

## **When can you finish someone else's...? Influencing the production of compound contributions in dialogue**

Christine Howes,<sup>1</sup> Patrick G.T. Healey,<sup>1</sup> Ruth Kempson,<sup>2</sup> and Eleni Gregoromichelaki<sup>2</sup>

<sup>1</sup>Queen Mary University of London, London, UK

<sup>2</sup>King's College London, London, UK

c.howes@qmul.ac.uk

Linguistic analyses of both intra-sentential grammar-internal and cross-sentential discourse patterns have until recently presumed that the dynamics of communication is solely the provenance of semantics/pragmatics, with syntax precluding in principle all reference to aspects of performance such as real-time incrementality (Jackendoff 2002). Yet it has long been known that parsing is incremental (see e.g. Crocker et al 2000), with more recent evidence demonstrating that production is equally incremental (see e.g. Konopka 2012). Within ongoing debates in these areas, there are two related concerns. To what extent are parsing/production decisions context-driven; or, conversely, expectation-driven? We suggest that critical evidence for these debates can be drawn from a study of compound contributions.

Compound contributions (CCs) – dialogue contributions that continue or complete an earlier contribution (see e.g. 1) – make up as much as 20% of spoken corpora (Purver et al, 2009), and appear to allow switch of speaker/hearer roles at arbitrary points in a sequence.

- 1) Daughter: Oh here dad, a good way to get those corners out  
Dad: is to stick yer finger inside.  
Daughter: well, that's one way. [from Lerner (1991)]

This phenomenon has been recently studied within the *Dynamic Syntax* framework (DS: Purver et al, 2010). With “syntax” defined in DS to provide constraints inducing incremental build up of semantic representation, it has been shown that successful modelling of CCs can be achieved across the broad array of construction-types and role-switches (see Gregoromichelaki et al, 2011). In DS, projection of such semantically-grounded structures as “syntax” is, by definition, expectation-driven, in that overall projection of structure is top-down and goal-directed. Such expectations as they unfold incrementally become satisfied by lexically-driven, hence broadly bottom-up, input reflecting time-linear processing dynamics. With speaker and hearer processing coupled via use of such procedures, this account correctly predicts that CCs can ensue at any point in an utterance sequence, with apparent fluency in the shift at each such point.

However, the further question that these data raise is what determines the use of CCs by participants; and are CCs initiated due to the same factors as cross-sentential switches, or, are they, instead, driven by locally-defined constraints? We report an experiment which indicates that predictability at the sub-sentential level is not sufficient to predict when interlocutors are likely to continue another's incomplete utterance, but that the production of such continuations involves an interaction of local predictability and some more global notion of predictability that could be based on common context (e.g. “topical coherence”, Schegloff 1990) or QUD (Ginzburg 2012).

This experiment used the Dialogue Experimental Toolkit (DiET) chat tool (Healey et al, 2003), a text-based chat interface through which interventions can be introduced into a dialogue in real time. To investigate the influence of different levels of predictability on how people respond to unfinished turns and their likelihood of producing a continuation, a number of genuine single contributions in dyadic conversations were artificially split into two parts. The first part was transmitted to the other participant as typed, with the turn truncated according to syntactic and lexical predictability. Since predictability depends on dialogue context and topic, values were calculated from a corpus of prior dialogues (53663 word tokens) collected using the same tool (i.e. text-based chat) and domain – the balloon task – an ethical dilemma in which participants must decide who to throw out of a hot-air balloon

that is losing height. Two entropy measures were calculated: part-of-speech entropy, to capture the “syntactic” predictability of one part-of-speech (POS) following another; and lexical entropy, to capture the predictability of a particular lexical item following a specific POS. To illustrate the difference: although determiners are predictably followed by nouns, there are lots of different nouns: determiners therefore have a low POS entropy (i.e. highly predictable next POS), and a high lexical entropy. A POS-tagger (Toutanova et al., 2003) analysed the typed turn in real time and triggered an intervention based on these entropy values, producing a range of truncation points with high, medium and low POS entropy, and, independently, high, medium and low lexical entropy.

The truncated turn was followed by a delay, during which the other person could respond if they wished. Any response was not relayed to the original sender, who was therefore unaware that their turn had been tampered with, and the rest of the original (interrupted) contribution was subsequently transmitted. Measures of local expectations were the online calculations of the syntactic and lexical predictability of the next word at each truncation point. Each truncated turn was classified as either contributing to an ongoing topic of discussion (loosely defined as something the dyad had already mentioned in a prior conversational turn), or introducing a new topic, as a measure of global (cross-sentential) predictability. Results show that participants were more likely to construct their response as a CC if it was about the current topic than if it was about something else (topic 59/121, 49% vs. off-topic 13/50, 26%;  $\chi^2(1) = 7.519, p = 0.006$ ). Moreover, a GEE model (Topic  $\times$  Lexical predictability  $\times$  POS predictability  $\text{Wald}\chi^2 = 8.635, p = 0.003$ , see figure 1) showed that if the next lexical item is unpredictable, participants were more likely to construct their response as a CC if a) the syntactic category is predictable and b) they were talking about some topic which was already being discussed, hence contextually salient. In other words, although people can continue an incomplete sentence at any syntactic point, they are unlikely to do so unless the possible continuations are sufficiently constrained by the unfolding context. Conversely, when an incomplete utterance does constrain the range of possible continuations, participants are able to produce a suitable continuation by utilising the specific grammar-defined constraints provided at that point. These results indicate that both local (e.g. syntactic) and global (e.g. “topicality”) sources of discourse expectations are available to interlocutors in dialogue at a sub-sentential level.

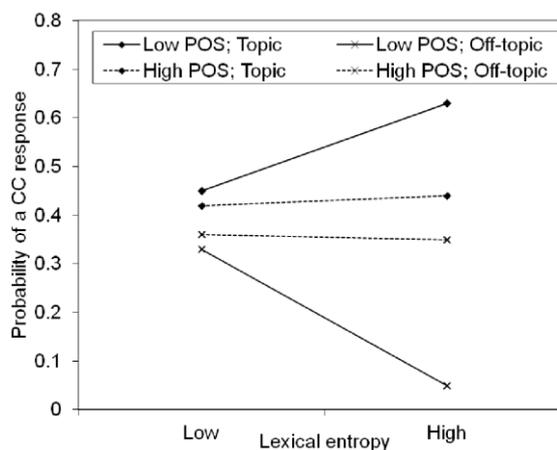


Figure 1: Marginal means of probability of a CC response

This interweaving of intra- and extra- grammatical mechanisms can be captured by encapsulated sentence-based grammar formalisms only if each such fragment is taken as an elliptical sentence string requiring completion via appropriate pragmatic inference steps (e.g. Stainton 2006). This leaves wide open the challenge of how to constrain such extra-grammatical steps so that their output will match the input to whatever further structure building is needed to ensure a wellformed result over all. Potential “constructional” resolutions (e.g. Ginzburg 2012) etc. will fail as CCs are available at any syntactic juncture, a

fact threatening to empty the concept of grammar-defined wellformedness of all content. However, DS, unlike other grammar formalisms, is a leading contender for modelling the seamlessness of such context/structure-building interdependence, because “syntactic” constraints are defined within a non-encapsulated domain-general system of structure-building licensing incremental interweavings of grammar-induced and context-induced actions at any point in the construction process. The more general conclusion is that expectations and their fulfilment are not some peripheral extra-grammatical phenomenon, but constitute the core of our linguistic-competence (see also Pickering and Garrod in press).

## References

- Crocker, M., Pickering, M. And Clifton, C. eds. 2000. *Architectures and Mechanisms for Language Processing*. Cambridge University Press.
- Ginzburg, J. 2012. *The Interactive Stance: Meaning for Conversation*. Oxford University Press.
- Gregoromichelaki, E. Kempson, R. Purver, M. Mills, G.J. Cann, R. Meyer-Viol, W. Healey, P. 2011. Incrementality and intention-recognition in utterance processing. *Dialogue & Discourse*, 2(1), 199-233.
- Healey, P. G. T., Purver, M., King, J., Ginzburg, J., and Mills, G. (2003) Experimenting with clarification in dialogue. In *Proceedings of 25th Cognitive Science Society Meeting*.
- Howes, C. Healey, P. Purver, M. Eshghi, A. 2012 Finishing each other's.... responding to incomplete contributions in dialogue. *Proceedings of 34th Cognitive Science Society Meeting*, 479-85.
- Jackendoff, R. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. OUP.
- Konopka, A. E. 2012. Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language*, 66, 143-162.
- Lerner, G. H. 1991. On the syntax of sentences-in-progress. *Language in Society* 20(3), 441-458.
- Pickering, M. & Garrod, S. (in press), 'An integrated theory of language production and comprehension'. *Behavioral and Brain Sciences* to appear .
- Purver, M. Howes, C. Gregoromichelaki, E. Healey, P. 2009. Split utterances in dialogue: A corpus study. In *Proceedings of the 10th Annual SIGDIAL Meeting on Discourse and Dialogue*, 262-271. Association for Computational Linguistics, London, UK.
- Purver, M. Gregoromichelaki, E. Meyer-Viol, W. Cann, R. 2010. Splitting the 'I's and crossing the 'You's: context, speech acts and grammar. In *Aspects of Semantics and Pragmatics of Dialogue. SemDial 2010, 14th Workshop on Semantics and Pragmatics of Dialogue*, 43-50.
- Schegloff, E. A. 1990. On the Organization of Sequences as a Source of “Coherence” in Talk-in-Interaction. In B. Dorval (ed.) *Conversational Organization and its Development*. Norwood, NJ: Ablex Publishing Co., pp. 51–77.
- Stainton, R. 2006. *Words and Thoughts* Oxford University Press.
- Toutanova, K., Klein, D., Manning, C., and Singer, Y. 2003. Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proceedings of HLT-NAACL*, pages 252–259.