**Conversational alignment and conversational fluency**

Julie E Boland (University of Michigan)

jeboland@umich.edu

Fluent conversation is a marvel of multi-tasking within the language domain: the listener must simultaneously comprehend the speaker, predict a turn transition point, and plan a response, yet turn transition time for English dyads averages just over 200ms (Levinson & Torreira, 2015). Despite the apparent demands on working memory and selective attention, Pickering and Garrod (2004, 2013) suggest that dialogue is actually easy, because it is facilitated by a collection of automatic cognitive operations when interlocutors are well-aligned (i.e., using the same words, phrases, and structures to discuss the same topics). We test three claims from this account (i) conversation makes minimal demands on central resources, (ii) alignment is necessary for fluent conversation, and (iii) lexical repetition between speakers is an essential component of the alignment advantage. We found partial support for (i) and (ii), but no support for (iii).

In two experiments, participants casually conversed with an experimenter, while in some blocks simultaneously performing a working memory task at two levels of difficulty. Experiment 1 used a visuo-spatial n-back for the secondary task and Experiment 2 used a visual letter n-back. In both experiments, there were measurable, but quite small, effects of the n-back on conversational fluency variables, such as turn transition time (Fig 1), turn length, speech rate, and disfluencies. The strongest predictor of conversational fluency, by far, in both experiments was turn type (question, answer, agree, comment), perhaps because of differences in how quickly participants could begin planning their responses and the degree of planning required (Fig 2). The robust turn-type effect suggests that the cognitive operations supporting utterance planning are not completely automatic. Across the two experiments, almost all conversations were well-aligned, using Latent Semantic Similarity (LSS, Landauer and Dumais, 1997) and Language Style Matching (LSM, Ireland and Pennebaker, 2010) as estimates of alignment. Despite strong overall alignment, there were some hints that degree of alignment influenced conversational fluency: LSS was a significant predictor of speech rate in Experiments 1 and 2 and a marginal predictor of transition time in Experiment 2; LSM was a marginal predictor of turn length in Experiment 1.

Because it was difficult to manipulate alignment in spontaneous conversation, Experiment 3 used a picture description paradigm to test the effects of two aspects of alignment (shared topic and shared vocabulary). Participants looked at a line drawing of a complex scene while listening to an auditory sentence, then described the image. In the Match condition, the auditory sentence was about the current image; in the Mismatch condition, an auditory sentence for a different image was substituted. A recognition probe encouraged attention to the auditory sentence. With training, participants learned to time their utterance to begin near the end of the auditory stimulus sentence, mimicking a conversational turn transition (descriptive statistics in Table 1). As expected, fluency was higher in the Match condition, with shorter transition times and more succinct descriptions than in the Mismatch condition. In contrast, the more content words the participant repeated from the auditory stimulus sentence, the longer the transition time and the more wordy the image description (see Fig 3). Thus, a shared topic increased fluency, as predicted. But the topic alignment advantage did not come from the most obvious source--direct lexical repetition of words in the auditory stimulus. Rather, the Match advantage in this paradigm must be due to more subtle phenomena, such as more accurate prediction of the end of the auditory stimulus, greater use of anaphoric expressions, and semantic priming.

References

Ireland, M. E. & Pennebacker, J. W. (2010). Language style matching in writing. *Jo Pers & Social Psych*, 99, 549-571.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory. *Psych Review*, 104, 211-240.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load Theory of Selective Attention and Cognitive Control. *JEP: General*, 133(3), 339-354.

Levinson, S. & Torreira, F. (2015). Time in turn-taking and its implications for processing models of language. *Front. Psychol.* 6:371.
Pickering, M. J. & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *BBS*, *27*, 169-190.
Pickering, M. J. & Garrod, S. (2013). An integrated theory of language production and comprehension. *BBS*, 36, 329-347.

**Experiment 2 Results**

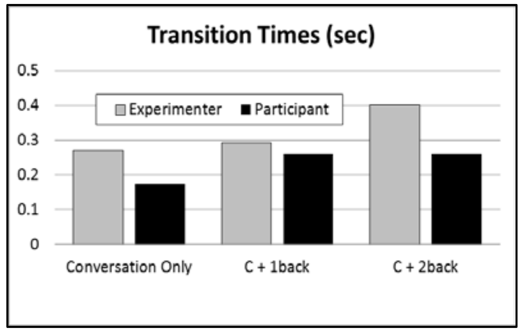


**Figure 1.** Transition times for spontaneous conversation between participant & experimenter: Participant times were slower in difficult 2back than in Conversation Only.

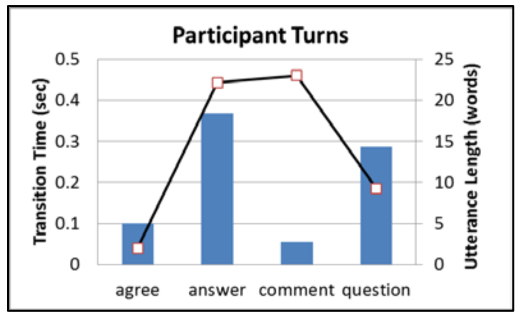


**Figure 2.** Type of utterance was a strong predictor of transition time (bars) and utterance length (line), but they did not co-vary. E.g., comments tended to be long utterances, but had fast transition times.

**Experiment 3 Results**

**Table 1.** Grand Means (standard error) for transition time between auditory stimulus offset and participant utterance onset, number of words in participant's utterance, participant's speech rate, percent of utterances containing disfluencies, and number of words in the participant's utterance repeated from auditory stimulus from the same trial.

| | **Transition Time** | **# of Words** | **Speech Rate** | **Disfluent %** | **#Repeated** |
|---|---|---|---|---|---|
| Match | 515 ms (22) | 10.01 (.17) | 3.23 w/s (.04) | 33 (2) | 1.42 (.05) |
| Mismatch | 575 ms (25) | 9.56 (.16) | 3.15 w/s (.04) | 33 (2) | 0.12 (.02) |

**Figure 3.** When controlling for number of repeated words, Match utterances had shorter transition times and fewer words. The more repeated words, the longer transition time and utterance length.