENCH SPEECH

Danielle Duez

CNRS URA 261, Laboratoire Parole et Langage
Université de Provence, France

ABSTRACT

The present study deals with the perception of hesitations in spontaneous speech. The results showed that half of the hesitations were identified by at least 75% of the listeners. Listener responses were also found to depend on hesitation type: silences were never perceived, contrary to filled pauses, lengthenings and repeats. For these hesitations, the listeners' responses tended to correlate with duration. These findings suggest that hesitations are perceived.

INTRODUCTION

The perception of hesitations in speech has not aroused much interest in speech research. Goldman-Eisler's claim [1] that hesitation pauses do not serve communication may be one of the reasons for this lack of interest. Following Goldman-Eisler's view the majority of the studies on the significance of hesitations have tested the hypothesis that hesitations may not be "heard". For example, in an experiment where decoders heard recorded utterances and attempted to reproduce them. Martin and Strange [2] observed that decoders placed relatively more of their hesitations at sentence breaks than did encoders. Moreover, instructions to reproduce hesitations increased hesitations and words but at the expense of the percent of correct words. The authors stressed the fact that decoders distribute pauses in accordance with the distributional scheme of pauses

Later studies on pause perception [3 and 4] confirmed the influence of the distributional scheme of pauses on pause perception; however, it was also shown that pauses occurring within constituents were not detected as well as between-phrase pauses, but they were not entirely ignored. Moreover, subjective pauses (i.e. perceived pauses which do not correspond to silent pauses) were also found to be related to significant vowel lengthening, which mostly correspond to the realization of a syntactic boundary, and/or the presence of a hesitation. This tends to prove that hesitations are perceived.

The purpose of the perceptual experiment reported here was to check whether hesitations are *actually* perceived. It was assumed that the accuracy of hesitation perception would depend on different factors such as the type, duration, and distribution of hesitations.

EXPERIMENTAL PROCEDURE

Subjects

The subjects were twenty-five native speakers of French. All were students at the University of Provence. They had no history of speaking or hearing disorders. They were paid for the task.

Materials

The stimuli used in the experiment were extracted from three hours of conversational speech produced by seven French male speakers. The conversations were digitized at a sampling rate of 16 KHz with a SUN computer. The presence of hesitations

was controlled both acoustically and perceptually. One hundred sentences were extracted from the conversations. They were selected according to the following two criteria: 1) each of the utterances began and ended with a terminal contour, and 2) each sentence contained one hesitation, or no hesitation. The limits of each utterance were determined auditorily and visually from the oscillographic and spectrographic waveforms. Out of the 100 utterances, 68 contained hesitations such as filled pauses (30 instances), repetitions of syllables and grammatical words (16 instances), silent pauses (12 instances) and lengthened syllables (10 instances). In the 32 remaining utterances, there was no hesitation.

Hesitations.

Four types of hesitation were defined.

-Filled pause (FP): any occurrence of the French hesitation interjections (euh, hum...)

-Repeat (R): any unintended repetition of a sequence of phonetic segments that is subsequently produced in its complete intended form.

Silent pause (S): any interval on the oscillographic trace where the amplitude is indistinguishable from that of the background noise. Silent hesitation pauses were restricted to interruptions located within a minor phrase, e.g. an article and a noun, a proposition and a verb...

Lengthened syllable (L): any abnormally lengthened syllable or vowel.

As false starts were always associated with another hesitation, they were not included in the corpus.

Testing procedure

The utterances were transferred to a 486 PC computer using a program written by Pavlovic et al. [5]. The subjects were told that the 100 utterances they would hear were extracted from conversations, and that each utterance contained no hesitation,

or one of the following hesitations, an "euh", a lengthened syllable, a silence, or a repeat. In the latter case, they were cautioned not to mistake a hesitation repeat for a repeat which has a semantic role. They were instructed to report the presence of a hesitation in a two-alternative forced choice task by clicking the appropriate response (yes or no) as soon as possible. Utterances were presented to subjects with headphones at a sound pressure level of 63 dB in a quiet room. Each subject took the test independently, at his own pace: after presenting an utterance, the computer waited for the subject to click the appropriate response before recording it. Then when the utterance was totally played, the next utterance was presented. An inter-utterance interval of 3 s was maintained and a different random order of utterances was used for each subject. A pretest of 8 utterances was included to familiarize subjects with the task and to give them a further chance to ask questions. The session lasted about twenty minutes.

Analysis

The limits of each hesitation were determined auditorily and acoustically by analyzing both the oscillographic trace of the signal and wide band spectrograms, each hesitation was measured in ms. For a lengthened syllable, the degree of vowel lengthening was estimated with respect to the mean length of non-prominent vowels for a within-phrase lengthening, and with respect to the mean length of prominent vowels for a lengthening occurring at a major break. This estimation was done for each speaker. Repeats were also expressed as the number of syllables and the number of times the same word was repeated. Filled pause duration ranged from 75 ms to 1554 ms, lengthenings from 243 ms to 854 ms, silences: from 231 ms to 776 ms, and repeats from 170 ms to 742 ms. The repeats ranged from

the repetition of the initial syllable in a word to the repetition of a bisyllabic word. The syntactic distribution of hesitations was studied as a function of their duration and their types. Three locations were considered: 1) within a minor phrase, 2) between minor and major phrases, and 3) between clauses. The number of listeners perceiving a hesitation was examined as a function of the type, duration and location of each hesitation. As subjects were given a forced binary choice, a hesitation was considered as perceived when reported by at least 19 listeners (i.e. 75% of the listeners).

RESULTS

Hesitation type

The mean and standard deviation of listener responses for each hesitation type are as follows. FP (M= 19.3, SD= 0.7); L (M=18.3, SD=1.3); R (M=19.4, SD=1) and S(M=8.8, SD=1.1). Hesitation type had a significant effect on mean number of responses ($F(3, 64)=20.8, p=0.0001$). A post-hoc test revealed that the mean number of responses obtained for S was significantly different from that obtained for FP, L and R ($p=0.0001$). The analysis of number of listeners perceiving each hesitation was consistent with this finding: 36 of the 68 hesitations were perceived by 75% of the listeners (19 out of the 30 FP, 6 out of the 10 L, and 11 out of the 16 R). No silence was perceived as a hesitation by 75% of the listeners.

Hesitation duration

The results suggest that duration is a cue to the perception of hesitations. A 75 ms FP was detected by only 10 listeners. FP's between 100 ms and 400ms were perceived by a number of listeners ranging from 12 to 18. Most of the FP's longer than 400 ms were perceived by the majority of listeners. There was only one exception: one 910-

ms FP was detected by only 18 listeners. Similar tendencies were observed for L's: L's greater than 450 ms were perceived as hesitations by at least 75% of the listeners while short lengthenings (between 200 ms and 500 ms) were perceived as hesitations by a number of listeners ranging from 9 to 12 listeners. For FP's and L's, there seems to be a duration threshold zone around 500 ms. Listener responses correlate less with duration for R, suggesting that listeners are also sensitive to the number of repetitions. The correlation coefficients obtained for FP, L and R are 0.5, 0.6 and 0.1, respectively.

Hesitation location

Listener responses tended to correlate more with duration for FP's located at syntactic boundaries ($r^2=0.5$) than for within-phrase FP ($r^2=0.3$). All hesitations longer than 500 ms were perceived by the majority of the listeners when located at phrase or clause boundaries, while a within-phrase hesitation as long as 500 ms was only detected by 13 listeners. However, the low number of cases does not allow us to test the significance of these differences.

CONCLUDING REMARKS

The finding that emerges from the present study is that hesitations are "heard". This finding is rather robust since listeners only had to identify one hesitation at a time. In normal communication they are used to hearing a succession of hesitations in the same location as speakers tend to accumulate hesitations when expressing themselves spontaneously. Listeners' sensitivity to hesitations is not uniform; it seems to depend on hesitation type, and hesitation duration. An extension of the present study is in progress. It should allow us to relate a hesitation perception threshold to a given duration and a specific location within the utterance.

The present finding has some implications for the role of hesitations in speech communication. Studies on the significance of pauses [7 and 8] have shown that hesitations hinder effective communication. Other studies have suggested that hesitations serve as an opportunity for the hearer to review and integrate what is to follow [9]. The role of hesitations in speech communication is probably complex. Hesitations may disturb components of speech, especially when they are long, but at the same time they may provide potentially useful information about the characteristics of the speaker's style and communication situation. A closer investigation of hesitations should bring us a better understanding of the function of paralinguistic and extralinguistic information in speech communication.

ACKNOWLEDGEMENT

Acknowledgements to M. Brousseau and R. Espesser for their technical assistance

REFERENCES

- [1] Goldman-Eisler, F. (1968). *Psycholinguistics. Experiments in spontaneous speech*, Academic Press, London and New York
- [2] Martin, J. G. and Strange, W. (1968), The perception of hesitation in spontaneous speech, *Perception and Psychophysics*, 3(6), pp. 427-438.
- [3] Duez, D. (1985) Perception of pauses in continuous speech, *Language and Speech*, 28(4), pp.377-384.
- [4] Duez, D. (1993) Acoustic correlates of subjective pauses, *Journal of Psycholinguistic Research*, 22 (1), pp. 21-39.
- [5] Pavlovic, C., Brousseau, M., Howells, D., Miller, D., Hazan, V., Faulkner, A. and Fourcin, A. (1995) Analytic assessment and training in speech and hearing using a poly-lingual

workstation, *EURAUD*, TIDE Congress, Paris, to appear.

[6] Duez, D. (1982) Silent and non silent pauses in three speech styles, *Language and Speech*, 25, pp.11-28.-

[7] Aaronson, D. (1968). Temporal course of perception in an immediate recall, *J. Exp. Psychol.* 76, pp.129-140

[8] Reich, S. S. (1980) Significance of pauses for speech perception, *Journal of Psycholinguistic Research*, 9(4), pp.379-389.

[9] O'Connell, D. C., Kowal, S. and Hörmann, H. (1969). Semantic determinants of pauses, *Psychol. Forsch.*, 33, pp.50-67.