

NIS-2070

Silent Pauses in Simulated Request-Refusal Type Dialogues. A Phonetic Analysis of German, Korean, and Japanese.

Yukihiro NISHINUMA

Laboratoire Parole et Langage, CNRS, France
yukihiro.nishinuma@lpl.univ-aix.fr

Akiko HAYASHI

Tokyo Gakugei University, Japan
hakiko@u-gakugei.ac.jp

Abstract

The present study analyzes the phonetic correlates of silent pauses in German, Korean and Japanese. Seventy two adult speakers participated to the experiment. Thirty six pairs of same-sex subjects were formed, i.e. 13 for German, 11 for Korean, and 12 for Japanese. Subjects interacted in 2 to 5 minute role-play dialogues of the 'request-refusal' type. One subject in each pair was instructed to play the role of a student asking for an appointment to the other, a professor, who was to refuse it. The corpus totaled 95 minutes in duration. The durations of silent pauses and of utterances were measured for the acoustic analysis using signal editors. Analyses of variance and other related statistics have been carried out on the measured data converted into logarithmic values. The following factors have been taken into account: language, gender, role, occurrence of pause, and utterance. Three statistically significant findings have been made: (1) a negative correlation between pause length and frequency; (2) a highly significant interaction between language and role, i.e. in each language, students and professors vary in their use of pauses; (3) a greater significance of pause than of utterance effect in interactional management.

Résumé

Nous présentons dans cette étude les résultats d'une analyse phonétique des paramètres de la pause silencieuse dans les dialogues de 3 langues : allemand, coréen et japonais. Au total, 36 paires de locuteurs adultes ont participé à l'enregistrement du corpus : 13 paires pour l'allemand, 12 paires pour le japonais et 11 paires pour le coréen. Les locuteurs ont simulé des dialogues de 2 à 5 minutes du type 'requête-refus'. Dans chaque groupe, un des locuteurs, jouant le rôle de l'étudiant, devait obtenir un rendez-vous avec l'autre, un professeur, qui devait le refuser. Les enregistrements totalisent 95 minutes de parole. Nous avons procédé, à l'aide d'éditeurs de signal, à la mesure des pauses et des énoncés du corpus. Les analyses de variance à partir des mesures converties en valeurs logarithmiques ont été effectuées avec les facteurs suivants : langue, sexe, rôle, fréquences des pauses et des énoncés. Trois faits statistiquement significatifs ont été relevés : (1) une corrélation négative entre la longueur et la fréquence des pauses ; (2) une interaction très significative entre les facteurs langue et rôle indiquant que dans chaque langue étudiants et professeurs ont une utilisation spécifique des pauses d'après leur rôle à jouer; (3) la pause a un poids plus important dans la gestion du dialogue que l'énoncé lui-même.

Key words : silent pause, dialogue, role-play, request-refusal, German, Japanese, Korean.

Mots clés : pause silencieuse, dialogue, requête-refus, allemand, japonais, coréen.

1. Introduction

Over the last years, greater attention has been given to increase the knowledge on pauses in speech, not only through the development of theories but that of applications. Thus the collecting of data has greatly diversified in linguistic discourse, in communication engineering, and in speech pathology. Corporuses now range from read to natural speech, and analysis procedures develop accordingly.

It is already well known that pauses concern all levels of the speech chain, from the semiotic concept to the physical articulation of the message (Levelt, 1989). Behind pauses, the complex relationships of the various levels of language activities can easily be figured out: breathing, swallowing, syntactic marking, lexical access, cognitive and sociolinguistic functions, and so forth.

This may explain why, unlike phoneme durations, pause lengths do not converge towards a reference value (mean, threshold). Actually, former studies propose several threshold values and agree on an estimate value for short pauses of about 200 msec. for read speech or predetermined utterances (Grosjean & Deschamps, 1972, 1973, 1975; Duez, 1982; Candea, 2000; Koiso & Den, 2000; Goldman-Eisler, 1958, 1961, 1968, 1972 ; Campione & Véronis, 2002).

As the choice of expressions becomes wider, speakers need more time, which has an impact on pause duration. Thus in the free description of a complex content, mean pause duration increases and is approximately between 500 msec. and 1000 msec. (Grosjean & Deschamps, 1975; Zellner, 1994). Three fourth of pauses would lie within that time range (Goldman-Eisler, 1968; Dalton & Hardcastle, 1977).

In situations where more than one speaker intervenes, i.e. conversation, interview, or dialogue, the choice of the final realization becomes a highly complex task given the setting, the interlocutor, and one's own internal criteria. This leads to an increase in the production and the duration of pauses (Goldman-Eisler, 1972; Gustafson-Čapková & Megyesi, 2001). In other words, pauses are an indicator of fluidity in production. When communication is easier, pauses are shorter, and when coding becomes difficult, the number of pauses increases (Grosjean & Deschamps, 1975; Pfauwadel, 1986; Zellner, 1994). Moreover, when there is a subject change within a conversation, pauses tend to be longer (Hansson, 1999; Gustafson-Čapková & Megyesi, 2001; Botinis et. Al., 2003).

In total contrast to this, very short pause thresholds have been observed. Weilhammer and Rabold (2003), in their study of the dialogues of three different languages, have reported a distribution of pauses with two peak values, i.e. a high mean value at approximately 400 msec. and a lower one at about 50 msec.. Levin et. al. (1965) also make mention of a short pause at around 80 msec..

Alongside these observations, we present, in the following pages, our own experimental results on cross-language variability of silent pauses within dialogues.

2. Experiment

2.1. Recordings

The languages studied here are German, Korean, and Japanese. As an introduction to a larger study on spontaneous conversation, the present experiment is limited to role-play dialogues in a request-refusal setting. Our objective is to maximize the possible contrast effect introduced by the induced situation.

In order to determine the influence of the sex factor, pairs with same-sex participants were formed. In each pair, a subject asked a favor to the other who was to refuse it according to the following scenario: a student (male or female) had called a week ago and obtained an appointment for an interview with a professor s/he didn't know. The student is being unexpectedly prevented from meeting the professor and calls him back on the day of the interview explaining the reason why s/he has been held up, trying to put off the interview. Seventy two subjects formed a total of 36 pairs: 5 pairs of males and 7 of females for German; 6 pairs of males and 7 of females for Japanese; 5 pairs of males and 6 of females for Korean. Conversations lasted from 2 to 5 minutes with a mean duration of 2.5 minutes. They totaled 95 minutes. Recordings were made in a very quiet room (recorder: Sony TCD-D100, microphone: SONY F-P5500).

2. 2. Phonetic Analysis

The recordings have been segmented into silent segments and speech segments, with a previously made orthographic text transcription and using spectrograms and oscillograms on speech signal editors (Praat, Soundedit16 2J, CoolEdit96). Unvoiced consonants have been considered as parts of the speech segments. Other segments assimilated to speaker sounds produced – cough, laughter, sigh, filled pauses – have been placed into a distinct category. No distinction has been established between pauses with or without breathing, their measurement requiring a specific equipment (Autesserre et. al., 1987). No distinction has been made either between pauses and portions of silence (in the discorsal sense). However, the vocal lengthenings of hesitations have been considered as parts of the speech segments. Lastly, in our coding, utterances (phrase, group of words) are always preceded by a pause which implies that dialogues are considered having a 1 msec. pause at their very beginning, thus facilitating calculations.

3. Statistical Analysis

It comprises a series of basic statistical analyzes and variance analysis (henceforth ANOVA). These have been performed with both the numeric data and with the measurements converted into logarithmic values (Unistat, Statview). Three factors have been taken into account, i.e. language, role, and sex.

3. 1. Pause Duration

As indicated in Figure 1, mean silent pause value approximates 500 msec.. However Standard Deviations are greater than mean values in all languages, revealing very important variations in silent pauses. Moreover, while in German and Korean, arithmetic mean and median values are close (German: 626msec. vs. 464msec.; Korean: 534msec. vs. 401msec.), an important difference has been observed for Japanese between those major indicators (517msec. vs. 274msec). Japanese thus seems to have a variation in pause durations different from the two other languages.

The ANOVA reveals a significant interaction between the *language* and *role* factors (with pause length as the dependent variable). In other words, each language has a specific distribution of pause durations according to speaker role. On the other hand, no interaction between *pause duration* and *sex* has been found for any language (see below).

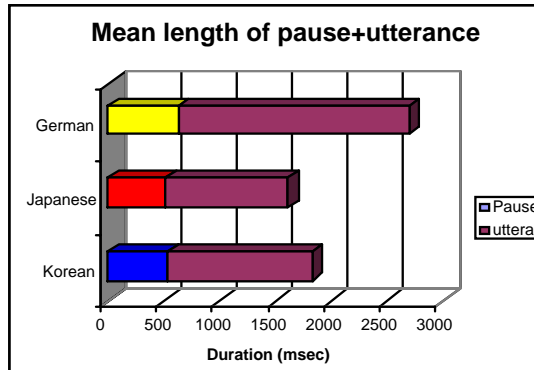


Figure 1. Mean pause and utterance durations for the three languages.

3.2. Utterance Duration

Our findings indicate that mean duration of German utterances is particularly long, i.e. over 2 seconds and twice longer than in the other two languages, a phenomenon confirmed by the ANOVA ($F(2, 2851) = 132.488, P < 0.00001$). Moreover, the sex and role factors do not significantly influence speech duration (cf. Figure 2).

Mean utterance duration (between 1 and 2 seconds), has a lower Standard Deviation than mean pause duration. The ratio between pause and utterance durations in our data is of 1 against 2 or 3, which roughly corresponds to the values found in other studies (Sugito, 1987; Autesserre et. al., 1989; Koopmans-van Beinum & Donzel, 1996).

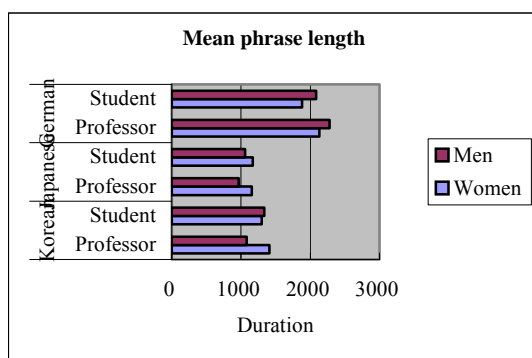


Figure 2. Mean utterance durations according to sex and role.

3.3. Pause Frequency

Pause frequency is significantly different in all three languages ($F(2, 1328) = 5.398, P < 0.0046$) and interaction with role is highly significant ($F(1, 1327) = 6.22, P < 0.0001$). Here again, there seems to be in each language two very distinct tendencies in the frequency of pauses according to speaker role: the *student* who asks a favor to his/her *professor* uses a

greater number of pauses and the professor who intends to refuse, has no need to break up his/her speech with pauses. This tendency appears quite clearly if we compare Figures 3a – 5b. In fact pauses in German are few but their durations are quite long, which is quite obvious for *male professors* (*male Korean professors* showing the same tendency).

An opposite tendency is found for *female Koran students*. As shown in Figure 5a, they use two times more pauses but very brief ones (around 400msec.) with durations reaching half the value of the pauses used by male Korean students. Overall, the frequency and length of pauses reveal a negative correlation, as can be noted in Figure 6.

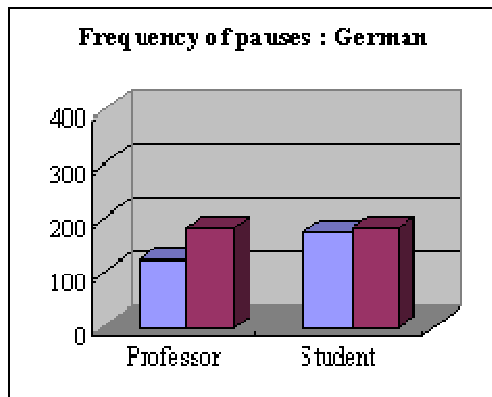


Figure 3a. Frequency of pauses for German.

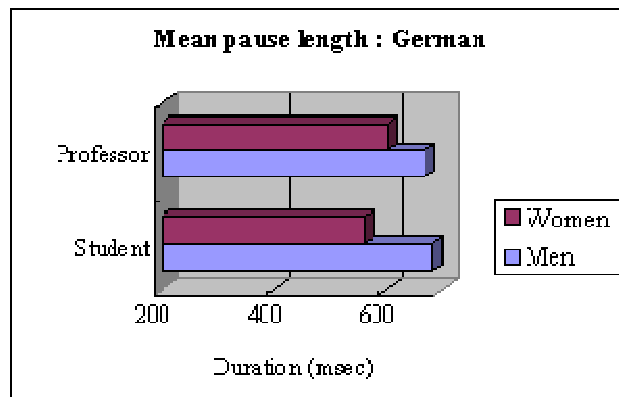


Figure 3b. Duration of pauses for German.

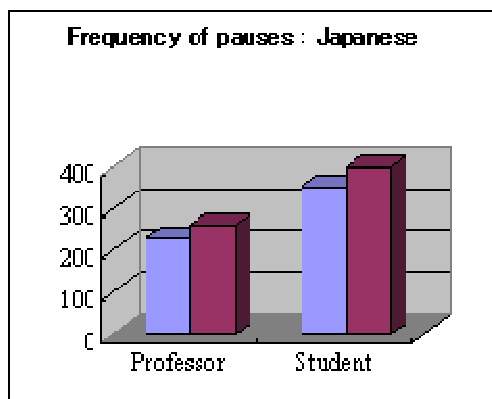


Figure 4a. Frequency of pauses for Japanese.

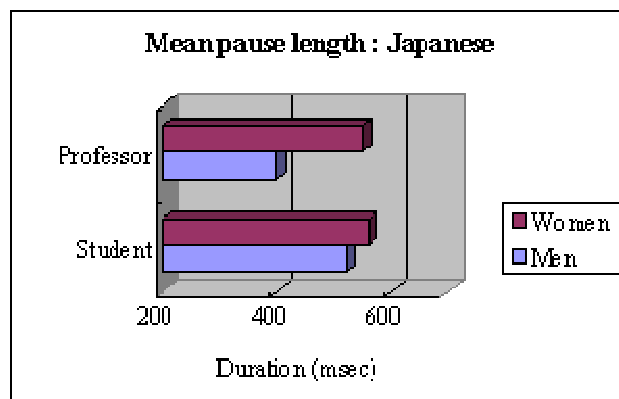


Figure 4b. Duration of pauses for Japanese.

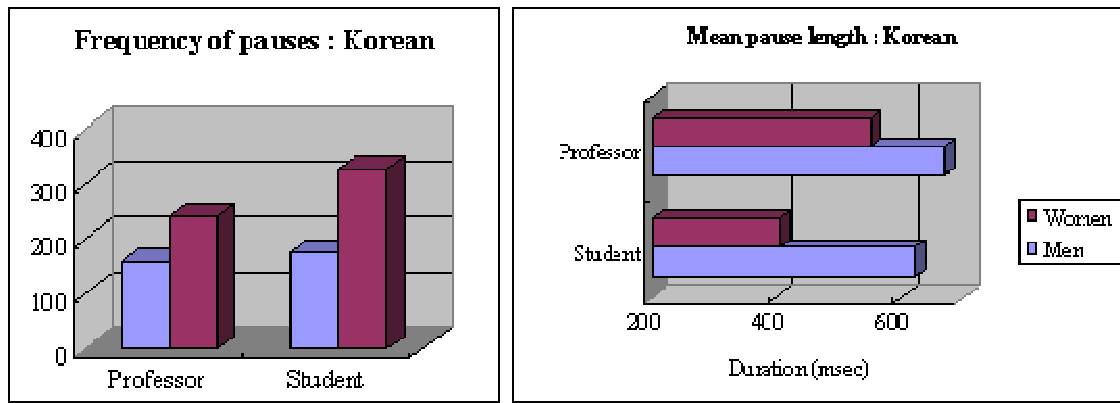


Figure 5a. Frequency of pauses for Korean.

Figure 3b. Duration of pauses for Korean.

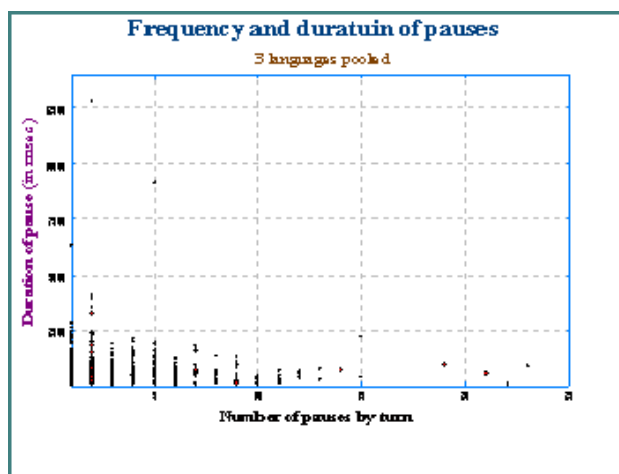


Figure 6. Frequency and duration of pauses.

3. 4. Utterance Frequency

Results show a significant interaction between *utterance frequency* and *role* ($F(1, 2849) = 17.008, P < .00001$). In all languages, *students* intervene more than the *professors*, especially in Japanese. As a rule, the solicitor needs to explain on repeated occasions the reason and the importance of the favor asked to the solicited person who gives a ready-made answer. In the case investigated in this study, the difference in the alleged social levels may restrain freedom of verbal expression, which in turn makes negotiations more difficult.

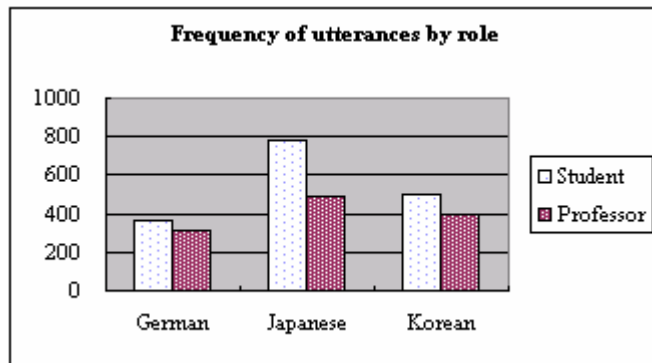


Figure 7. Frequency of utterances according to role.

4. Discussion

4. 1. Statistical Distribution of Pauses

Measured data for pauses show a typical logarithmic distribution with the three following characteristics: (i) positive values; (ii) an important peak on the left side of the distribution curve; (iii) a flat distribution on the right (Limpert et al., 2001). However former studies have shown that pause data in logarithmic values have a normal distribution (Zelher1998; Campione & Véronis, 2000; Koiso & Den, 2000; Weilhammer & Rabold, 2003). Figures 8a and 8b illustrate how the conversion of data from msec into logarithmic values affects the distribution of pause durations. The distributional normality according to utterance length has further been observed (Figure 9b).

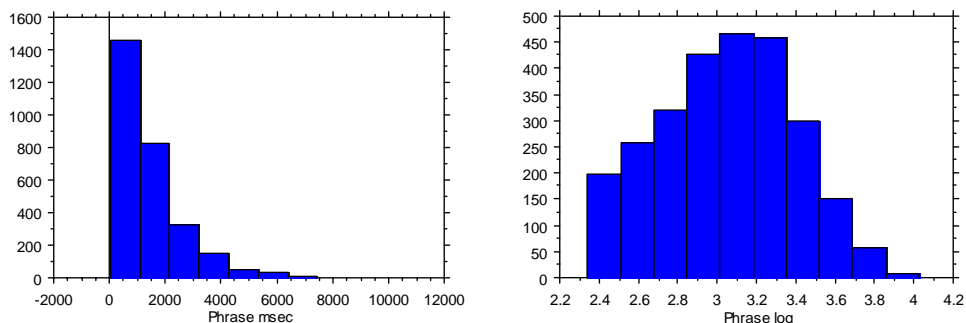


Figure 8a. Distribution of pauses in msec. Figure 8b. Distribution of pauses in log. values.

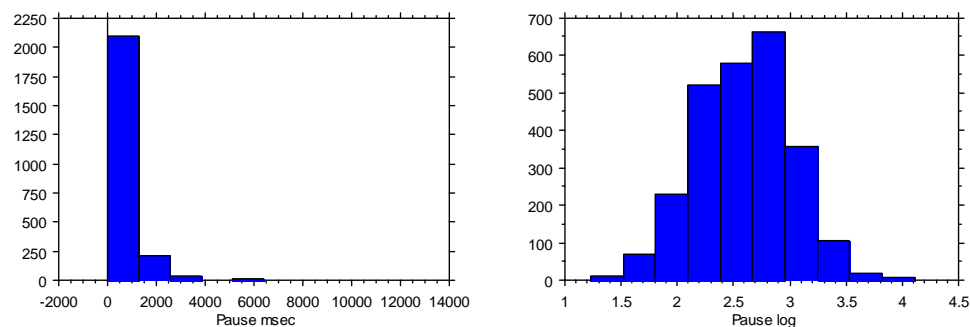


Figure 9a. Distribution of utterances in msec. Figure 9b. Distribution of utterances in log. values.

4. 2. Categories of Pauses

For half a century, pauses have been classified under two categories: short and long (Goldman-Eisler, 1958). More recently, by combining various data, i.e. of read speech from several European languages, and of dialogues in French, a three- category classification has been proposed by Campione and Véronis (2002).

For all three languages studied here, our data show a normal distribution in the log scale with two large flat sides. Actually, the limits of the high and low quartiles lie at 204 msec. and 689 msec., which means that 25% of pauses are shorter than 200 msec. and another 25% are longer than 700 msec.. Within such a distribution, it is legitimate to suppose that in each quartile, a secondary peak may be found in addition to the central peak. By applying the estimation method proposed by the afore mentioned authors, three areas of pauses have been determined as shown in Table I. Our estimate is in accordance with the values traditionally reported in the descriptions of pauses: around 200 –250 msec., 500 msec., and above 1 second.

	Short Pauses	Average Pauses	Long Pauses
German	210 msec	564 msec	1334 msec
Japanese	121 msec	434 msec	1355 msec
Korean	175 msec	513 msec	1171 msec

Table I. Thresholds estimates for the three categories of pauses.

Short pauses occur in our corpus when speakers use a fast speed or give a familiar routine explanation. Their durations roughly correspond to syllable duration, and from a perceptive standpoint, they serve as markers separating utterances. Mean pauses are used at normal turn-takings the typical case being silent pauses actually perceived as such. Long pauses, more specific to conversation, have not been observed in read speech. Pause duration increases with the complexity of projected production or of vocabulary, or when production or the understanding of context is difficult. In sum, the issue discussed here is the trace of difficulty in communication (Bestgen, 1998).

4. 3. Sociolinguistic Aspects

Besides the quantitative aspect, the sociolinguistic aspect of pauses is also necessary to understand the dynamic functionality of dialogues to compare cultural communities. For example, a portion of silence or a long silent pause does not have a positive value in the American society (Scollon 1985). However, in Japanese, silence does not have just a negative aspect to it (Sugiyama Lebra, 1989). Moreover, Hayashi & Kim (2002), in their study of silence portions superior to 1 second, have provided evidence of the importance of silence by the non respect of the interlocutors' speech. In German, it is considered impolite to interrupt someone but once the right to speak has been obtained, one can make the most of it. This sociolinguistic aspect can clearly be observed in our corpus. For example, dialogues unfold in a specific way in Japanese, i.e. interlocutors briefly indicate their approval with words and gestures; speakers and interlocutors complementing each other. In sum, a dialogue is a string of words shared and built in common (Mizutani, 1988).

Difference in expressivity has often been associated with difference of sex. However evidence of the influence of sex on silent pauses would has not been provided (Botinis et. al.

2003). The present study shows Korean is quite specific in that regard: the sex factor in relationship with the role factor reveals a statistically significant effect. That issue remains an open field for future studies on sociolinguistic behavior.

5. Conclusions

In this study, we have examined silent pauses in simulated request-refusal type dialogs for three different languages, i.e. German, Japanese, and Korean. The phonetic and statistical analyzes have led to the following findings:

- (1) There is a negative correlation between pause length and pause frequency
- (2) Speaker role appears a prominent factor for pause duration in all three languages
- (3) Pauses have more weight than utterances in interactional management.

Evidence has thus been provided of the crucial place of silent pauses in conversational strategies.

References

- Autesserre, D., Guaitella, I. & Nishinuma, Y. (1989): "Aerodynamic foundations for a new typology of silent pauses in dialogue", *Eurospeech 89*, 260-263.
- Bestgen, Y. (1998): "Segmentation markers as trace and signal of discourse structure", *Journal of Pragmatics*, 29, 753-763.
- Botinis, A., Gawronska, B., Erlendsson, B. & Torstensson, N. (2003): "Syntactic correlates and discourse correlates of pause distribution and duration in German, Swedish, and Icelandic", *ICPhS*, Barcelona, Spain, August 2003, 1229-1232.
- Botinis, A., Gawronska, B., Katsika, A. & Panagopoulou, D. (2003): "Prosodic speech production and thematic segmentation", *Phonum*, 9, 113-116.
- Campione, E. & Véronis, J. (2002): "A Large-Scale Multilingual Study of Silent Pause Duration", *Speech Prosody 2002*, 199-202.
- Candea, M. (2000): *Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané*, Thèse de doctorat, Université de Paris III.
- Dalton, P. & Hardcastle, W. J. (1977): *Disorders of Fluency and their Effects on Communication*, X, Edward Arnold, London.
- Duez, D. (2001): "Caractéristiques acoustiques et phonétiques des pauses remplies dans la conversation en français", *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, 20, 31-48.
- Duez, D. (1982): "Salient pauses and non salient pauses in three speech style", *Language and Speech*, 25(7), 11-28.
- Goldman-Eisler, F., 1972, "Pauses, clauses, sentences", *Language and Speech*, 15, 103-113.
- Goldman-Eisler, F., (1968): *Psycholinguistics: experiments in spontaneous speech*, Academic Press, London.
- Goldman-Eisler, F., (1961) : "The distribution of pause duration in speech", *Language and Speech*, 4 , 232-237.
- Goldman-Eisler, F., (1958): "Speech production and the predictability of words in context", *Quarterly Journal of Experimental Psychology*, 10, 96-106.
- Grosjean, F. & Collins, M., (1979): "Breathing, pausing, and reading", *Phonetica*, 36(2), pp.98-114.
- Grosjean, F. & Deschamps, A. (1975): "Analyse des variables temporelles de l'anglais et du français spontanés", *Phonetica*, 31, 144-184.
- Grosjean, F. & Deschamps, A. (1973): "Analyse des variables temporelles du français spontané", *Phonetica*, 28, 191-226.
- Grosjean, F. & Deschamps, A. (1972): "Analyse des variables temporelles du français spontané", *Phonetica*, 26, 130-156.
- Gustafson-Čapková, S. & Megyesi, B. (2001): "A Comparative Study of Pauses in Dialogues and Read Speech", *Eurospeech*, 931-935.
- Hansson, P. (1999): "Discourse Markers in Dialogue", *FONETIK* 99, 65-68.
- Hayashi, A., & Kim, Y. (2002) "Kontextualisierungshinweise in Gesprächen und deren Interpretationen - Kontrastive Analyse von deutschen, japanischen und koreanischen Gesprächen-" IV. *Ost-West-Kolloquium fuer Sprachwissenschaft der Ost-West-Gesellschaft fuer Sprach- und Kulturforschung e.V. Berlin, XII Forschungstagung zu Sprachdidaktik und Curriculumentwicklung des Instituts fuer*

- Sprache und Kultur der Konan Universitaet. Internationales Kolloquium "Sprachforschung und Sprachdidaktik".* 20, Kobe, Japan.
- Koiso, H. Den, Y. (2000) : "How is the Smooth Transition between Speakers Realized ? - Consideration Based on an Analysis of a Spoken Dialogue Corpus -", *Ninchi-Kagaku*, 7(1), 93-106.
- Koopmans-van Beinum, F. J. & Donzel, M. E. (1996): "Relationship between Discourse Structure and Dynamic Speech Rate", *ICSLP 96*, 1724-1727.
- Levelt, W. J. M. (1989): *Speaking: From intention to articulation*, Cambridge, Mass., MIT Press.
- Levin, H., Silverman, I. & Ford, B. (1965): "Hesitations in children's speech during explanation and description", *Verbal Learning and Verbal Behavior*, 6, 560-564.
- Limpert, E., Stahel, W. A. & Abbt, M. (2001): "Log-normal distributions across the sciences: Keys and Clues", *Bioscience*, 51(5), 341-352.
- Maynard K., S. (1997) *Japanese Communication: Language and Thought in Context*. Honolulu, University of Hawaii Pres.
- Mizutani, N. (1988) : "Aizuchiron", *Nihongo-gaku*, 7-13, 4-11 (in Japanese).
- Pfauwadel, M.-C. (1986): *Etre begue*, Paris, Retz.
- Saint-Bonnet, M. & Boe, J.-L., 1977, "Les pauses et les groupes rythmiques: leur durée et distribution en fonction de la vitesse d'élocution", *JEP*, 337-343, GALF
- Scollon, R. (1985): "The Machine Stops: Silence in the Metaphor of Malfunction", Tannen, D. & Saville-Troike, M. (eds.), *Perspectives on Silence*. 21-30. Norwood :Ablex
- Sugito, M., 1987, "Characteristics of durations and function of speech pauses in discourse", *Studies in Phonetics and Speech Communication*, 1, 53-68.
- Sugiyama Lebra, T. (1989) : "The Cultural Significance of Silence in Japanese Communication", *Multilingua: Journal of Cross-Cultural and Interlanguage Communication*, 6-4, 343-357.
- Weilhammer, K. & Rabold, S. (2003): "Durational Aspects in Turn Taking", *ICPhS*, 931-934.
- Van Donzel, M. & Koopmans-van Beinum, F. J. (1996): "Pausing Strategies in Discourse in dutch", *ICSLP 96*, 1029-1032.
- Zellner, B. (1994): "Pauses and the temporal structure of speech", in E. Keller (Ed.) *Fundamentals of speech synthesis and speech recognition*, 41-62, John Wiley.