Cross-dialectal Turn Exchange Rhythm in English Interviews

Janice Fon

Graduate Institute of Linguistics National Taiwan University, Taiwan jfon@ntu.edu.tw

Abstract

This study investigates factors underlying the exchange rhythm in Singapore English using six cross-dialectal interviews from the NIECSSE corpus. Exchange intervals (EIs), defined as the latency interval between the onsets of an exchange pair, were measured and two exchange types, question-answer and confirmation, were identified and labeled using Praat. Results showed that EIs in Singapore English were generally limited to a narrow range. Over 90% of the turn exchanges were unmarked-next-position, making this dialect closer to Midwestern and Californian than New York English. In addition, EIs were reflective of the cognitive load. Exchange pairs requiring more cognitive processing tend to have longer EIs, rendering a mismatch between adjacency coupling and rhythmic structure. Due to different levels of social insecurity and social expectations, EIs also vary with gender, with female speakers having EIs twice as long as males. However, individual talker style did not seem to be a deciding factor when all other potential factors were partialled out.

1. Introduction

Unlike written dialogs, which contain punctuation marks that clearly demarcate exchange boundaries, spoken language often obscures its punctuation equivalents, if one believes they even exist in the first place. Previous studies showed that when participants of a conversation are from the same dialect or with the same type of personality, they are able to cooperate and coordinate so that most of the conversations are conducted without going much awry; on the other hand, when participants are from different dialects or are with very different personalities, conversations might not be conducted as smoothly due to different exchange paces and negative impressions of conversation partners might thus be formed ([1] & [2]) It is therefore interesting to examine how exchange tempo is patterned in spontaneous conversation.

Previous studies on turn-exchanges mainly came from a discourse perspective. Researchers were in general more concerned about how speakers in a conversation compete for the floor. Sacks et al [3] proposed a speech exchange system which claimed that turn-taking is governed by a local management system. Speakers realize transition relevance places (TRPs) for potential turn-exchanges, and thus it is possible for conversation participants to avoid overlaps to a large extent. Subsequent studies have shown that there are certain linguistic features, mainly prosodic, that are relevant for TRPs (e.g., English: [4], [5], & [6]; German: [7] & [8]; Dutch: [9]), although some have also mentioned vowel quality as a potential cue [4].

However, it is undeniable that even with cues at TRPs, overlapping speech still occurs in our everyday life at a fairly frequent rate. Based on whether and how turn-exchanges overlap, Jefferson [10] proposed three categories, *transitional*, *recognitional*, and *progressional*.

Transitional onsets refer to turn-exchanges that come at or after a TRP, which is the majority of cases. There are three subtypes, unmarked-next-position onsets, latched onsets, and terminal onsets. An unmarked-next-position onset occurs when Speaker A continues until he reaches a TRP and stops. After a pause, Speaker B begins. A latched onset, on the other hand, is similar to the unmarked onset, but there is no perceptual pause in-between the two turns. Finally, terminal onsets refer to exchanges where the second speaker comes in a little before the TRP, resulting in a slight overlap between the last few segments of Speaker A's turn and the first few segments of Speaker B's turn. Jefferson [10] claimed that the first type is the most common. However, using a Thanksgiving dinner conversation, Tannen [1] showed that preference can be dialect-dependent. Midwesterners and Californians indeed prefer unmarked-next-position onsets to convey attentiveness, but New Yorkers, especially New York Jews of Eastern European descent actually prefer latched or even terminal onsets to show involvement, rapport, and interest.

Recognitional and progressional onsets come substantially before a TRP, the difference depending on the location of the cut-in point. The former refers to a turn exchange that occurs when the hearer decides that he has received enough information to respond, while the latter refers to a cut-in that occurs at a hesitation or stutter point. There are two subtypes of recognitional onsets, *item-targeted* and *thrust-projective*. The former occurs when only one word is overlapped while the latter refers to an overlap that may include a whole phrase.

TRPs for turn exchanges are generally indicated by a battery of fixed intonation patterns. However, the exact realizations seem to depend on individual languages and even dialects of the same language. For example, Wells & Macfarlane [6] identified two TRP-projecting accent patterns in West Midland English. One has a low nuclear accent and a low boundary tone, and the other has a high nuclear accent and a low boundary tone. However, for Tyneside English, Local *et al.* [4] recognized two different patterns, one a pitch step-up and the other a pitch drop at the end of the turn, along with accompanying cues in duration, amplitude, and vowel quality.

Based on previous studies, it is evident that most turn initiations are transitional onsets. Languages have different ways of phonetically coding turn initiations and endings and speakers can readily produce and recognize these cues so that conversations can gracefully proceed in most cases. However, it is unclear from these studies how exchange pace is determined. Specifically, from the viewpoint of an incoming speaker, once a TRP is identified, how much latitude in pause is he granted before it is considered too late to chime in harmoniously? Is there an optimal pause size, and if there is, is it dependent on dialect preferences [1], personal styles [2], or other factors such as cognitive loading [11]? Although Jefferson [10] subcategorized transitional onsets into three subtypes based on the existence of empty beats in-between the two turns, her model does not predict how different underlying mechanisms might be at work in determining onset types.

2. Aims

There are four specific aims in this study. First, we would like to see which type(s) of transitional onsets is(are) the most preferred in Singapore English, and consequently, the optimal range of exchange pauses. If Singapore English is more concerned with showing attentiveness, like Midwestern or Californian English, then we should find most of the transitional onsets to be unmarked-next-position [1]. On the other hand, if Singapore English values involvement, rapport, and interest more, then it should pattern more like New York English in that latched and terminal onsets are preferred.

Secondly, we would like to examine whether exchange pauses are reflective of cognitive loading, which would predict that different exchange types should be accompanied by different exchange pause duration. Specifically, questionanswer pairs should then have longer exchange pauses than confirmation pairs, as the former requires heavier cognitive processing than the latter. However, Scollon [2] claimed that exchange pauses are impervious to cognitive loading, which would predict that there should be no difference in pausing between question-answer and confirmation pairs.

Thirdly, we would like to also look into possible gender effects. As Asian females are in general socially more insecure than males [12], their exchange pauses might also be able to reflect this, indicating a gender difference. If this is the case, we would predict exchange pauses for female speakers to be longer than those for male speakers.

Finally, we would like to see if exchange pauses are reflective of individual speaking styles. This would predict a talker effect indicating high intra-talker reliability in addition to high inter-talker variability when other potential factors are partialled out.

3. Method

3.1. Data

Data included 6 interview recordings of approximately 5 min each, 3 males and 3 females from the NIECSSE corpus [13]. The interviewer was a British male, speaking British English (B), and the interviewees were all Singaporean, speaking Singapore English (S), the official language of Singapore. All of the interviewees are very fluent if not considered native speakers of English. They all acquired Mandarin and/or Hokkien (a Chinese Min dialect) as their first language(s), but started acquiring and using English on a daily basis after they entered the elementary school system. Although most of them still use Mandarin and other Chinese dialects to talk to their family members, some even to their friends, all of them use English regularly and consistently to talk to friends and strangers, and some even to their brothers and sisters.

3.2. Measurements

In an interview, there are two directions of exchanges: from an interviewer to the interviewee ($B\Rightarrow$ S), and from the interviewee to the interviewer ($S\Rightarrow$ B). To make analyses more interpretable, only the former is considered in this study since we have only one interviewer, and $S \Rightarrow B$ would not be very representative and reliable. Two types of turn exchanges were identified and included in this study, question-answer (QA) and confirmation (CF). QA pairs are exchanges in which one speaker poses a question which the other speaker answers. A turn is qualified as a question if it is syntactically phrased as such, as in (1), or if it is said with a question intonation, as shown in Figure 1. It is apparent from the pitch track that Subject M2 was using a final rise on *then* (in addition to the diminishing amplitude on the word and the hesitant lengthening of the word *autumn*) to indicate that it was a question, even though syntactically it was not phrased as such.

(1) B: Is that a good program?
S: Er ... it's quite interesting. (M1-130, 131)



Figure 1: An example of a QA pair (M3-30, 31). EI refers to exchange interval (see below). S: subject; I: interviewer.

CF pairs are exchanges in which a speaker confirms what the other speaker just said in the previous turn, as shown in (2). Usually the confirmation response is indicated by *yeah*, *right*, etc., but sometimes repetition would also be used. Repetition can be complete, as in (3), or partial, as in (4).

- (2) B: There aren't too many koala bearsS: Yeah. (M3-133, 134)
- (3) B: In <u>in Melbourne</u>. S: <u>In Melbourne</u>. (F1-140, 141)
- (4) B: That that's <u>sad</u>.
 S: Pretty uh <u>sad</u>. (M3-84, 85)

3.3. Measurements

Exchange interval (EI) was used in this study as a measure for exchange rhythm. It is defined as the latency between the end of one turn and the beginning of the next (see Figure 1). If EI is positive, it implies an unmarked-next-position onset. If EI is zero, it implies a latched onset. Finally, if EI is negative, it implies a terminal onset, as only transitional onsets were included in this study. All measurements were done using Praat (www.praat.org).

4. Results

4.1. Summary statistics

In total, there were 182 B \Rightarrow S exchanges. Female interviewees provided almost twice as many exchanges as male interviewees. Table 1 shows that QA was the most common, probably due to the nature of the corpus being in an interview format.

 Table 1: Summary statistics of exchange types. M:

 male; F: female.

	0A	CF	Others	Total
M1	28	2	1	31
M2	8	0	1	9
M3	15	6	1	22
F1	34	15	1	50
F2	22	10	0	13
F3	18	10	2	43
T-5	10	/	2	100
Total	136	40	6	182

Of the 182 exchanges, only 11 were terminal (6.04%; M = -92.43 ms, SE = 32.35), 5 were latched (2.75%), and the rest were unmarked-next-position onsets (91.21%; M = 442.93 ms, SE = 28.09). The mean overall EI was 398.41 ms (SE = 27.83), ranging from -370.25 ms to 2705.41 ms. Figure 2 shows the distribution of EIs, the central 50% of which were between 119.53 and 618.46 ms.



Figure 2: *EI distribution for* $B \Rightarrow S$ *exchanges*.

4.2. Analysis

A Gender (2) × Exchange type (QA, CF) two-way ANOVA showed that both of the main effects were significant [Gender: F(1, 172) = 5.92, p < .05, $\hat{\eta}^2 = .03$; Type: F(1, 172) = 21.97, p < .0001, $\hat{\eta}^2 = .11$]. No interaction effect was found (Figure 3). Post-hoc pairwise comparisons using Bonferroni's adjustments showed that female speakers had longer EIs than male speakers (p < .05), and EIs for QA pairs were longer than those for CF pairs (p < .0001).



Figure 3: EIs for QA and CF.

Since there is a significant gender difference regarding EI, two separate analyses were done to test individual differences, one on male, and the other on female, in order to partial out the gender effect. As seen in Table 1, since male speakers in general did not produce many CF cases, a one-way ANOVA on only QA pairs were done to test individual variations for males. No significant effect was found [F(2, 48) = 1.67, ns.]. A similar analyses was done on female speakers, but including both QA and CF pairs. Again, no significant effect was found [F(2, 114) = 1.61, ns.].

5. Discussion

EI seems to be distributed within a relatively small range. 50% of them occurred within a span of about 500 ms, implying that this is probably the preferred range to conduct a smooth conversation in Singapore English. Interestingly, the majority of transitional onsets in the corpus were unmarkednext-position (> 90%). In other words, pause at the exchange point is likely to be the default for Singapore English, which makes it closer to Midwestern and Californian than New York English in style. However, the predominance of unmarkednext-position onsets might also be due to the formal setting of the interview since the interviewer is a university lecturer, and the interviewees are all his students or of his students' age (*i.e.*, M2). It is possible that the interviewees were trying to show respect to the interviewer by being more attentive than involved. More studies would be needed to tease the two factors apart.

With regards to the effect of exchange type, we found that EIs for QA pairs were longer than those for CF pairs. This implies that cognitive loading is a deciding factor in determining the duration of EIs, contrary to Scollon's [2] claim. This finding is also interesting in that it creates a mismatch between adjacency and rhythmic structures, rendering a pair with tighter adjacency coupling (i.e., QA) to have longer pauses in-between than one with looser coupling (i.e., CF). Similar findings were also reported in Tannen [1]. In interchanges where the turn exchange styles of the participants are in conflict, interlocutors tend to attach more weight to dialect-specific exchange styles than to coupling tightness of adjacency pairs, resulting in uncooperative conversations. More studies would be needed to determine for certain whether coupling tightness has any role at all in governing turn exchanges.

We also found a gender effect regarding EI. Female speakers displayed EIs that were about twice as long as those by males, regardless of exchange types. This is especially true for CF pairs. This might have something to do with Asian societies being in general partial to males, and there are differential social expectations for different genders. For example, *smart* and *successful* are "good" traits for males but not for females, while *meek* and *gentle* are "good" traits for females but not for males. As a consequence, females might feel more insecure in social interactions [12], and might thus be more cautious in answering questions and giving confirmations than males, which results in longer pondering time. Males, on the other hand, were more likely to give answers immediately to show that they were intelligent and decisive, fulfilling social expectations.

Finally, unlike what Scollon [2] had claimed, we did not find any effect regarding individual variations. However, this does not imply individual styles have no place in exchange rhythm, given the fact that we only included about five minutes of recordings of relatively formal register (in school, with a lecturer, etc.) from each speaker in this study. More data are probably needed in order to clarify this.

6. Conclusion

This study looks at the exchange rhythm of Singapore English in a formal cross-dialectal interview setting. It seems that turn taking in this dialect is more like Midwestern and Californian English in that the majority of the transitional onsets cluster to a relatively small range, and are unmarked-next-position, implying that in conversation, Singaporeans value attentiveness more than involvement. Contrary to some previous studies, the exchange rhythm is reflective of cognitive loading, with exchange pairs requiring heavier cognitive processing showing longer EIs, even at the risk of disrupting adjacency structure. Gender also has an effect on EI, with females displaying intervals that are about twice as long as males, implying that the degree of social insecurity might also be a determining factor on exchange rhythm. However, no evidence of individual EI styles was found.

7. Acknowledgments

I would like to thank David Deterding for allowing me to have access to the NIESSE corpus and also for his useful comments on a previous version of the paper. I would also like to thank the anonymous reviewer from SP2006 for this paper for his valuable comments.

8. References

- D. Tannen, "New York Jewish Conversational Style," International Journal of the Sociology of Language, vol. 30, pp. 133-149, 1981.
- [2] R. Scollon, "The machine stops: Silence in the metaphor of malfunction.," in *Perspectives on silence* D. Tannen and M. Saville-Troike, Eds. Norwood, NJ Ablex, 1985 pp. 21-30
- [3] H. Sacks, E. A. Schegloff, and G. Jefferson, "A simplest systematics for the organization of turn-taking for conversation," *Language*, vol. 50, pp. 696-735, 1974.
- [4] J. K. Local, J. Kelly, and W. H. G. Wells, "Towards a phonology of conversation: Turn-taking in Tyneside English," *Linguistics*, vol. 22, pp. 411 - 437, 1986.
- [5] C. Ford and S. Thompson, "Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns," in *Interaction*

and Grammar, E. Ochs, E. Schegloff, and S. Thompson, Eds. Cambridge: Cambridge University Press, 1996, pp. 134-184.

- [6] B. Wells and S. Macfarlane, "Prosody as an interactional resource: Turn-projection and overlap," *Language and Speech*, vol. 41, pp. 265-294, 1998.
- [7] P. Auer, "On the prosody and syntax of turncontinuations," in *Prosody in Conversation: Interactional Studies*, E. Couper-Kuhlen and M. Selting, Eds. Cambridge: Cambridge University Press, 1996, pp. 11-56.
- [8] M. Selting, "On the interplay of syntax and prosody in the constitution of turn-constructional units and turns in conversation," *Interaction-based Studies of Language*. *Special issue of Pragmatics*, vol. 6, pp. 371-388, 1996.
- [9] J. Caspers, "Who's next? The melodic marking of question versus continuation in Dutch," *Language and Speech*, vol. 41, pp. 375-398, 1998.
- [10] G. Jefferson, "Two explorations of the organization of overlapping talk in conversation: Notes on some orderliness of overlap onset," *Tilburg Papers in Language and Literature*, vol. 28, 1983.
- [11] F. Goldman-Eisler, Psycholinguistics: Experiments in Spontaneous Speech. New York: Academic Press, 1968.
- [12] P. Eckert and S. McConnell-Ginet, "Communities of practice: where language, gender, and power all live," in *Locating Power: Proceedings of the Second Berkeley Women and Language Conference*, K. Hall, M. Bucholtz, and B. Moonwomon, Eds. Berkeley, CA: Berkeley University, 1992, pp. 89-99.
- [13] D. Deterding and E. L. Low, "The NIE corpus of spoken Singapore English (NIESSE)," SAAL Quarterly, vol. 56, pp. 2-5, 2001.