

TOWARDS AN INTEGRATED UNDERSTANDING OF SPEECH OVERLAPS IN CONVERSATION

Jiahong Yuan, Mark Liberman, Christopher Cieri

University of Pennsylvania

jiahong@ling.upenn.edu, myl@cis.upenn.edu, ccieri@ldc.upenn.edu

ABSTRACT

We investigate factors that affect speech overlaps in conversation, using large corpora of conversational telephone speech. We analyzed two types of speech overlaps: 1. One side takes over the turn before the other side finishes (turn-taking type); 2. One side speaks in the middle of the other side's turn (backchannel type). We found that Japanese conversations have more short turn-taking type of overlap segments than the other languages. In general, females make more speech overlaps of both types than males; and both males and females make more overlaps when talking to females than talking to males. People make fewer overlaps when talking with strangers than talking with familiars, and the frequency of speech overlaps is significantly affected by conversation topics. Finally, the two conversation sides are highly correlated on their frequencies of using turn-taking type of overlaps but not backchannel type.

Keywords: overlap, interruption, backchannel, conversational speech, corpus phonetics

1. INTRODUCTION

We investigate speech overlaps (two speakers speaking simultaneously) in conversation. Previous studies have shown contradictory results on gender as well as linguistic/cultural differences in speech overlaps. Kajikawa [6] compared the overlap ratios in English and Japanese conversations reported in the literature and stated that Japanese adult conversations have more frequent overlaps (45% in Japanese vs. 5% in English). In a comparison study of turn-taking in English and Japanese, Furo [4] showed, however, that the Japanese data have fewer instances of overlaps than the English data.

Within the language and gender literature one of the most contested areas is whether men interrupt their conversation partners more often than do women. Totally different results have been obtained (review: [5], [1]): 1. men use more interruptions; 2. women use more interruptions; 3. no difference between the sexes. 'interruption' has been distinguished from 'overlap' from different perspectives (review: [4]): 'interruption' is 'deeper

intrusion' and longer than 'overlap' [11]; 'interruption' refers to the participants' perception whereas 'overlap' is essentially a descriptive term [2]; etc. Although 'overlap' and 'interruption' are different, they are highly related, we hope our analysis of overlaps can also shed light on the question of how men and women are different in using interruptions.

Besides linguistic/cultural factors and sex, situational factors such as conversation topic, speaker relationship, etc. may also affect the participants' use of speech overlaps, as suggested in many studies [4, 1, 11]. In this study, we investigate the effects of all these factors on speech overlaps using large speech corpora. We regard this work as a first step towards a comprehensive analysis. Because our research is based on published corpora, we look forward to future investigations that will supplement or revise our results.

2. DATA AND METHOD

We make use of large corpora of conversational telephone speech, including the CallHome corpora in Arabic (LDC97S45), English (LDC97S42), German (LDC97S43), Japanese (LDC96S37), Mandarin (LDC96S34), and Spanish (LDC96S35), which contain conversations between family members and friends, and the Fisher English corpus (Part I, LDC2004S13), which contains conversations between strangers. We excluded the conversations in which there are more than one person talking at one side and the conversations that have poor recording quality, and used 69, 163, 79, 64, 44, and 103 conversations from the Arabic, English, German, Japanese, Mandarin, and Spanish CallHome corpora respectively, and 5455 conversations from the Fisher corpus. For each conversation, we analyzed five-minute data, from minute 3 to minute 8.

To automatically obtain the speech overlaps of the two sides in a conversation, we segmented each side into pause and speaking segments using intensity as the criterion (using a Praat script from

[8], and time stamped the overlap segments in which both sides were speaking. We distinguished two types of speech overlaps in our analysis: 1. One side takes over the turn before the other side finishes (Turn-taking type); 2. One side speaks in the middle of the other side's turn (Backchannel type).

To verify the performance of the automatic procedure of obtaining speech overlaps, we randomly selected four conversations from the English, Japanese, and Mandarin CallHome corpora respectively, and hand labeled pause and speaking segments in the conversations. The overlaps calculated from using the manual segmentations are comparable to the results obtained by using the automatic segmentations by Praat: The mean of the Turn-taking type overlaps in these conversations is 10.6 by hand vs. 9.2 by Praat; and the mean of the Backchannel type is 11.6 by hand vs. 9.0 by Praat. The individual numbers of overlaps, for both types, are shown in Figure 1.

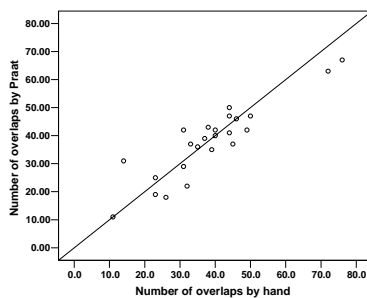


Figure 1. Number of overlaps by hand vs. by Praat.

In the following analysis we calculated for each conversation the speech overlaps made by each of the conversation sides respectively. For example, if A takes over the turn before B finishes, then the overlap (Turn-taking type) belongs to A but not B; and if A speaks in the middle of B's turn, the overlap (Backchannel type) is A's but not B's. Using the overlaps of each conversation side instead of the entire conversation makes our analysis consistent across all the factors, including language, speaker relationship, conversation topic, speaker sex, and correlation between conversation sides.

3. RESULTS

3.1. Language differences

Figure 2 shows the mean number of overlaps made by one conversation side in 5-minute

conversation. It is interesting that Japanese has more short turn-taking type of overlaps than the other languages. One possible explanation for this is that Japanese has more frequent speaker changes than the other languages [4], and hence have more utterances in a five-minute conversation. Therefore, the proportion of the utterances overlapped in Japanese would not be very different from the other languages.

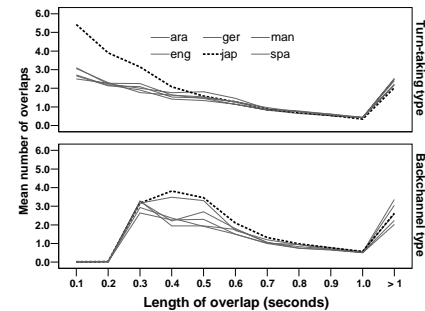


Figure 2. Number of overlaps made by one conversation side in 5-minute conversation.

As shown in Figure 3, Japanese does have more utterances than the other languages in the same length of conversation. However, the proportion of utterances having a short turn-taking overlap in Japanese is still significantly higher than the other languages, as shown in Figure 4.

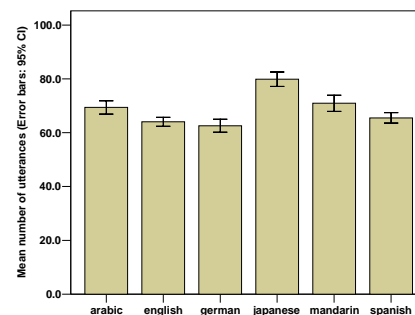


Figure 3. Number of utterances made by one conversation side in 5-minute conversation.

It has been claimed that there are more backchannels in Japanese conversations than in English ([4], [7]). A striking result from our analysis is that Japanese has more short Turn-taking type of overlaps than the other languages.

Overlaps of the turn-taking-type are not necessarily destructive. There are constructive turn-taking overlaps, in which a listener talks along with the speaker not to interrupt the speaker, but to show enthusiastic listenership and participation [10]. The constructive turn-taking overlap is an instance of high-involvement interaction. Japan is,

however, often seen as a high consideration culture [4]. Further studies will be needed to determine whether the fact that Japanese has more short Turn-taking type overlaps is a cultural characteristic or a linguistic characteristic, or both.

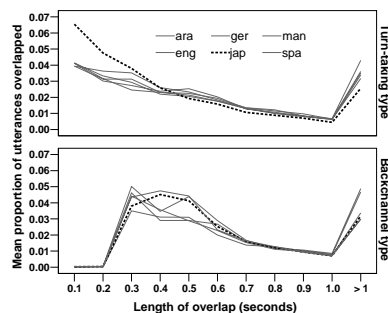


Figure 4. Proportion of utterances overlapped made by one conversation side in 5-minute conversation.

3.2. Effects of speaker sex

From our analysis, in the CallHome corpora females tend to use slightly more overlaps of both types than males. Although the difference is consistent across all the six languages, it is not statistically significant. This is probably because the effects of speaker sex on speech overlaps are dependent on many other factors, for example, conversation topics, relationship between the speakers, the sex of the other conversation side, etc, and the CallHome corpora are not large enough for these factors to be balanced. To explore the real effect of speaker sex, we use the much larger Fisher corpus.

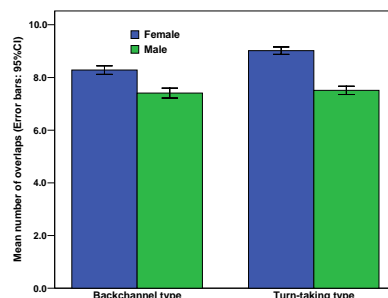


Figure 5. Number of overlaps made by females vs. males speakers in 5-minute conversation in Fisher.

Figure 5 shows the difference between the two sexes in Fisher. Clearly females make more frequent overlaps than males, which is consistent with the differences shown in the CallHome corpora. In Figure 6, we add the sex of the other conversation side, and compare four categories: females when talking with a female, females when talking with a male, males when talking with a

female, and males when talking with a male. We can see that both males and females tend to make more speech overlaps when talking to females than talking to males.

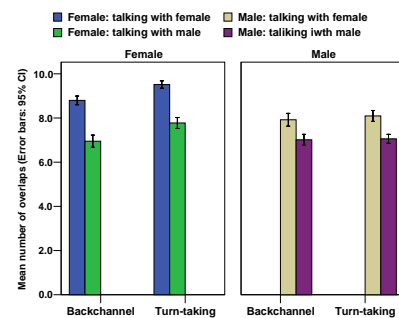


Figure 6. Number of overlaps made by female vs. male speakers when talking a female vs. a male.

Our analysis shows that it's not that men are more likely to interrupt and overlap the others' speech than women, it is that women's speech are more likely to be interrupted and overlapped than men's. This might shed some lights on explaining the contradictory results in the literature about who (men or women) tend to use more interruptions.

3.3. Effects of speaker relationship and conversation topic

From Figure 7 we can see that compared with the CallHome English corpus, the speech overlaps in the Fisher corpus are much less frequent for both males and females. That is, when talking with strangers, people use less overlaps of both types.

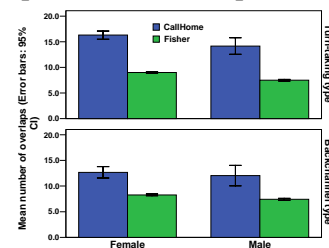


Figure 7. Number of overlaps in CallHome vs. Fisher.

We also examined the effects of conversation topics on speech overlaps. The mean number of overlaps ranges from 6.1 to 9.3 for the backchannel type of overlaps, and from 5.8 to 10.3 for the turn-taking type of overlaps across different topics. The effects of conversation topics on the two types of overlaps seem to be consistent: if there are more Backchannel overlaps in a conversation topic, there will be also more turn-taking overlaps in the same topic. This result is

shown in Figure 8, in which the mean number of overlaps for each conversation topic is drawn on a scatter plot. We can see that there is a linear correlation between the two types of overlaps across conversation topics.

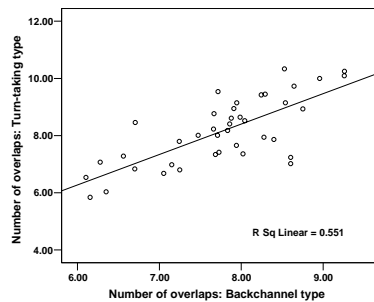


Figure 8. Correlation between the two types of overlaps across conversation topics.

3.4. Correlation between the conversation sides

In a conversation, if one side makes frequent speech overlaps, will the other side do so too? By doing a correlation test between the overlaps made by the two conversation sides, we found that the two sides are more correlated on the turn-taking type of overlaps ($r = .544$) than on the backchannel type of overlaps ($r = .229$). Because conversation topics significantly affect both types of overlaps, we also did correlation tests for each conversation topic separately, the results are shown in Figure 9. It is clear that for all the conversation topics the correlation between the two conversation sides is higher on the turn-taking type of overlaps than on the backchannel type. All the associated p-values are less than .0001 for the correlation tests on the turn-taking type of overlaps; on the backchannel type, however, 20 of the 40 associated p-values are larger than .01, and 8 of them are larger than .1.

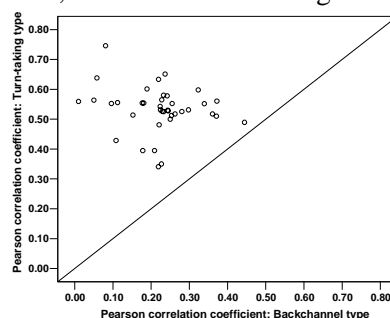


Figure 9. Correlation between the two conversation sides.

Our analysis shows that if one conversation side makes more turn-taking type of overlaps, so does the other side. Perhaps this is a sort of

accomodation to conversational style, or perhaps more animated conversations lead to more turn-taking overlaps by both participants. The two sides are less correlated in the use of backchannel-type overlaps, perhaps because backchannels are sometimes used mainly on one side of an asymmetrical conversation.

4. DISCUSSION

The organization of turn-taking has been widely studied since [9] and [3], including when and how an overlap may occur in conversation. There are many systematic reasons for the occurrence of an overlap, including competing for an early start for the next turn, projectability of possible completion or transition-relevance places [9]. This study was based on an automated (and extremely coarse) binary division of overlap types. For a better understanding of our results, for example why there are more short turn-taking type of overlaps in Japanese telephone conversations, we will need to make finer distinctions in the nature and function of overlaps.

5. REFERENCES

- [1] Anderson, K.J., Leaper, C. 1998. Meta-analysis of gender effects on conversational interruption: who, what, when, where, and how. *Sex Roles*, 39, 225-252.
- [2] Bennett, A. 1981. Interruptions and the Interpretation of Conversation. *Discourse Processes* 4, 171-188.
- [3] Duncan, S. 1972. Some signals and rules for taking speaking turns in conversations. *Journal of personality and social psychology*, 23, 283-292.
- [4] Furo H. 2001. *Turn-Taking in English and Japanese*. Routledge: New York.
- [5] James, D., Clarke, S. 1993. Women, men, and interruptions: a critical review. In: Tannen, D. (eds), *Gender and Conversational Interaction*. Oxford University Press, 231-280.
- [6] Kajikawa, S., Amano, S., Kondo, T. 2004. Speech overlap in Japanese mother-child conversations. *Journal of Child Language*, 31, 1-16.
- [7] Maynard, S.K. 1986. On back channel behavior in Japanese and English casual conversation. *Linguistics* 24, 1079-1108.
- [8] Miitta Lennes' Praat scripts. <http://www.helsinki.fi/~lennes/praat-scripts/>
- [9] Sacks, H., Schegloff, E., Jefferson, G. 1974. A simple systematics for the organization of turn-taking for conversation. *Language*, 50, 696-735.
- [10] Tannen, D. 1993. *Gender and discourse*. New York: Oxford University Press.
- [11] Zimmerman, D., West, C. 1975. Sex roles, interruptions and silences in conversation. In Thorne, B., N. Henley. N. (eds.), *Language and sex: differences and dominance*. Rowley, MA: Newbury House, 105-129.