Grammar Formalisms and Explanations of Dialogue

Ruth Kempson, Ronnie Cann, Matthew Purver
King's College London, Edinburgh, CSLI Stanford
kempson@dcs.kcl.ac.uk,r.cann@ed.ac.uk,mpurver@stanford.edu

1 Dialogue: The Challenge of Split Utterances

Contrary to what for years has been a widespread belief that conversational data are riddled with mistakes, false starts, and simply ungrammatical data, Pickering and Garrod (2004) extended a challenge primarily to psycholinguists, but also to theoretical linguists that their models should be evaluatable by how well they provide a basis for characterizing the patterns systematically displayed in conversational dialogue, this being the core language data. In this paper we argue that by adopting a Dynamic Syntax perspective, with the time-linear dynamics of building up structured semantic representations relative to context built in to the architecture of the grammar formalism, these patterns are directly predicted, while still retaining the assumption of grammar as a system which underpins language use but does not fully dictate it, i.e. without writing conversational rules or a grammar of conversation.

First, dialogue is replete with pronouns and ellipsis (where the context is essential to interpreting the pronoun or "completing" the uttered fragment), so that commonality of content for a subpart of what interlocutors say to each other is by definition shared. Secondly, even when speakers and hearers do not choose forms which literally pick up interpretation from context, nevertheless they systematically re-use the tools the other person in the dialogue has used wherever possible, so that interlocutors will replicate the type of structure with which they are addressed, whether a subcategorisation choice or more globally, and rather than shifting to an alternative mode of presentation, they will repeat words, and in repeating those words, they will retain the same interpretation assumed by their interlocutor (so-called alignment phenomena) (1):

(1) A: What did you buy Eliot?
B: Some lego.
A: I bought him some playdoh.

In addition, interlocutors regularly finish each other's utterances, a hearer shifting into being a speaker with respect to a single structure, and v.versa, apparently effortlessly,

a phenomenon like ellipsis and pronoun control (except 1st and 2nd pers pronouns), equally controlled early in language acquisition:

(2) A: Where shall we go for
B: Christmas? Nowhere.
C: Unless Granny is expecting us.

(3) A: Old Macdonald had a farm E-i-e-i-o. And on that farm he had a
B: Pig.
A: And the pig goes
B: oink oink.

One might argue that such phenomena display nothing more than pragmatic tendencies, driven by some form of Gricean or Relevance-theoretic pressure (eg Sperber and Wilson, 1995), that is irrelevant to grammar design, given the feeding relation between grammar formalism and pragmatics. However, this leaves entirely unexplained their systematicity and early emergence in language acquisition. Such data are extremely hard to explain given conventional assumptions about the relation between grammar formalism and language performance. Despite attempts to reduce the diversity (eg Kamp, 1984), both anaphora and ellipsis phenomena remain recalcitrant, heterogeneous phenomena that have to be split into grammar-internal and discourse distributions (eg Hankamer and Sag, 1976). Yet this division is contrary to the evidence of split utterances, and the fact that fragments may contain anaphoric devices which require interpretation as though they were part of a larger whole. Indeed any form of dependence can be split across the context structure within which some fragment is to be interpreted and the fragment itself:

(4) Who did every girl worry about? Herself/her mother

(5) A: Do you know whether John has handed in
B: his term-paper?
A: or even any problem-sets?

Even the very existence of switch-utterances is problematic given the assumption that parsing and production are separate applications of the use-neutral grammar formalism, for the fluency with which they take place is contra-predicted. The challenge is to express such regularities without enforcing a shift into articulating a dialogue grammar as though conversational interaction was rule-governed, a challenge which we shall argue is met by Dynamic Syntax (DS Kempson et al., 2001; Cann et al., 2005).

2 DS: Context-dependent Structure-building

In DS, syntax is defined as progressive building of tree representations of content relative to context and following the dynamics of parsing. The outcome from any one

such sequence of partial trees is a complete tree with topnode decorated with a propositional formula, each dominated node decorated with a sub-term. All syntactic/lexical processes are defined as contributing to monotonic tree growth across partial trees to such a complete tree, from an initial state which contains a single-node tree decorated solely with a requirement for a propositional value. As a lexical example of update actions encoding underspecified values, anaphoric expressions project an underspecified formula value necessitating update (eg pronouns, auxiliaries). Long-distance dependency illustrates structural underspecification, inducing introduction from some initial state of a node that dominated by the top proposition-requiring node without at that point further specification of that relation. In all cases of underspecification, update must be provided either from the construction process itself (6)-(7)), or from the context (8)-(10), with (10) an unfixed tree relation updated indexically :

(6) It is likely that I am wrong.

(7) John, Mary says Sue dislikes.

(8) John came in and he fainted.

(9) Mary, I like; and Harry too.

(10) Bacon and eggs, please.

Context is defined as a record of such structures and the actions used in building them; and production is defined to make use of the same system of structure-building relative to context, this being no more than an alternative application of the grammar formalism. Indeed, the structure-building process of production differs only from parsing in having in addition a so-called 'goal' tree against which all putative update transitions have to be checked as to whether each subsumes the intended goal tree.

2.1 Ellipsis

With context defined as a record of both structures and procedures used in building up such structures, the divergent ellipsis patterns can all be seen as constituting different ways in which construal can be dictated by context (either re-using context-recorded content (strict VP ellipsis construal, or re-using structure (fragment answers, where the parsing process develops the very structure initiated in the context) or re-using context-recorded actions (stripping, sloppy VP ellipsis). Choice of interpretation for anaphoric and elliptical expressions is triggered by the presence at the interpretation site of a metavariable which is defined as requiring a suitable type of value.

2.2 Coordination of Parsing/Production and Switch Utterances

Switch-utterance phenomena are directly predicted, since the production mechanism in following the dynamics imposed by the grammar formalism follows the same incre-

mental steps of tree construction as the parser: both parser and producer build the same (partial) semantic representation. Like parsing, production is context-dependent, licensing generation success with actions contextually stored, enabling successful communication without needing to repeat an explicit sentential string. The motivation for use of context in production is very strong, since it enables bypass of the otherwise considerable task of incremental search in the lexicon for appropriate words.

Given the account of ellipsis, the parse process can be initiated from any partial tree: in particular interpretation for fragment answers to questions (1) take the structure constructed from the question as the point of departure for their process of construal. In following the same dynamics, the generation process may equally take any partial tree as its starting point for the analogous generation of such fragments. And with such allowance for either parse process or generation process to start from any partial tree, switch between parse and production modes will be licensed at any arbitrary transition in the construction process; and switch for an individual from some parsing activity to production and vice-versa is predicted to be effortless, and very free. Notice that, as we expect, it is not restricted to major constituent edges, as illustrated by (2)-(3).

2.3 Context-dependent Wellformedness

Given a construction process shared by both parsing and production mechanisms, we define a concept of context-dependent wellformedness. Fragments, like other strings, are wellformed iff they lead to a complete representation of some propositional content (predicting their wellformednedss with respect to restricted contexts):

(11) A string ϕ is well-formed iff an utterance of ϕ is well-formed in a context $\mathfrak{C}$:

$$\exists \mathfrak{C}[P_0 \quad \overrightarrow{A_\phi, \mathfrak{C}} \quad \{\ldots, \langle T_\phi, \phi, A_\phi \rangle, \ldots\}]$$

where (as above) $P_0 = \{\langle T_0, \emptyset, \emptyset \rangle\}$ is the standard initial state (a single-node tree with only a propositional type-requirement, a null sequence of words, and a null sequence of actions); A_ϕ is the set of lexical, computational and pragmatic actions used in parsing ϕ on a strictly time-linear basis; and T_ϕ is complete (has no outstanding requirements).

This definition distinguishes the first part of a split utterance from the subsequent fragment. A fragment is well-formed as long as it leads to a complete tree (with no outstanding requirements) of propositional type. Hence the fragment replies (1)-(5) are all wellformed, relative to restricted contexts, but the interrupted utterance is not. Notice that the context licensing such fragments may be extremely restricted, as in (3).

Alignment phenomena (1) can be explained in terms similar to those underpinning ellipsis construal - minimizing production costs with actions from immediate context rather than a full-lexicon search. Parallelism at morphological lexical syntactic and semantic levels can all be explained as different aspects of re-use of context. Such

alignment across strings in context differs from ellipsis in that words are re-generated; but nevertheless the minimisation of production costs remains the operative constraint since no lexicon search is needed if words and actions are recovered from context.

3 A Grammar of Dialogue?

It might seem from this shift to characterising context-dependent concepts of well-formedness which allow a natural characterisation of split utterances, that we are advocating a position in which a grammar articulates what constitutes a well-formed dialogue. However, we would argue, to the contrary, there is no such shift. The difference between ellipsis and pronouns on the one hand, and alignment patterning on the other, is that the former involve use of context only as licensed by some specific trigger given by the linguistic input, hence require a grammar-internal specification of that input, whereas alignment is a fully pragmatic phenomenon arising solely from relevance constraints. In adopting a grammar formalism with the built-in dynamics of a parsing, we obtain the advantage of being able to explain the systematicity of patterns displayed in conversational dialogue without thereby being forced into the assumption of a grammar encoding conversational rules. With the Pickering and Garrod challenge arguably being met by DS, this debate opens up a broader swathe of evidence for evaluating grammar formalisms, since all conversational fragments and switch-utterance phenomena now fall within the remit we expect grammar formalisms to reflect.

References

Cann, R., R. Kempson, and L. Marten (2005). The Dynamics of Language. Number 35 in Syntax and Semantics. Elsevier, Oxford.

Hankamer, J. and I. Sag (1976). Deep and surface anaphora. Linguistic Inquiry, **7**:391–428.

Kamp, H. (1984). A theory of meaning and semantic representation. In T. Janssen and M. Stokhof, eds., Truth, Interpretation, and Information, pp. 1–34. Foris, Dordrecht.

Kempson, R., W. Meyer-Viol, and D. Gabbay (2001). Dynamic Syntax: The Flow of Language Understanding. Blackwell.

Pickering, M. and S. Garrod (2004). Towards a mechanistic account of dialogue. Behavioral and Brain Sciences, **27**:169–226.

Sperber, D. and D. Wilson (1995). Relevance: Cognition and Communication (2nd editn). Blackwell.