How do speakers coordinate their utterances?
Evidence for prediction of another's utterances in a joint language task

Chiara Gambi[1], Uschi Cop[2], and Martin J. Pickering[1]
[1]University of Edinburgh, [2]University of Ghent
contact: C.Gambi@sms.ed.ac.uk

Research into how two people coordinate to achieve a common goal revealed that co-actors model each other's actions and integrate predictions about their partner's actions into their own action planning (Sebanz et al., 2006). In conversation, interlocutors coordinate their utterances by implementing timely and smooth turn-taking. Addressees anticipate the end of the speaker's turn in order to get their contribution ready in time (e.g., De Ruiter et al., 2006).

In three experiments (32 participants each, tested in pairs), we tested whether people predict their partner's utterances in a joint language task. We adapted a task developed for the experimental study of self-repairs (Hartsuiker et al., 2008). Two participants (A and B) took turns in naming pictures; on a small percentage of trials (change trials; see Figure 1) an initial picture (e.g., of an apple) changed into a target picture (e.g., of a chair). The task was to stop naming the initial picture as soon as the change was detected and immediately resume with the target name (e.g., (*Ap-Chair*)$_A$; Exp. 1), stop naming the initial while the partner resumed with the target name (e.g., (*Ap-*)$_A$, (*Chair*)$_B$; Exp. 2), or stop naming with no resumption (e.g., (*Ap-*)$_A$; Exp. 3). As in Hartsuiker et al., initial names were completed more often in Exp. 1 than in Exp. 3. This is consistent with a decision to avoid the cost of stopping speech and instead invest resources in planning the target name, while carrying on with the articulation of the initial name (Tydgat et al., 2011).
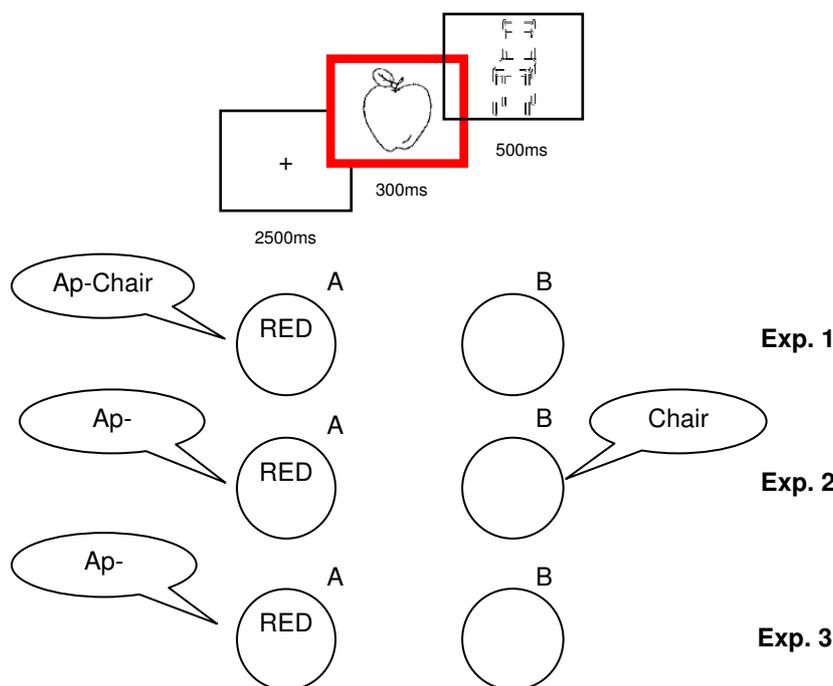

Figure 1. Summary of the procedure for a change trial in Experiments 1 to 3. Participant A named pictures surrounded by a red frame.

Additionally, we manipulated target planning difficulty in all three experiments. Target pictures were visually-degraded (see Figure 1) or intact. The duration of the initial name was longer before degraded targets in Exp. 1 (but not in Exp. 3). Planning the

resumption interferes with the process of stopping speech in a way that suggests they share a limited pool of resources (Hartsuiker et al., 2008).

If participants in Exp. 2 predict that their partner will resume, they might follow the same strategy as participants in Exp. 1 and withdraw resources from stopping, even though they need not invest resources in target planning. If so, we expected initial names to be completed more often in Exp. 2 than in Exp. 3. However, if they do not plan the resumption in their own production system (i.e., they do not retrieve the semantics, syntax and phonology of the target name), there should be no effect of target planning difficulty on initial duration in Exp. 2.

Both predictions were confirmed. Participants completed more often in Exp.2 than in Exp. 3 (63.2 % vs. 46.5 %, see Table 1; β=-0.99, z=-2.16, p<.05). Initial duration was not affected by target degradation (degraded: 362 vs. intact: 356ms, n.s.). We conclude that participants in Exp. 1 and 2 were anticipating that a word would be uttered, either by themselves or by their partner. We argue that they computed such predictions using a forward model (Pickering & Garrod, 2007),whereas they appeared not to plan their partner's utterance (unlike their own utterance) in their own production system.

	Initial Response Type			
Experiment	Completed	Interrupted	Skipped	**Tot.**
Exp. 1 (N=963)	83.4	6.5	10.1	100
Exp. 2 (N= 962)	63.2	22.0	14.8	100
Exp. 3 (N=914)	46.5	25.8	27.7	100
Mean	64.7	18.0	17.3	100

Table 1. Percentages of Initial Response Type by Experiment (Completed: *Apple*, Interrupted = *Ap-*, Skipped = Ø); N is the total number of responses in the experiment.

References:
De Ruiter, J.P., Mitterer, H., & Enfield, N.J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language, 82*, 515-535.
Hartsuiker, R.J., Catchpole, C.M., de Jong, N.H., & Pickering, M.J. (2008). Concurrent processing of words and their replacements during speech. *Cognition, 108,* 601-607.
Pickering, M.J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11,* 105-110.
Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences, 10*, 70-76.
Tydgat, I., Stevens, M., Harstuiker, R.J., & Pickering M.J. (2011) Deciding where to stop speaking. *Journal of Memory and Language, 64,* 359-380.