# TEMPORAL VARIABLES IN LECTURES IN THE JAPANESE LANGUAGE

*Michiko Watanabe*

Graduate Division of International and Interdisciplinary Studies
The University of Tokyo
mich@mx5.nisiq.net

## ABSTRACT

In second language input studies, speaking speed is regarded as one of the most influential factors in comprehension. However, research in this area has mainly been conducted on written texts read aloud. The present study investigated temporal variables, such as articulation rate and ratio and frequency of fillers and silent pauses, in three university lectures given in Japanese. It was found that the total duration ratio of fillers was as great as that of silent pauses. It also became clear that, for individual speakers, articulation rate and frequency of fillers are relatively constant, while frequency of silent pauses varies depending on discourse section. Of total pause ratio, pause frequency and articulation rate, the latter correlated best with listener ratings of speech speed. The findings suggest that spontaneous speech requires methods of speech speed measurement different from those for read speech.

## 1. INTRODUCTION

In foreign language learning and teaching studies, speech speed has been regarded as one of the most influential factors in comprehension. Previous studies suggest that slower speed enhances comprehension (Flowerdew 1994). On the other hand, it has been shown that exaggeratedly slow delivery is a hindrance rather than a help (Griffith 1990).

The term most commonly used in the field to express speech speed is speech rate, which is defined as number of words, syllables or morae uttered per unit time including pauses. Here, however, the term articulation rate will be used to denote number of words, syllables or morae uttered per unit time exclusive of pauses. In earlier studies, overall speech rate was the major concern, and experiments were conducted in which speech rate was modified by using machines to vary the speed of all speech stimuli by a uniform rate. Comprehension tests results were then compared. Grosjean, however, (1972, reported in Griffiths 1990) was unique in manipulating articulation rate, pause frequency and pause duration individually. It was consequently found that the manipulation of the first two parameters in particular led to significant increases in comprehension test scores.

Thanks to recent technical development, speech can now be analysed more easily and precisely (Kiritani 1997). More sophisticated, pinpoint experiments have become possible. Of the three variables listed above, pause frequency has attracted the greatest attention. By moderating number of pauses, Sugito (1990) demonstrated that pause frequency influenced the degree of understanding and recall of texts; those with more pauses were better recalled. Kohno (1994) reported that texts with clause and phrase pauses were better understood than those with sentence pauses only, while those with pauses after each word were less comprehended. Hirozane (1994) found that texts with less frequent pauses were perceived as faster when articulation rate and total pause duration were held constant. He concluded that perceived speech speed was affected not only by the average quantity of phonetic signals but also by pause distribution.

Compared with pause phenomena, articulation rate has been relatively little studied. Kohno (1994:89) argued that the influence of articulation rate on listener perception of speech speed was slight compared with that of pauses. Grosjean and Lane (1974), however, argued that articulation rates were as important to listeners' subjective perception of speech speed as pause frequency. Kimura (1997) reported that an experimental group which listened to a text with an articulation rate 1.2 times lower than the original, but without any change in the number or length of pauses, scored significantly higher in a comprehension test than a control group which listened to the original text, and argued that articulation rate also affected comprehension.

Concerning real life listening, one limitation of previous studies is that most of them are based on read speech. Spontaneous speech has features such as hesitation, repeats and false starts which are not found in read speech. It has been observed that such features hinder learners' comprehension (Voss 1979, Fukao *et al.* 1991). To understand real life listening, it is indispensable to investigate spontaneous speech. Another problem with previous studies is that it is often impossible to compare their results because the measurement method and the way of moderating speech speed vary from study to study. Griffiths (1991) suggested that researchers should adopt comparable and reproducible methodology based on speech science.

# •. PRESENT STUDY

In the present study, the temporal variables of three university lectures in Japanese were investigated. Lectures were chosen as material because they are important comprehension targets for overseas students. As lectures are usually prepared, they cannot be said to be completely spontaneous. However, in the sense that the lecturer adjusts his or her address to the student audience according to the latter's previous knowledge and response, they are different from read speech.

## 2.1. PURPOSES

One purpose of the present study was to investigate hesitation phenomena in Japanese-language lectures. As mentioned earlier, these phenomena have been reported to cause comprehension problems for non-native listeners. Knowledge of the quantitative and qualitative characteristics of hesitation would therefore certainly appear to be useful for both learners and teachers. The present study concentrated on the quantitative aspect.

The other purpose was to find out which temporal variable correlates best with listener ratings of speech speed. The same speech material was used as stimuli in the author's previous research study of lecture comprehension (Watanabe 1997). In the latter study, subjects were asked after listening to passages to rate the appropriateness of the speaker's speed to their comprehension level. Interestingly, the ratings given by a non-native speaker group and a native speaker group were very similar. One lecture (lecture 2) was rated as slightly fast, another (lecture 3) as appropriate and the third (lecture 1) as intermediate by the majority of both subject groups. Speech rates did not correspond to listener ratings. The individual factors of speech rate therefore need to be more closely examined.

## 2. 2. MATERIAL

Three video-taped university lectures were used. The titles were 'Introduction to Japanese Language Studies', 'Economic Systems' and 'Science History', referred to hereinafter as lectures 1, 2 and 3 respectively. All were introductory lectures. In the author's previous research, excerpts from each lecture with a duration of between seven and eight minutes were used as stimuli for comprehension tests. In the present study, to allow comparison of temporal factors within the speech of the same individual as well as between speakers, two speech segments were further extracted from each excerpt for measurement. In each case, the first segment was the part where the lecturer introduced the topic of the day, while the second was a passage where the lecturer talked about one of the main subjects. The duration of each speech sample was as follows:

Lecture 1-1 (first segment of lecture 1): 36 sec.
Lecture 1-2 (second segment of lecture 1): 36 sec.
Lecture 2-1 (first segment of lecture 2): 40 sec.
Lecture 2-2 (second segment of lecture 2): 31 sec.
Lecture 3-1 (first segment of lecture 3): 44 sec.
Lecture 3-2 (second segment of lecture 3): 45 sec.

## 2.3. METHODS OF ANALYSIS

The speech samples were analysed on a personal computer using the *Onseirokubunken* speech analysis system. The following elements of each excerpt were measured: the duration of continuous

streams of speech between silent pauses (hereinafter called breath groups), the number of morae in each breath group, the duration and number of silent pauses of at least 150ms, and the same two parameters for fillers. The term filler is used to denote expressions of hesitation, such as *ah, er, um, mm* in English and *ano, ma, eeto, e* in Japanese.

Articulation rate was computed in the following way: the number of morae per second was first calculated for each breath group, then averaged for each speech sample. The averaged value is here called articulation rate. Fillers were omitted in the calculation of articulation rate. The ratio of the duration of silent pauses to the length of the whole speech segment and the ratio of fillers to the length of the whole speech segment was computed. These measurement methods basically follow Misono and Kiritani's (1994) study of two lectures in Japanese. In the present study, however, fillers were investigated separately from silent pauses, whereas in Misono and Kiritani's study, they were treated as a variety of pause.

# 3. RESULTS



Figure 1: percentage of total duration

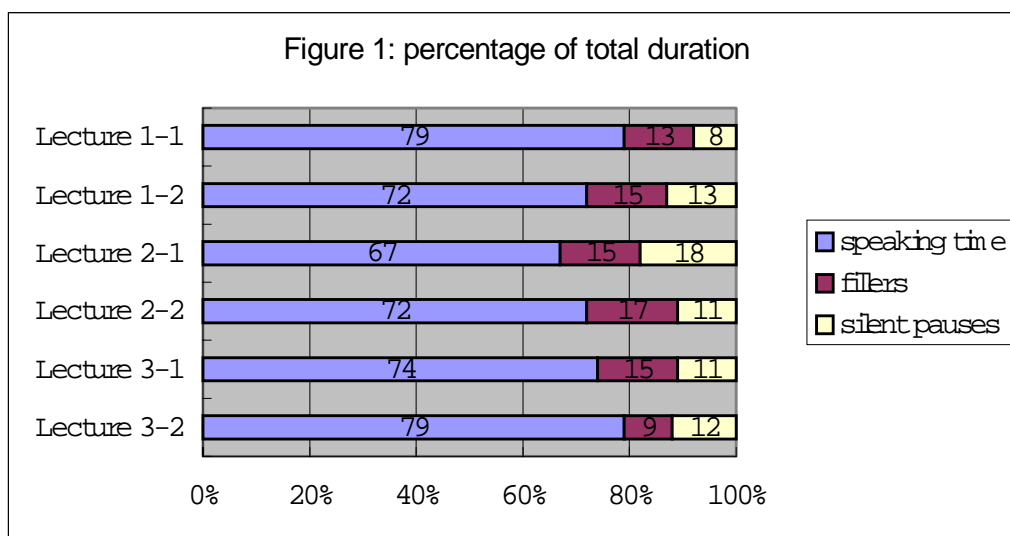| | speaking time | fillers | silent pauses |
|---|---|---|---|
| Lecture 1-1 | 79 | 13 | 8 |
| Lecture 1-2 | 72 | 15 | 13 |
| Lecture 2-1 | 67 | 15 | 18 |
| Lecture 2-2 | 72 | 17 | 11 |
| Lecture 3-1 | 74 | 15 | 11 |
| Lecture 3-2 | 79 | 9 | 12 |

Figure 1 shows the percentage duration of fillers and silent pauses in each speech sample. It establishes that the ranges of the ratios of fillers and of silent pauses in the six samples are almost identical: 9 to 17% for fillers and 8 to 18% for silent pauses. The sum of the ratios of fillers and silent pauses is between 21 and 33%, which is a little less than the 33 to 38% found by Misono and Kiritani (1994). No excerpt had a particularly elevated ratio of fillers or silent pauses. Of the three lectures, lecture 2 had the smallest speaking time ratio on average.



Figure 2: Frequency of fillers and silent pauses

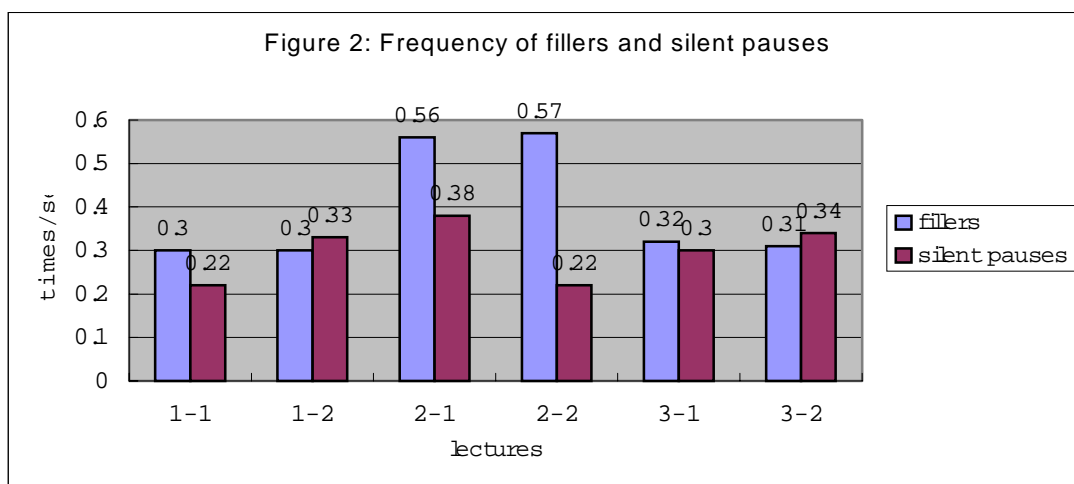| lectures | fillers | silent pauses |
|---|---|---|
| 1-1 | 0.3 | 0.22 |
| 1-2 | 0.3 | 0.33 |
| 2-1 | 0.56 | 0.38 |
| 2-2 | 0.57 | 0.22 |
| 3-1 | 0.32 | 0.3 |
| 3-2 | 0.31 | 0.34 |

Figure 2 shows frequencies of fillers and silent pauses. Fillers appear as often as silent pauses with

two of the speakers, but twice as often with lecturer 2. While pause frequencies vary more within the speech of the individual than among speakers, filler frequencies are quite constant for individual speakers but vary between speakers.
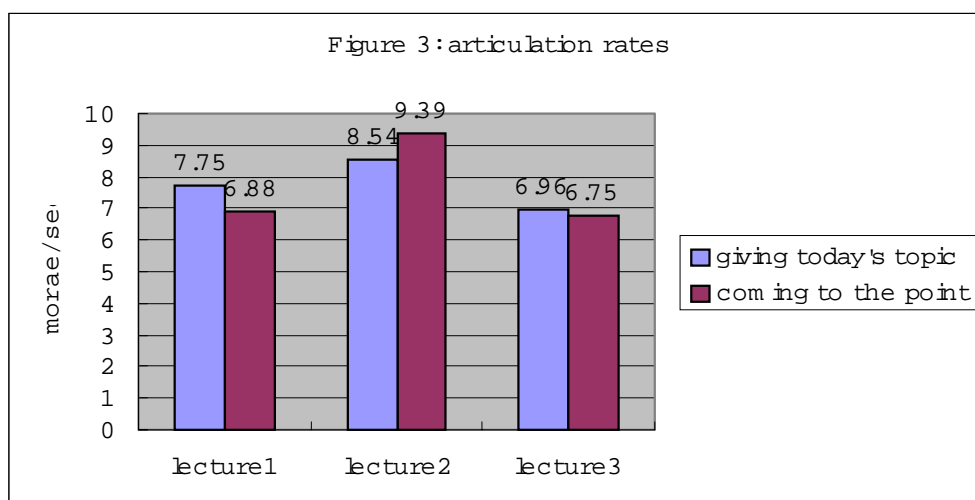


Figure 3 shows the articulation rates of the six speech excerpts. The range is 6.8 to 9.4 morae/sec. Misono and Kiritani's (1994) results also fall in this range. Articulation rates vary more between speakers than within the speech of the individual. Lecture 2 has the highest articulation rate and lecture 3 the lowest. This order correlates with the speed rating obtained in the author's previous study.

**Table 1:** The mean duration of breath groups and the average number of morae

| lectures | BG (sec) | SD | morae | SD |
|---|---|---|---|---|
| 1-1 | 1.59 | 1.14 | 12.3 | 9.3 |
| 1-2 | 1.43 | 1.09 | 9.8 | 9.57 |
| 2-1 | 1.19 | 0.73 | 9.67 | 6.23 |
| 2-2 | 1.12 | 0.85 | 9.96 | 7.49 |
| 3-1 | 1.44 | 0.93 | 10.04 | 7.31 |
| 3-2 | 1.4 | 0.71 | 9.82 | 5.86 |

For the six speech segments, Table 1 shows the mean duration of breath groups and the average number of morae per breath group (incl. standard deviation). The values for mean duration of breath group indicate that this parameter is speaker-dependent. Lecturer 2 has the shortest mean duration, and lecturer 1 the longest. On the other hand, there is only a slight difference between speakers in the average number of morae. These results correspond to the differences in articulation rates.

## 4. CONCLUSIONS AND DISCUSSION

In the lectures investigated, the percentage duration of fillers was as high as that of silent pauses. The frequency of fillers was at least as high as that of silent pauses. These findings show that the quantity of fillers is not negligible. It seems that the influence of fillers in comprehension is worth further investigation. Whether a listener takes a rest during filled pauses or attempts to identify meanings may make a great difference to the listening process, as more pauses enhance comprehension.

Another finding was that filler frequency and articulation rate are speaker-dependent and relatively constant in a given speech while frequency of silent pauses varies within the speech of the individual depending on discourse section. For read speech such as news casts and text readings, the unit known as speech rate is most commonly used to express speech speed. However, the results of the

present study indicate that speech rate is not a reliable index for more spontaneous speech because the percentage duration of pauses varies within a single speech act. It would seem that, in studying the speed of spontaneous speech, there is a need to distinguish between quantity of phonetic signals uttered per unit and sectional speed of delivery. Speech rate expresses the former, but not the latter. Speech with high articulation rate and many pauses may have the same speech rate as speech with low articulation rate and few pauses. The application of articulation rate seems to be indispensable to accurately express speech speed under such circumstances.

Among the temporal variables examined, only articulation rate showed correlation with listener ratings of speed. This result suggests that articulation rate may be the best index of listener perception of speed of spontaneous speech. However, more controlled study is needed before any conclusion can be drawn about the most influential factors in listener perception and comprehension of spontaneous speech.

# 5. REFERENCES

1. Flowerdew, J. "Research of relevance to L2 lecture comprehension," *Academic Listening*, Flowerdew, J. (Ed.), Cambridge, Cambridge University Press: 7-29, 1994.

*2.* Fukao, Y., Mizuta, S., and Ohtsubo, K. "Developing listening comprehension material for academic purposes for advanced learners of Japanese," *Proceedings of autumn meeting of the society for teaching Japanese as a foreign language* : 37-42, 1991.

3. Griffiths, R. "Speech rate and NNS comprehension: a preliminary study in time-benefit analysis," *Language Learning* 40 (3): 311-336, 1990.

4. Griffiths, R. "Pausological research in an L2 context: A rationale, and review of selected studies," *Applied Linguistics* 12 (4): 345-364, 1991.

5. Grosjean, F. *Le rôle joué par trois variables temporelles dans la compréhension orale de l'anglais étudié comme seconde langue et perception de la vitesse de lecteurs et d'auditeurs*, unpublished doctoral dissertation, Université de paris •, Paris, France, 1972.

6. Grosjean, F. and Lane, H. "Effect of two temporal variables on the listener's perception of reading rate," *Journal of Experimental Psychology* 102: 893-896, 1974.

7. Hirozane, Y. "Effects of the number of pauses on perceived rate of speech," *Journal of the Phonetic Society of Japan* 205: 63-65, 1994.

8. Kimura, S. "Reconsideration of the influence of speech speed on Listening Comprehension," *Journal of Language Processing* 1: 60-68, 1997.

9. Kiritani, S. "Speech analysis using personal computer," *Accent, Intonation, Rhythm and Pauses*, Kunihiro, T., Hirose, H. and Kohno, M. (Eds.) , Tokyo, Sanseido, 1997.

10. Kohno, M. *Reports of studies on the role of rhythm in recognition and production of speech, supported by a grant form the Ministry of Education, Science and Culture,* 1994.

11. Misono, Y. and Kiritani, S. "Characteristics of pauses in Japanese lecture-style speech: Comparison with newscasts," *International Journal of Psycholinguistics* 10(2): 167-180, 1994.

12. Sugito, M. "On the role of pauses in production and perception of discourse," *Proceedings of ICSLP* 1990: 513-515, 1990.

13. Voss, B. "Hesitation phenomena as sources of perceptual errors for non-native speakers," *Language and Speech* 22(2): 129-144,1979.

14. Watanabe, M. *Factors influencing understanding of lectures by international students in Japan*, unpublished master's thesis, Tokyo University of Foreign Studies, 1997.