Splitting the *"I"*s and Crossing the *"You"*s: Context, Speech Acts and Grammar

Matthew Purver
Department of Computer Science,
Queen Mary University of London
m.purver@qmul.ac.uk

Eleni Gregoromichelaki, Wilfried Meyer-Viol
Department of Philosophy, King's College London
eleni.gregor@kcl.ac.uk
wilfried.meyer-viol@kcl.ac.uk

Ronnie Cann
Linguistics and English Language, University of Edinburgh
r.cann@ed.ac.uk

Abstract

The occurrence of split utterances (SUs) in dialogue raises many puzzles for grammar formalisms, from formal to pragmatic and even philosophical issues. This paper presents an account of some of the formal details that grammars need to incorporate in order to accommodate them. Using *Dynamic Syntax* (DS), we illustrate how by incorporating essential interaction with the context into the grammar itself, we can deal with speaker change in SUs: not only its effects on indexicals *I* and *you*, but also the multiple illocutionary forces that can arise. We also introduce a *Split Turn Taking Puzzle* (STTP) showing that the current speaker and the agent of the resulting speech act are not necessarily the same.

1 Introduction

Split utterances (SUs) – utterances split across more than one speaker or dialogue contribution – are common in spontaneous conversation and provide an important source of data that can be used to test the adequacy of linguistic theories (Purver et al., 2009). Previous work has suggested that Dynamic Syntax (DS) (Kempson et al., 2001) is well placed to analyse these phenomena as it is strictly word-by-word incremental, allowing an account of speaker changes at any point (Purver et al., 2006) and interruptive phenomena such as mid-sentence clarification sequences (Gargett et al., 2009). However, other less incremental grammar formalisms have also been applied to particular kinds of SUs: Poesio and Rieser (2010) use LTAG[1] in an analysis of *collaborative completions*. But given that such accounts employ grammars which license *strings of words*, a direct account for SUs is prevented, as the reference and binding of indexical speaker/addressee pronouns changes as the speaker transition occurs:

(1) A: Did you give me back
B: your penknife? It's on the table.

(2) A: I heard a shout. Did you
B: Burn myself? No, luckily.

A grammar must rule out the sentence *did you burn myself?* as ungrammatical if spoken by one single speaker, but allow it as grammatical if the identity of the speaker changes as in (2) – this will be hard for a string-based account. In (1), *you* and *your* must be able to take different referents due to the speaker change between them. In contrast, as DS defines grammatical constraints in terms of the incremental construction of semantic content (rather than through licensing strings via an independent layer of syntax over strings), we show that such examples are not problematic given an independently motivated definition of the lexical entries for indexicals.

SUs can also perform diverse dialogue functions, with the speech acts associated with the individual speakers' contributions often being different – see (Purver et al., 2009). In (1)-(2), B's continuations seem to function as clarifications of A's intended queries. Others have pointed out that continuations can function as e.g. adjuncts (Fernández and Ginzburg, 2002) or clarification requests (Purver et al., 2003); and Poesio and Rieser (2010) show how a completion (in their terms) can get its function in the dialogue (in their case, to act as a collaborative contribution to a plan). A full account of SUs therefore requires some representation of such dialogue function information in the model of context that guides dialogue interpretation and production.

[1] Unlike DS, LTAG must be supplemented by a parsing and/or generation model (a set of defeasible inference rules for Poesio and Rieser) to derive the incrementality required.

DS, however, currently incorporates no notion of illocutionary force or dialogue act type, as it is assumed that derivation of such information is not linguistically determined. In the case of SUs, it has been assumed that the *grammar* itself provides adequate means of continuing/taking over somebody else's utterance, and that this does not *necessarily* involve strategic reflection or fully-formed intentions as to what function the utterance should perform: this provides the possibility for speakers to 'blurt out' utterances without necessarily having any specific plans/intentions in mind, and for hearers to respond without reflection as to the speaker's plan. But, as pointed out in (Kempson et al., 2007; Gargett et al., 2009), this is not an in-principle objection to the specification of speech act information as part of the representation derived by the parse of an utterance, as DS provides mechanisms for allowing the inclusion of optional inferred information. We present here an extension to DS which allows it to include such information explicitly and draw the distinctions relevant for SUs.

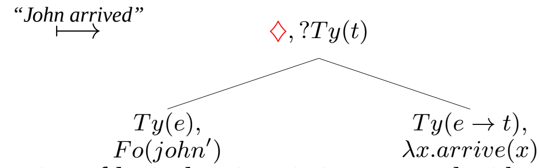
We also show how this extension is motivated by the resolution of the *Split Turn Taking Puzzle* (STTP). This is a version of Ginzburg (1997)'s Turn Taking Puzzle applied to SUs, where it appears that distinct empirical results are obtained: given a SU split between two people, the possible interpretations of a subsequent "Why?" depend not on the most recent speaker, but on who can be taken as the agent of the speech act performed – which may be distinct from the notion of 'speaker' tracked by indexical pronouns like *I* and *my*.

2 Combining Dynamic Syntax with TTR

2.1 Dynamic Syntax (DS)

DS combines word-by-word incrementality with context-dependent, goal-directed parsing defined over partial trees. Importantly, these trees are semantic objects, rather than reflecting syntax or word order. Parsing in DS relies on the execution of licensed *actions*, as incorporated in lexical entries (as in (3) below); such actions resolve outstanding requirements (here, $?Ty(e)$) to decorate the tree with information about semantic type Ty and content (formula) Fo:

(3) *john:*
IF $?Ty(e)$
THEN put$(Ty(e))$
put$(Fo(john'))$
ELSE abort

(4)


Application of lexical actions is interspersed with the execution of computational rules which provide the predictive element in the parse and provide the compositional combinatorics. For example, eventual type deduction and function application is achieved by means of the rule of *Elimination* (5). This derives the value of a mother node's semantic type Ty and content Fo from that of its daughters, in (4) providing the values $Ty(t), Fo(arrived(john))$ at the top node:

(5) *Elimination:*
IF $?Ty(T_1)$,
$\downarrow_0 (Ty(T_2), Fo(\alpha))$
$\downarrow_1 (Ty(T_2 \rightarrow T_1), Fo(\beta))$
THEN put$(Ty(T_1))$
put$(Fo(\beta(\alpha)))$
ELSE abort

Grammaticality is then defined in terms of a resulting complete (requirement-free) tree. Generation is defined in terms of parsing, and therefore also functions with partial trees, uses the same action definitions, and has the same context-dependence, incrementality and predictivity.

DS is thus well-placed to account for SUs: equal incrementality in parsing and generation, and the use of the same partial tree representations, allows the successful processing of "interruptive" SUs with speaker changes at any point. As *goal trees* (planned messages driving generation) may also be partial, utterances may be produced before a total propositional message has been constructed, and completions may be analysed without necessarily involving "guessing". The parser-turned-producer has just to access a word that seems to them an appropriate completion, without *necessarily* considering whether it matches the previous speaker's intention.

DS doesn't incorporate a notion of dialogue act type (in contrast to e.g. Ginzburg et al. (2003)) as it is assumed that the linguistically provided information is highly underspecified, namely just an indication of sentence mood as declarative, interrogative, imperative.[2] However, as the DS formalism

[2] Such specifications are currently encoded as features translatable into use-neutral procedural instructions, unless there are "grammaticised" associations between moods and speech acts, an empirical issue to be decided on a language-by-language basis.

is designed to interact with context incrementally at any point, the possibility of deriving speech act information from context exists; although the interface to enable this must be specified. For this reason, we now turn to TTR, a transparent representation format allowing the specification and interaction of multiple types of information.

2.2 Type Theory with Records (TTR)

TTR has already been used in dialogue modelling (Cooper and Ginzburg, 2002; Ginzburg, forthcoming). Tokens (*records*) and types (*record types*) are treated uniformly as structured representations – sequences of $label : type$ pair *fields* – with the result that their interaction can be modelled in a single system, as required when dealing with metacommunicative uses of language such as 'repair'-constructions or grounding.

Here, the attraction of TTR is that it allows the stratification of multiple types of information, using distinct field labels. The device of *dependent types* allows linking of information between fields, as types can depend on types occurring earlier in the record (higher up in the graphical representation). This allows us to separate contextual information (e.g. information about conversational events, including speaker, addressee, time, location etc.) from the semantic content directly derived from the linguistic string, but allow interaction between the two; this is what we need for phenomena like resolution of ellipsis or assigning values to indexicals and anaphoric elements.

2.2.1 Using TTR in DS

TTR has not, however, been defined in an incremental manner.[3] Here, then, we use TTR representations within the DS vocabulary of trees and actions, replacing the unstructured content of the $Fo()$ labels with TTR record types, and interpreting $Ty()$ simple type labels (and requirements) as referring to *final* TTR field type. Compare the modified lexical entry and eventual tree representation below with the ones displayed in (3)-(4):

(6)

john:	
IF	$?Ty(e)$
THEN	put($Ty(e)$)
	put($\left[\ x : john\ \right]$)
ELSE	abort

(7) $\overset{\text{"John arrived"}}{\longmapsto}$

$$\Diamond, Ty(t), \left[\begin{array}{l} x : john \\ p : arrive(x) \end{array}\right]$$

$$Ty(e), \left[\ x : john\ \right] \qquad Ty(e \rightarrow t), \lambda\left[\ x : e\ \right].\left[\begin{array}{l} x : e \\ p : arrive(x) \end{array}\right]$$

Function application and type deduction will now apply under a suitably modified rule of the DS *Elimination* process; see (9) below.

3 Utterance Events

An account of SUs must explain how indexical pronouns can assume distinct values around a change of speaker (and addressee). We therefore require some record of the utterance event/situation which includes information about speaker/addressee identity. Note that the availability of utterance events to the semantics is independently motivated by e.g. event reference via anaphora ("what do you mean by *that*?") (see also Poesio and Rieser 2010).

We assume that utterance events should at minimum record participant information and who is uttering which particular word(-string). We therefore introduce a partition within the TTR representation of content, with utterance event information held in a *context* (or $ctxt$) field, and linguistically derived semantic content in a *content* (or $cont$) field. The $ctxt$ field is itself structured, containing the required information about utterance event, speaker and addressee; we assume this is available directly from the real-time context of utterance:[4]

(8)
$$\left[\begin{array}{lcl} a & : & participantA \\ b & : & participantB \\ u & : & utt-event \\ s_s & : & spkr(u,a) \\ s_a & : & addr(u,b) \end{array}\right]$$

In a fuller treatment, this utterance context information should also include further information such as time of utterance, world etc, but we omit these here for simplicity.

The DS *Elimination* process must now perform beta-reduction (as before) for the $cont$ field, and TTR extension (i.e. concatenation (Cooper, 1998), shown here as $\oplus$) for the $ctxt$ field, as shown in (9), (10). Parsing a two-word utterance *John arrived* spoken by one speaker, A, will therefore now result in a representation as in Fig 1.

[3] Work is underway to introduce incrementality in the TTR model via the subtyping relation (White (in prep); Meyer-Viol (in prep)). Here we pursue a more conservative strategy.

[4] This is a simplification, of course: determination of addressee is not trivial – see (Goffman, 1981) amongst others.

Figure 1: Tree structure derived from *John arrived* spoken by a single speaker $participantA$

$$\diamondsuit, Ty(t), \begin{bmatrix} ctxt: & \begin{bmatrix} a & : & participantA \\ u_0 & : & utt-event \\ s_{s0} & : & spkr(u_0, a) \\ u_1 & : & utt-event \\ s_{s1} & : & spkr(u_1, a) \end{bmatrix} \\ cont: & \begin{bmatrix} x & : & john \\ p & : & arrive(x) \end{bmatrix} \end{bmatrix}$$

$$Ty(e), \begin{bmatrix} ctxt: & \begin{bmatrix} u_0 & : & utt-event \\ s_{s0} & : & spkr(u_0, a) \end{bmatrix} \\ cont: & \begin{bmatrix} x & : & john \end{bmatrix} \end{bmatrix} \qquad Ty(e \rightarrow t), \begin{bmatrix} ctxt: & \begin{bmatrix} u_1 & : & utt-event \\ s_{s1} & : & spkr(u_1, a) \end{bmatrix} \\ cont: & \lambda[x].\begin{bmatrix} p & : & arrive(x) \end{bmatrix} \end{bmatrix}$$

(9) *Elimination:*

IF $?Ty(T_1)$,
$\downarrow_0 (Ty(T_2), \begin{bmatrix} ctxt: & c_1 \\ cont: & \alpha \end{bmatrix})$
$\downarrow_1 (Ty(T_2 \rightarrow T_1), \begin{bmatrix} ctxt: & c_2 \\ cont: & \beta \end{bmatrix})$
THEN put$(Ty(T_1), \begin{bmatrix} ctxt: & c_1 \oplus c_2 \\ cont: & \beta(\alpha) \end{bmatrix})$
ELSE abort

(10)

$$\diamondsuit, Ty(t), \begin{bmatrix} ctxt: & c_1 \oplus c_2 \\ cont: & \begin{bmatrix} x & : & e \\ p & : & f(x) \end{bmatrix} \end{bmatrix}$$

$$Ty(e), \begin{bmatrix} ctxt: & c_2 \\ cont: & \begin{bmatrix} x & : & e \end{bmatrix} \end{bmatrix} \qquad Ty(e \rightarrow t), \begin{bmatrix} ctxt: & c_1 \\ cont: & \lambda\begin{bmatrix} x & : & e \end{bmatrix}.\begin{bmatrix} x & : & e \\ p & : & f(x) \end{bmatrix} \end{bmatrix}$$

3.1 Indexical Pronouns

Importantly, the definitions of TTR mean that semantic $cont$ information can depend on values in the earlier $ctxt$ context field (although not vice versa). Given this, an explanation of the reference of *I* and *you* becomes expressible. First-person pronouns are defined to take their semantic value from the value of the speaker information in $ctxt$; second-person pronouns from the addressee ($\mathbf{x}$ and $\mathbf{u}$ are rule-level variables binding terms on the nodes where the rules apply).

(11) *I*:
IF $?Ty(e), \begin{bmatrix} ctxt & : & \begin{bmatrix} s_s : spkr(\mathbf{u}, \mathbf{x}) \end{bmatrix} \end{bmatrix}$
THEN put$(Ty(e))$, put$(Fo(\mathbf{x}))$
ELSE abort

(12) *You*:[5]
IF $?Ty(e), \begin{bmatrix} ctxt & : & \begin{bmatrix} s_a : addr(\mathbf{u}, \mathbf{x}) \end{bmatrix} \end{bmatrix}$
THEN put$(Ty(e))$, put$(Fo(\mathbf{x}))$
ELSE abort

As grammatical constraints in DS are phrased in terms of semantic features (rather than syntactic features), the grammaticality of examples like (2) now becomes almost trivial. While a syntactic account would have trouble explaining how *myself* can be co-referential with its antecedent *you*, there is no such problem here: as *you* uttered by A and *myself* uttered by B annotate the trees with co-referential semantic variables. The lexical entries for reflexives such as *myself* must check for a suitably co-referential subject elsewhere in the tree (via the co-argument constraint $\uparrow_0\uparrow_{1*}\downarrow_0 \ Fo(\mathbf{x})$), and here, this will be available:

(13) *myself*:
IF $?Ty(e), \begin{bmatrix} ctxt & : & \begin{bmatrix} s_s : spkr(\mathbf{u}, \mathbf{x}) \end{bmatrix} \end{bmatrix}$,
$\uparrow_0\uparrow_{1*}\downarrow_0 \ Fo(\mathbf{x})$
THEN put$(Ty(e))$, put$(Fo(\mathbf{x}))$
ELSE abort

4 Speech acts

Purver et al. (2009) show that SUs are often not straightforward in speech act terms: sometimes they continue/complete the original speech act; sometimes they perform a new one, clarifying/confirming a suggested completion; sometimes they are ambiguous and/or multifunctional. In order to express these important differences, we need the ability to represent and reason about speech act information (see e.g. (Ginzburg et al., 2003; Asher and Lascarides, 2003)).

Importantly, we would like any inferences about speech acts to be *optional*. A parser should enable these inferences when the appropriate function of the turn is at issue (e.g. in cases of 'repair'), but they should not have to be derived for intelligibility or the determination of grammaticality. They should also be derivable retrospectively: as a result of an interlocutor's feedback, one can assign a particular force (even to one's own contribution) that had not occurred to them beforehand.

Any computational rules that introduce such inferences must therefore be available in the grammar but optional (except where the association of a specific construction with a particular interpre-

[5] A more complex set of actions may be required to account for the fact that *you* may be singular or plural in reference, may include the hearer or not and may be generic.

tation has been grammaticised); and the resulting representations should be kept distinct from those derived directly from the parsing of linguistic input. DS already provides a mechanism which suits these requirements: the use of LINKed trees (trees which share some semantic variable), as in the analysis of non-sentential fragments (Gargett et al., 2009) and relative clauses (Kempson et al., 2001). This device of LINKed trees expresses the cognitive reality of distinguished local domains as evinced by standard syntactic tests, e.g. island-constraints and binding restrictions (see e.g. (Gregoromichelaki, 2006)). As TTR currently does not provide the means for such syntax-semantics interface restrictions we retain the notion of LINKed trees here.

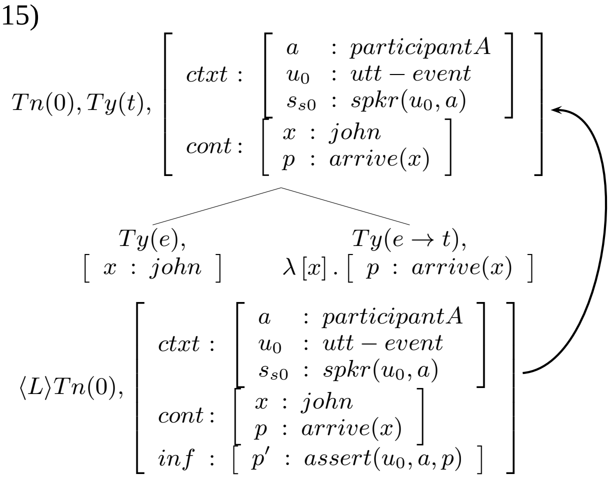
As speech act information can be highly underspecified and context-dependent, we do not wish to assume here either a fixed range of speech acts or a fixed set of inferences from linguistic form to speech act type. We therefore take the rules introducing such information to be of the form sketched in (14). When applied, this rule will introduce a new LINKed tree and provide a Fo value $\mathbf{A}(\mathbf{V}, \mathbf{U}, \mathbf{F}(\mathbf{p}))$ where $\mathbf{A}$ is a metavariable ranging over speech act specifications, $\mathbf{V}$ the agent responsible for the speech act, $\mathbf{U}$ an utterance event (or sequence of events), and $\mathbf{F}$ some function over the semantic content of the utterance ($\mathbf{p}$ and $\mathbf{x}$ are rule-level variables binding terms on the nodes where the rules apply):[6]

(14)

IF	$Ty(\mathbf{x}), Fo(\mathbf{p})$
THEN	make(L), go(L)
	put($\mathbf{A}(\mathbf{U}, \mathbf{V}, \mathbf{F}(\mathbf{p}))$)
ELSE	abort

In order to distinguish content that is derived directly on the basis of linguistically provided information and content derived on the basis of such inferences we introduce a partition in the TTR representation: we take the *cont* field to indicate the (linguistically-derived) truth conditional content and introduce an *inf* field for the speech act content derived by means of such rules (this roughly corresponds to the *explicature*/*high level explicature* distinction in Relevance Theory). So, for illustration, a suitable (optional) rule for assertions might perhaps apply to $Ty(t)$ trees with proposition p and speaker a, allowing one to infer the extra content $assert(a, p)$:

(15)



4.1 SUs and speech acts

Given this, we can outline an account of SUs in which the same linguistic input can be construed as performing different possible speech acts (perhaps simultaneously). Consider the simple (and constructed) example in (16):

(16) A: John ...
B: arrived?

There are (at least) two possible readings of the resulting collaboratively produced contribution: one in which B is (co-)querying whether John arrived; and one in which B is clarifying A's original speech act, i.e., B is asking whether A was asking that John arrived. The tree resulting from parsing (or producing) this SU will be similar to the one in Fig 1 above, except that, due to the speaker change, the second utterance event u_1 is shown as spoken by B (see the unboxed part of Fig 2).

Applications of computational rules as in (14) above allow us to infer the speech act information corresponding to the two possible readings, deriving LINKed sub-trees which indicate speech acts performed by whichever participant is taken as the agent. One possible rule would derive the simple "co-querying" reading (based on the interrogative intonation and the identity of the final speaker B) adding the speech act proposition that B is asking whether John arrived – see the upper box in Fig 2. An alternative rule would derive the "clarificational" reading shown in the lower box. Of course, other inferences may also be possible.[7]

[6] The nature of **F** will depend on speech act type; for an assertion, it may simply be the identity operator; for irony, negation (see e.g. Asher and Lascarides (2003) for suggestions on how speech act type may relate to semantics).

[7] If such inferences become grammaticised, i.e. a particular construction is associated with a particular act (e.g. *clarification*), only one rule may be available. This is an empirical issue which we set aside here, but see (Ginzburg, forthcoming).

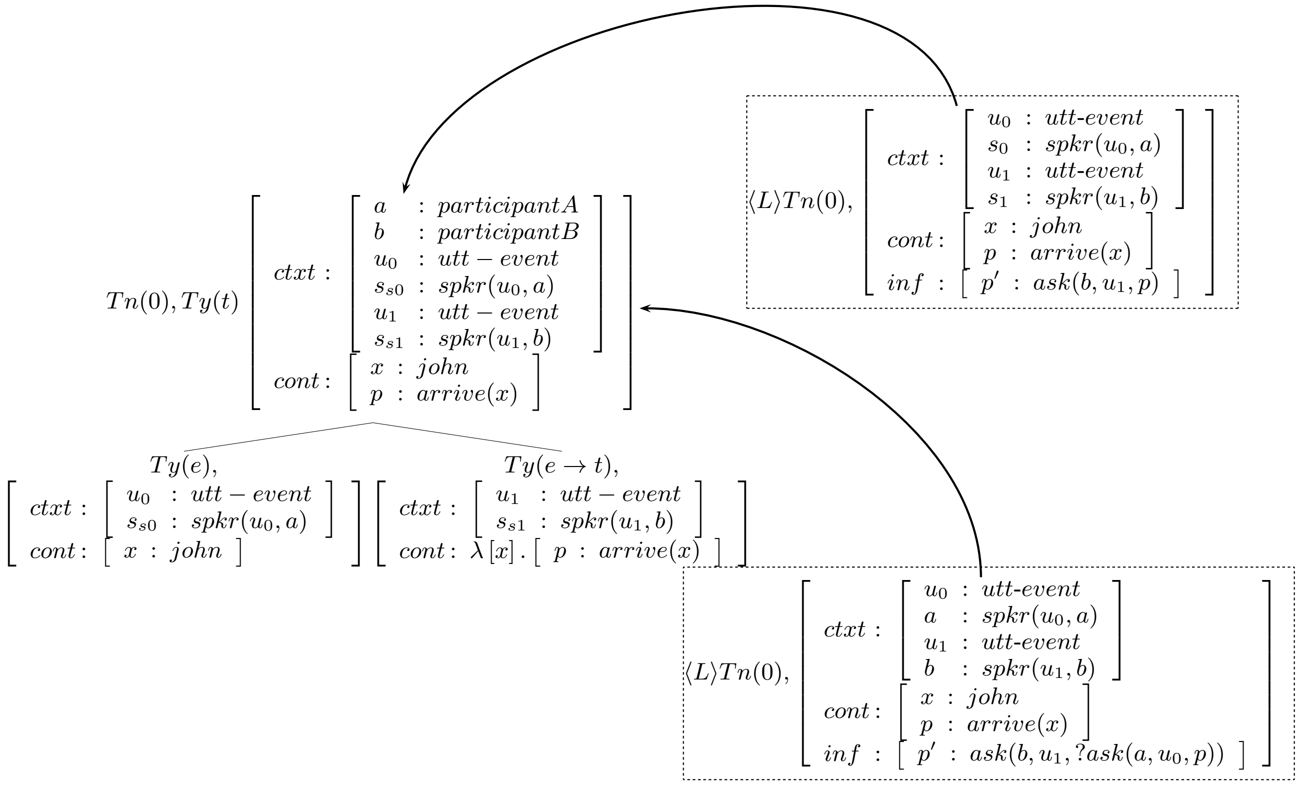
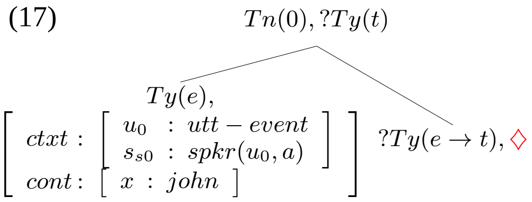


Figure 2: SU-derived tree

Note that Fig 1 and Fig 2 display representations of the final state that a parser might be in after B's contribution; from an incremental processing point of view, we are also interested in the state at the transition point (the change in speaker). Without considering any speech act inference, the tree at this transition point will be as follows:

(17)



This tree is partial (i.e. incomplete, having as yet unsatisfied requirements), but in itself is enough for B to begin generating – provided that they have some suitable message in mind (encoded as a *goal tree* in DS) which is subsumed by this partial tree. There is no requirement for B (or indeed A) to complete this tree, or perform any inference about speech acts, in order to begin generation (or, in A's case, parsing). In cases where B's continuation matches what the original speaker A could have intended to convey, the appearance would be one of "guessing", even though B has not performed any kind of inference regarding A's speech act. In fact, as (18) shows, completions of another speaker's utterance by no means need to be what the original speaker actually had in mind:

(18) 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)]

Such continuations can be completely the opposite of what the original speaker might have intended as in what we will call "hostile continuations" or "devious suggestions" – which are nevertheless collaboratively constructed from a syntactic point of view:

(19) (A and B arguing:)
A: In fact what this shows is
B: that you are an idiot

(20) (A mother, B son)
A: This afternoon first you'll do your homework, then wash the dishes and then
B: you'll give me £10?

Given a suitable model of the domain at hand, B, sometimes, will presumably be able to determine the content of A's intended speech act and represent it as such, i.e., as a speech act emanating from A, in their goal tree (see e.g. Poesio and Rieser (2010)). We take this not to be an essential process for the production of SUs, although it could be necessary in cases where B's next move is specifically intended as a confirmation request for such a representation.

4.1.1 The Split Turn-Taking Puzzle

Ginzburg (1997) describes a Turn-Taking Puzzle (TTP), which, he argues, shows that options for ellipsis resolution are distinct for speaker and hearer. This is illustrated by means of *why*-fragments:

(21) A: Which members of our team own a parakeet?
B: Why? (= *'Why are you asking which members of our team own a parakeet?'*)

(22) A: Which members of our team own a parakeet? Why?
(a) = *'Why own a parakeet?'*
(b)# *'Why am I asking this?'*

(23) A: Which members of our team own a parakeet? Why am I asking this question?

According to Ginzburg, the reading in which *why* queries the intended speech act (the why_{meta} reading) is available when asked by B (21) but unavailable when asked by the original speaker A (22). However, this is not simply due to coherence or plausibility, as it is available in (23) when expressed by non-elliptical means. Its unavailability must therefore be related to the way context is structured differentially for speaker and hearer.

Our explanation of this puzzle takes the why_{meta} interpretation as querying the intention/plan[8] behind the original speaker's speech act.[9] Since ellipsis resolution requires the potential for immediate accessibility of a salient representation, the infelicity of (22b) shows that the speaker's own intention behind their speech act is, in general, not salient enough for them to question it through *why*-ellipsis[10] (in Ginzburg's formulation such a fact does not belong in the TOPICAL FACTS field; however, this fact obtaining is not impossible, as (23) shows). Under this explanation, the TTP then reveals which agent takes responsibility for performing the relevant speech act, and hence can be queried about their intentions behind this act. In terms of (Goffman, 1981)'s distinctions among "speaker"-roles, the relevant agent is the 'Principal'. This can be evident in cases of SUs in multi-party dialogue. Now the utterer of a completion (the *final* "speaker" in the general sense discussed so far, and as indexed by pronouns like *my*) can felicitously ask elliptical why_{meta} questions of the *original* speaker (we will call this phenomenon the STTP, or Split Turn-Taking Puzzle):

(24) A to C: Have you finished sharpening ...
B to C/A: my loppers? B to A: Why?
(a) = *'Why are you asking C whether she has finished sharpening my loppers?'*
A to B: Because I want her to sharpen my secateurs too.

We can explain B's *why*-fragment interpretation in (24a) if we assume that although B's fragment *my loppers?* completes A's question, B does not necessarily assume responsibility for the performance of the speech act. That is, A must be taken as the agent of the querying speech act even though there is a sequence of utterance events which A and B have performed severally.[11] The availability of the why_{meta} reading then follows, even though apparently in contrast to (22b).

In some cases, then, even though the turn is collaboratively constructed, the original speaker maintains the authority or responsibility for the turn even though it was completed by somebody else. In other cases, see e.g. the hostile completions (19) and devious suggestions (20), this is not the case: the eventual content derived has to be taken as solely attributable to the second speaker. Notice however that in all cases (except those of direct quotation), the content of indexicals like *my* and *you* tracks directly the actual speaker/addressee, irrespective of who is taking responsibility for the content (or speech act performance). Even in helping out somebody to finish their sentence such indexicals will track the actual utterer/listener:

(25) Child (playing with toy garden tools): Give me my . . .
Mum: your secateurs. Here they are, in fact these are loppers.

(26) A: Next cut it with your ...
B: my loppers. No, this we cut with the secateurs.

[8]Note that this approach does not necessitate that speech act and therefore intention information is available PRIOR to the processing of the *why*-question: instead, seeking to interpret such questions can be the trigger for optional (speech-act inducing) rules to apply. Hence, this approach is perfectly compatible with the general view on intentions as post-facto constructs (see e.g. Suchman (2007)) and the fact that conversational participants negotiate the content of speech acts with such assignments able to emerge retrospectively.

[9]As (Ginzburg, forthcoming) notes, recognition of this intention is *not* necessary for grounding.

[10]However, it is not impossible:

(i) A: Piss off. Why? Probably because I hate your guts.

[11]In fact, the specification of the why-fragment as why_{meta} can be taken to trigger the inference that A is solely responsible for the query as B dissociates himself from it.

This provides evidence for the dissociation of speech act performance and performance of the utterance event: these are two distinct actions whose agent might coincide but not necessarily so (these two roles roughly correspond to Goffman (1981)'s 'Author/Principal' and 'Animator'). Most accounts conflate the two: Lascarides and Asher (2009) argue that each time a speaker makes a conversational move they undertake a *public commitment*. However, SU examples such as (1)-(2) and (25)-(26) show that the person undertaking the public commitment (the 'Principal') does not necessarily coincide with actual utterer (the 'Animator'). We therefore conclude that the notion of 'commitment' should be correlated with something else, namely, who is performing (the agent of) the associated speech act (which could be the two speakers jointly but not necessarily and not only for SUs). Speech act inference rules as outlined in (14) must therefore maintain the flexibility to assign the inferred speech act to any of the speakers involved, and not only the final one.

5 Conclusions

The STTP and the multifunctionality of SU fragments motivates our claim that information manipulated during a parse has to be distinguished at three levels: semantic content which is directly derived on the basis of the linguistic input, context specifications arising from the utterance situation (utterance events) and optional speech act information. Formulation of this information in a DS-TTR combined formalism allows the interactions required for appropriate processing of SUs.

Acknowledgments

We gratefully acknowledge valuable support and advice from Ruth Kempson and Greg Mills. Thanks also to Pat Healey, Graham White, Arash Eshghi and Chris Howes. This work was supported by DynDial (ESRC-RES-062-23-0962).

References

N. Asher and A. Lascarides. 2003. *Logics of Conversation*. Cambridge University Press.

R. Cooper and J. Ginzburg. 2002. Using dependent record types in clarification ellipsis. In *Proceedings of the 6th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL)*.

R. Cooper. 1998. Information states, attitudes and dependent record types. In *Proceedings of ITALLC-98*.

R. Fernández and J. Ginzburg. 2002. Non-sentential utterances: A corpus-based study. *Traitement Automatique des Langues*, 43(2).

A. Gargett, E. Gregoromichelaki, R. Kempson, M. Purver, and Y. Sato. 2009. Grammar resources for modelling dialogue dynamically. *Cognitive Neurodynamics*, 3(4):347–363.

J. Ginzburg, I. Sag, and M. Purver. 2003. Integrating conversational move types in the grammar of conversation. In *Perspectives on Dialogue in the New Millennium*, volume 114 of *Pragmatics and Beyond New Series*, pages 25–42. John Benjamins.

J. Ginzburg. 1997. On some semantic consequences of turn taking. In *Proceedings of the 11th Amsterdam Colloquium on Formal Semantics and Logic*.

J. Ginzburg. forthcoming. *The Interactive Stance: Meaning for Conversation*. Oxford University Press.

E. Goffman. 1981. *Forms of Talk*. University of Pennsylvania Press.

E. Gregoromichelaki. 2006. *Conditionals: A Dynamic Syntax Account*. Ph.D. thesis, King's College London.

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

R. Kempson, A. Gargett, and E. Gregoromichelaki. 2007. Clarification requests: An incremental account. In *Proceedings of the 11th Workshop on the Semantics and Pragmatics of Dialogue (DECALOG)*.

A. Lascarides and N. Asher. 2009. Agreement, disputes and commitments in dialogue. *Journal of Semantics*, 26(2):109–158.

G. Lerner. 1991. On the syntax of sentences-in-progress. *Language in Society*, pages 441–458.

M. Poesio and H. Rieser. 2010. Completions, coordination, and alignment in dialogue. *Dialogue and Discourse*, 1:1–89.

M. Purver, J. Ginzburg, and P. Healey. 2003. On the means for clarification in dialogue. In *Current and New Directions in Discourse & Dialogue*, pages 235–255. Kluwer Academic Publishers.

M. Purver, R. Cann, and R. Kempson. 2006. Grammars as parsers: Meeting the dialogue challenge. *Research on Language and Computation*, 4(2-3):289–326.

M. Purver, C. Howes, E. Gregoromichelaki, and P. G. T. Healey. 2009. Split utterances in dialogue: a corpus study. In *Proceedings of SIGDIAL*.

L. Suchman. 2007. *Human-machine reconfigurations: Plans and situated actions*. Cambridge University Press.