

Taking turns with the unpredictable

Laurel Brehm

laurel.brehm@mpi.nl

The mean turn gap between speakers is about 200 ms in several corpora (Stivers et al, 2009; Levinson & Torreira, 2015), with a large range around this estimate. Existing work suggests that variation in turn gaps comes from difficulty in predicting the prior turn's timing: as long as there is enough information available to plan a response, having a good estimate of when the prior turn will end decreases turn gaps more than having a clear prediction of what the prior turn will say (e.g. Brehm & Meyer, 2021; Corps, Crossley, Gambi, & Pickering, 2018; Magyari & de Ruiter, 2012). The current work pits the role of prompt content against prompt timing in coordination by asking speakers to time their speech onset to the offset of items varied in name agreement (low NA: *couch/sofa* versus high NA: *tiger*; Vitkovitch & Tyrell, 1995), which makes the prompt label more or less predictable in length and content.

A set of Dutch stimuli was developed from existing norms (Decuyper, Brysbaert, & Meyer, 2021). Each picture had three labels: a *modal* name (the most common response), and two lower NA names. One was a *same-length* name (same number of syllables as the modal name), and the other a *different-length* name (at least one more or one less syllable than the modal name); see Table 1. We used these to test the role of name agreement in name-picture verification (Experiment 1) and the role of name agreement and word length in a coordinated production task (Experiment 2). The degree to which both experiments match indexes the role of prompt content in coordinating speech timing; differences show additional influences of prompt length in turn taking.

Experiment 1 (N=120 participants) was a name-picture verification study (e.g. Stadthagen-Gonzalez et al., 2009) with 30 critical items and 168 high NA fillers. In this paradigm, a picture was presented, followed by a recorded name at an ISI reflecting the time needed to produce the picture name (M = 960 ms). Participants had to judge whether the picture and recording matched. Each participant received half matching and half mismatching trials and only one version of each critical item. Judgment accuracy was high for all fillers (match = 98%; mismatch = 99%), and high for 24/30 of the critical items (by-item range = 75%-100%). RT on correct match trials correlated weakly but reliably with name agreement ($r = -0.092$; see Figure 1), such that as name agreement increased, response time decreased. This suggests more processing time is needed to comprehend names of low NA pictures.

Experiment 2, ongoing, (target N=60 participants, 10 collected so far) was a production study designed to test whether dispreferred names disrupted turn taking in a simple coordinated speech production task. Participants listened to recordings of the 24 best critical items from Experiment 1 and 156 high NA fillers and were tasked with producing the name of their own high NA picture immediately after the recording ended (see Figure 2). The interval between recording offset and speaker production onset (turn gap) was measured. Once data collection is complete, turn gaps will be calculated for each condition in Table 1: the prediction is that if speakers are predicting prior turn offset from the picture's modal name, then different-length long items will elicit interruptions, different-length short items will elicit long gaps, and same-length items will elicit equivalent gaps to the modal items. In contrast, if speakers are predicting prior turn offset from predicted word content, any non-modal label should increase turn gaps. These patterns will be tested statistically using a model fitting procedure: we will hold out the non-modal items while fitting a model to the modal critical items and fillers with only the predictor of name agreement, and measure the accuracy of model predictions for the held-out conditions. This will test whether coordination is reliably different for various types of dispreferred names compared to the modal baseline, as would be expected if participants use expectations about word length in conversational turn-taking more than word content. Condition differences and the overall effect of name agreement in the resulting statistical model will then inform us about what type of predictions about word content are used to time conversational turn-taking.

Table 1: Sample critical items. NA = name agreement.



	Modal Name		Same Length Name		Different Length Name	
	Label	NA	Label	NA	Label	NA
	<i>theedoek</i>	40.00%	<i>handdoek</i>	31.43%	(short) <i>doek</i>	8.57%
	<i>poeder</i>	29.03%	<i>make-up</i>	22.58%	(long) <i>poederdoos</i>	9.68%

Figure 1: Experiment 1 results. Correct match trial RTs by NA and label condition.

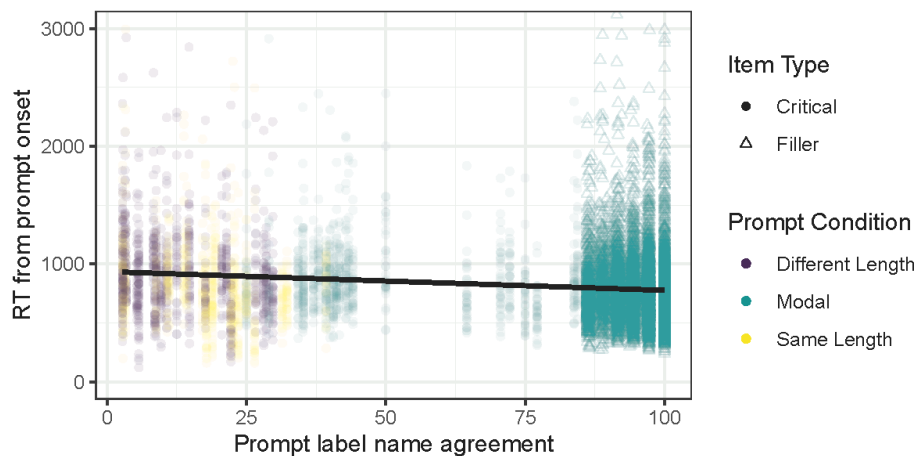
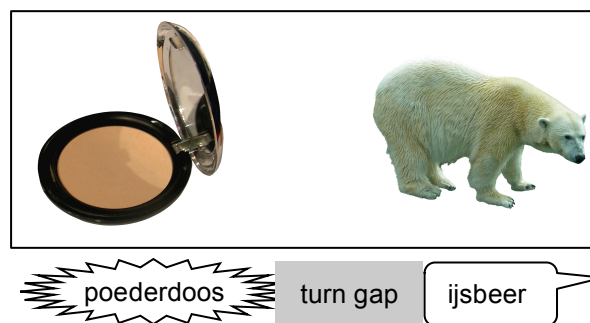


Figure 2: Experiment 2 paradigm. Recording plays first word; participant produces second word, and interval between words (turn gap) is measured.



References:

- Brehm, L., & Meyer, A.S. (2021). Planning when to say: Dissociating cue use in utterance initiation using cross-validation. *JEP:G*
- Corps, R. E., Crossley, A., Gambi, C., & Pickering, M. J. (2018). Early preparation during turn-taking: Listeners use content predictions to determine what to say but not when to say it. *Cognition*, 175, 77-95.
- Decyper, C., Brysbaert, M., & Meyer, A.S., (accessed 2021). Dutch norms for BOSS pictures. <https://osf.io/kwu87/>
- Magyari, L., & de Ruiter, J. P. (2012). Prediction of turn-ends based on anticipation of upcoming words. *Front. Psychol* 3, 376.
- Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Front. Psychol.*, 6, 731.
- Vitkovitch, M., & Tyrrell, L. (1995). Sources of disagreement in object naming. *QJEP Section A*, 48(4), 822-848.
- Stadthagen-Gonzalez, H., Damian, M. F., Pérez, M. A., Bowers, J. S., & Marín, J. (2009). Name-picture verification as a control measure for object naming: A task analysis and norms for a large set of pictures. *QJEP*, 62(8), 1581-1597.
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., ... & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *PNAS*, 106(26), 10587-10592.