Quotation via Dialogical Interaction

Jonathan Ginzburg, Robin Cooper

Quotation via Dialogical Interaction

Jonathan Ginzburg · **Robin Cooper**

Abstract Quotation has been much studied in philosophy. Given that quotation allows one to diagonalize out of any grammar, there have been comparatively few attempts within the linguistic literature to develop an account within a formal linguistic theory. Nonetheless, given the ubiquity of quotation in natural language, linguists need to explicate the formal mechanisms it employs. The central claim of this paper is that once one assumes a dialogical perspective on language such as provided by the KoS[1] framework, formalized in a rich type theory like Type Theory with Records (TTR), much of the mystery evaporates. In particular, one can utilize as denotations for quotative constructions entities that are independently motivated for dialogue processing—utterance types and locutionary propositions, austinian propositions about speech events.

Keywords Quotation · Dialogue · Type Theory with Records

1 Introduction

Quotation has been much studied in philosophy (see Cappelen and Lepore (2012); De Brabanter (2010) for some recent surveys). There have been comparatively few attempts within the linguistic literature to develop an account within a linguistic theory.

Jonathan Ginzburg
CLILLAC-ARP & LabEx–Empirical Foundations of Linguistics (EFL), Université Paris-Diderot
E-mail: yonatan.ginzburg@univ-paris-diderot.fr

Robin Cooper
Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg, Box 200, 405 30 Göteborg, Sweden

[1] KoS is not an acronym, despite emphasizing a Konversationally Oriented Semantics.

Partee (1973) is a classic paper in this area. More recent proposals include Geurts and Maier (2005); Potts (2007); Bonami and Godard (2008). This apparent lack of interest in the phenomenon on the part of linguists might be attributed to the fact that quotation is a means of *diagonalization* (in the Cantorian sense) out of any grammar—for any expression *e* deemed ungrammatical by grammar Γ, one can produce via quotation a well formed string that includes *e*, hence undermining Γ. Thus, we can quote something that is ungrammatical in our own language as in (1a) or something that is in a different language to the one we are speaking (1b) or even something that doesn't seem to involve speech sounds that occur in any natural language (1c) or sounds made by inanimate objects (1d).

(1) a. Damien, who's only four years old, said 'I go'ed to Grandma's'
     b. Pelle, whose native language is Swedish, said 'Jag har varit hos mormor' (meaning "I've been at Grandma's")
     c. Morry aimed the toothbrush at David and went '[æʔæʔæʔæʔ]' (Partee, 1973)
     d. The blender went 'plplplpl'

Linguists might be forgiven if they do not regard it as being part of the job of somebody writing a formal grammar of English to include everything that can occur between quotation marks in sentences like those in (1). From a linguistic perspective, one could easily form the opinion that quotation is on the margins of linguistic data, something that a formal grammarian could, and possibly should, ignore. In a similar vein, a wide range of philosophical accounts (e.g. Davidson (1979); Werning (2012), among many others) assume that what occurs between the quotation marks in the surface syntax is not part of the sentence in which those quotation marks occur. This, however, presents challenges for dealing with examples where the form or the content of the quotation is referred to from outside the quotation as in (2).

(2) a. 'I talk better English than the both of youse!' shouted Charles, thereby convincing me that he didn't. (Partee, 1973, ex. 20)
     b. The sign says 'George Washington slept here', but I don't believe he really did. (Partee, 1973, ex. 26)
     c. What he actually said was, 'It's clear that you've given this problem a great deal of thought,' but he meant quite the opposite. (Partee, 1973, ex. 32)

And indeed there is no little amount of evidence that quotation is subject to general grammatical principles governing word order, ability to be embedded and psuedo-clefted, and semantic selection (Postal, 2004; Bonami and Godard, 2008).

Moreover, all natural languages seem to have quotation of some kind. Children use direct quotation from their earliest utterances (Ginzburg and Moradlou, 2013). Indirect quotation is acquired substantially later, but is nevertheless standardly present in the adult language.[2,3] Given the ubiquity of quotation in natural language, linguists need to explicate the mechanisms it employs. Indeed, one is obligated to do so

---

[2] Everett, 2012 claims that Pirahã lacks *indirect* quotation.

[3] We limit our discussion and treatment here to 'pure', 'direct' and 'indirect' quotation, putting aside the recently much discussed 'mixed' and 'scare' quotation. We believe the account we develop can be extended to these, but that remains a future project.

in a way that offers an answer to the question: **why, rather than being a heterodox linguistic process, is in fact quotation so straightforward**?

The short answer, we suggest, is that this is because quotation involves entities and mechanisms utilized ubiquitously during dialogue processing. In slightly more detail—dialogue processing involves a periodic interaction concerning whether an utterance has been understood (*grounding* (Clark, 1996)) and if not the need to engage in clarification. On one model of this process developed in the KoS framework (see (Ginzburg, 2012)) this involves assessing the truth of a *locutionary proposition*—a proposition constructed from the utterance event and a linguistic *sign* made available by grammatical resources available to the dialogue participants—assessing whether the utterance can be fully characterized or not by the sign and if not inferring a clarification question. On this view, *locutionary propositions*, of which signs (speech event types) are fundamental components, are the essential means for representing utterances in an interlocuter's view of the context—their *dialogue gameboard*. We will propose, in section 4, that signs and *locutionary propositions* offer simple and semantically parsimonious denotations for (pure/direct) quotation, thereby directly tying them to utterance representations used ubiquitously in dialogue.

The central claim of this paper, then, is that once one adopts a semantic perspective rooted in dialogue and utilizing a *rich type theory* many of the recalcitrant issues associated with a semantic account of quotation dissolve. This perspective incorporates four fundamental assumptions:

1. *Language as a system in flux.* Cooper and Ranta (2008); Cooper (2012) argue that natural language grammar should be regarded in terms of a collection of resources that a linguistic agent has available in order to build local (possibly formal in the sense of Montague, 1974) languages on the fly appropriate to a particular situation. This is meant to provide a basis for the fact that speakers coordinate both the form and meaning of their language with their interlocutors (Cooper and Larsson, 2009; Larsson and Cooper, 2009), a fact which is normally abstracted away from in formal treatments of grammar but which has attracted a good deal of attention in the psychological literature on language (e.g. Brennan and Clark, 1996; Healey, 1997; Healey et al, 2007). The constant micro-adjustment of language as we talk with others (particularly, but not solely, on the part of children acquiring the language) is the driving mechanism in linguistic change.

   For quotation this means that it is no longer surprising that a speaker should create an *ad hoc* language blending resources from different languages or dialects which the speaker has available and even including non-linguistic resources. Since we are not in the business of building one monolithic formal grammar for English but rather explaining what resources a linguistic agent has to have available to construct a local language for a particular purpose, the formal problems associated with these kinds of quotations do not arise.

2. *Linguistic events.* In our previous work (most recently Ginzburg, 2012; Cooper, 2012) we take seriously the fact that natural language grammar concerns linguistic events and that this should be reflected in grammatical theory. Our inspiration for this comes originally from early work in situation semantics (Barwise and

Perry, 1983) which showed us how to make grammatical sense out of speech act theory (Austin, 1962; Searle, 1969, and, of course, much subsequent literature). For quotation this means that we analyze direct quotation as referring to a particular linguistic (speech, reading or writing) event. This situates our approach within the family of demonstrative theories (de Cornulier, 1978; Clark and Gerrig, 1990).

3. *Event types as central to linguistic processing and understanding.* On our event based approach to language it is important that linguistic generalizations and procedures be expressed not in terms of the events themselves but in terms of *types* of events (or situations). This use of event types enables us to express rules which predict possible future events. It also allows us to characterize certain aspects of events and *underspecify* the nature of an event by using a type to which the event in question belongs but to which other events could also belong. The formalism we will use for this is Type Theory with Records (Cooper, 2012, TTR) which incorporates ideas from situation theory and Martin-Löf Type Theory (Martin-Löf, 1984; Nordström et al, 1990).

   The relevance of this for quotation has to do with the fact that while we are referring to a previous speech event with a direct quotation we cannot reproduce that event directly (except perhaps by playing a recording of it, which would not count as a quotation). The direct quotation of the previous speech event is a new event which shares a type with the original event. Which type we choose to illustrate with the quotation varies depending on the context and the resources we have available.

   Types of linguistic events also play a role in our analysis of pure quotation. Words (like pieces of music, Cooper, 2013b) can be thought of as types of events, either types of utterances or types of written occurrences. Thus when we say that "Anna" begins with "a" we mean that any written occurrence of the type "Anna" begins with an instance of the letter "a", itself a type of written occurrence of a letter. Alternatively, we mean that any utterance event of the type "Anna" begins with an event of an utterance of the type which is phoneme /æ/, itself a type of utterance event.

4. *Language as a sign and feature based system.* It will be important for our approach to quotation to adopt the notion of linguistic sign (de Saussure, 1916), that is a pairing of linguistic form and content. For us the linguistic form will be an event and we will be interested in types of signs, that is types of form-content pairs where the form and the content are required to be of certain types. Our approach to this will be based largely on a type theoretical version of Head-Driven Phrase Structure Grammar (HPSG, Sag et al, 2003). The adaptation of de Saussure's notion of sign in HPSG uses feature structures. We will model feature structures as record types.

   Using such structures gives us a simple minded but powerful notion of similarity between types where similarity can be construed as sharing some features but not necessarily all. Such a notion of similarity is central to a theory of quotation. For example, a quotation in a different language from the original may share the content of the original but not the form. Furthermore, the content may not be exactly the content of the original but similar enough for the purposes of the dialogue. In exploiting the notion of linguistic sign in quotation we are following

the lead of Bonami and Godard (2008) and also Potts (2007) who essentially uses a sign-based (though not feature-based) approach although he does not refer to it explicitly as such.

The rest of the paper is structured as following: Section 2 provides a brief survey of relevant data and relates them to these four aspects. Section 3 introduces TTR and show how it enables the modelling of dialogical interaction. Section 4 applies these tools to develop analyses of pure and direct quotation.

2 Some data

In this section we will present some data on quotation that motivates our account: on pure quotation it justifies our decision to postulate a sign–denoting construction; for direct quotation we provide data suggesting that a similarity measure and that grammatical resources are contextual parameters of such uses; we demonstrate that direct quotation is selected by certain predicates, raising the issue of what the semantic type of such uses is; finally, we demonstrate the close relationship between direct and indirect quotation.

2.1 Pure Quotation

Pure quotation involves *general, rule–like* statements about utterance *types*. As example (3) shows, these can target various dimensions of the linguistic sign, but they do not directly concern any utterance tokens as such:

(3)   a. **syntax**:
            'Bo' is a noun
            'Bo left' is a declarative sentence
       b. **phonology**:
            'Bo' starts with b
            'Bo left' consists of two monosyllabic words
       c. **semantics**
            'Bo' is used to refer to a person named 'Bo'
            'Bo left' predicates leaving of a person named 'Bo'
       d. **context**:
            'Bye' is used as the final move in a conversation

To capture the data in (3), one needs to assume that quotation has a use in which it refers to an entity that incorporates information about the variety of linguistic dimensions (phonological, morphosyntactic, semantic, . . .) relative to a particular language—it is crucial, e.g., that neither intra-linguistic, nor cross-linguistic homophones (English 'rapport' (*relationship*) and French 'rapport' (*report*); Hebrew 'xatuna' (*wedding*) and Georgian 'xatuna' (a proper name)) be conflated. Hence, breezy assumptions about the pure quotation as 'denoting the expression' cannot be maintained.[4]

---

[4] The *locus classicus* for this view is Quine: 'Take quotation marks: applied to any sort of expression, what they produce is a singular term (naming, as it happens, the expression inside).' Quine (1960). Geurts

By shifting the content of a proper name to a type of occurrence of a proper name we can explain apparent failures of compositionality and existential generalization which have been discussed in the literature such as in (4).

(4)   a. Karl was looking for "Mark Twain" in the phone book.
          $\not\to$ Karl was looking for "Samuel Clemens" in the phone book.
      b. "Anne" begins with a vowel.
          $\not\to \exists x$ "x" begins with a vowel (Werning 2012)

We conceive of the use of quotation marks in written text as a symbolic way of representing that the content of the word or phrase within the quotation marks is to be shifted to such a type. So, for example, if we were to treat proper names following the classical Montague (1974) treatment as denoting the set of properties of some object, the quoted version of the proper name "Anne" could be the set of properties of the type of strings of phoneme utterances that consist of an utterance whose phonological dimension has the phoneme /æ/ followed by the phoneme /n/. This could be derived from the sign for *Anne* by replacing the content of the sign with a noun-phrase denotation derived from the original sign.

Note that on this conception, while (4b) indeed represents a failure of existential generalization, if we allow quantification over types and predicates of types in the way we have suggested the existential generalization in (5) holds.

(5)   a. "Anne" begins with a vowel.
      b. $\to \exists x$ x begins with a vowel

The point, of course, of (4b) is that we cannot consider there to be a quotation function, ".", which takes the content of an expression *a* and yields the content of "*a*". Thus in a strict sense quotation is not compositional. (For a recent discussion of compositionality and quotation see Pagin and Westerståhl, 2010.) In the case where the quotation refers to a type of utterance event, we might say that "." is defined not on the content but on the whole sign, picking out the phonology of the sign as the content of the quoted phrase. (This is an essential part of the proposal by Potts, 2007.)

2.2 Direct quotation

Unlike pure quotation, direct quotation involves reference to a particular linguistic event or situation (either a speech act or a textual situation of a given text containing a particular subtext).[5] We consider two contextual parameters that seem crucial for an account and have not been explicitly considered in previous work—similarity measures and grammatical resources.

For a start, there is a question of whether quotation involves reference to an event or, following Clark and Gerrig (1990); Clark (1996), a *demonstration* of an event. For

---

and Maier (2005) say: 'If the quotation is a 'pure' one, it denotes a linguistic expression and its semantic type may differ from that of the quoted expression. For example, in [the example *'Monosyllabic' is not monosyllabic.*] quotation maps the semantic type of the adjective 'monosyllabic' to that of singular terms, and the resulting expression denotes the word 'monosyllabic'.

[5] We will ignore the possibility that texts can be regarded abstractly as *types* of physical manifestations, although we find this plausible.

an illuminating discussion of this issue see Recanati (2010), Chap. 7. Our discussion of direct quotation here is closely allied with Recanati's discussion. We will follow Recanati in talking of reference in cases of what he calls closed quotation, that is, when the quoted material plays a grammatical function in an utterance. The speaker who is quoting a previous linguistic event cannot, of course, reproduce the exact previous event. It is in the nature of events in general that they happen only once. The closest we can come to repeating an event is to create a new event which is similar to the previous one. This is the notion of demonstration introduced by Clark and Gerrig. The kind of reference that Recanati ascribes to what he calls *closed* quotation is achieved by demonstrating the event which is referred to. On our type theoretic approach there are two ways in which one can analyze similarity between two events. One way is to say that there is a type which the two linguistic events (that is, the one being quoted and the demonstration) share. Recanati discusses (6).

(6)   . . . And then Greta Garbo said, 'I want to be alone!'

In demonstrating Garbo's utterance the speaker produces an utterance which shares a type with the original utterance being referred to. In terms of a type of sign which specifies both the form and the content of the utterance we can say that the demonstration and the original share a sign type which specifies the phonological form of the utterance to be of the phonological type represented by "I want to be alone" and the content type corresponding to *Greta Garbo wants to be alone*. Note that in order for this to be a quotation of Greta Garbo's utterance, more is required than that the two speech events share just any type. After all, there are much more general types to which both belong such as *Event*, the type of events or *SpeechEvent*, the type of speech events. One cannot be said to be quoting somebody by producing some random speech event simply because the original and the quotation are both speech events. The type which the two events share has to be specific enough for it to count as similarity.

We want to claim that this need for specificity leads us ultimately not to look at similarity in terms of sharing types but rather to develop similarity measures on types themselves. Let us consider a non-linguistic example first. Suppose that I want to demonstrate the way that John Cleese walks in his Minister of Funny Walks sketch. I am not able to do the walks so they are exactly like those of John Cleese but I might be able to produce walks of a similar type. There might well be types to which both Cleese's performance and my demonstration belong but there is not obviously a guarantee that these types will be specific enough to count as guaranteeing similarity. For language, Recanati discusses Sperber and Wilson's (1986) example (7).

(7)   He said 'We're going to close early tonight'

which, in Recanati's words, "can be used to report an utterance in Italian, provided the Italian sentence had a meaning sufficiently similar to that of the English sentence 'We're going to close early tonight' ". In terms of signs this means that the phonological form of the sentence does not have to be the same and the content only has to be "sufficiently similar". For us the content of a declarative sentence like *we're going to close early tonight* will be defined in terms of a type of situation which could be described by the sentence. Thus it will be expedient for us to consider similarity mea-

sures on the types themselves rather than characterizing similarity in terms of shared types.

A question that arises on this kind of view is whether there is a single similarity measure which can be used in all cases of quotation. It seems that this is not the case. The fact that we can quote utterances that were made in a different language, as in the discussion of (7), suggests that it can be sufficient for the content of the quoting utterance to be the same. Here there is no requirement that the form of the utterance be similar to the original utterance. Even if the quotation is in the same language as the original it normally cleans up disfluencies such as hesitations or mispronunciations unless these are a relevant part of the demonstration involved in the quotation. This point is made clearly in the quotation in (8).[6]

(8)   *Aucun entretien n'est reproduit tel quel. Le journaliste élimine les scories du langage parlé (ne serait-ce que les 'euh' et les 'ah'), les répétitions, les bouts de phrase inutiles, et supprime les fautes de français, sauf à vouloir faire écho à un certain exotisme . . .*
        No conversation is reproduced as such. The journalist eliminates the debris of spoken language (and that's not only the 'um's and 'uh's), the repetitions, the of useless phrases. She corrects the grammatical errors, unless she wishes to convey a certain exoticism

                                                                        R. Solé, Le Monde 7/03/2003

However, the requirements on similarity are not so lax in (9) and (10).

(9)   a. John said, "gorse is beautiful."
      b. ↛ John said, "furze is beautiful." (Werning 2012)

(10)   A: John said "I left". B: No he didn't. He said "I leaved".

In (9) one imagines that part of the point of the quotation is to show which word was used and similarly in (10) *B*'s point is to highlight how John's original utterance deviated from the normal past tense for *leave*. Further examples of different similarity measures are given in (11).

(11) a. A: I can't fight back, it's finished. B: So what you're saying is 'I'm a loser', right?
      b. Jo asked 'Am I crazy?' in French.
      c. Jean said 'Je le ferai' which I think means 'I will do it'.

---

[6] See also the following from the economist Brad DeLong:

It is conventional–at least in the circles in which I travel–to clean-up transcripts. Raw human speech looks silly in print. You eliminate the "ummms" and the "you knows" and the false starts. You collapse into one grammatical sentence those times when people start a sentence, go back and start it again, and then go back and start it again and finally finish it.
The convention is to edit the transcription of what they said into something different–into what they would have written if they had been not speaking but writing, sentence by sentence and paragraph by paragraph, and editing as they go.
But this seems to go rather beyond that, so much so as to convey a substantially distorted impression of how the symposium actually went . . .

http://delong.typepad.com/sdj/2013/12/thursday-idiocy-weblogging.html

d. Jean said 'Je le ferai' which I think means that he will do it.

These differences in the similarity measures used in different contexts are not surprising if the whole linguistic system follows the principle of flux. We would suggest that it is not the job of a linguistic theory to specify exactly which similarity measures are used but rather to exhibit a linguistic theory which enables the definition of relevant similarity measures.

Agents who know a language have grammatical resources which will allow them to assign grammatical types, such as sign types, to speech events they observe. In computational terms we can think of the types assigned to a speech event (utterance) as the result of parsing the utterance. It is these types, not types in general, which have to be similar between the original event and the quotation event. We assume that agents have a collection of grammatical resources and that they draw on different subcollections of these resources on different occasions, depending, for example, on which language or dialect they are speaking, whether they are speaking about theoretical physics or the family outing at the weekend or are involved in a technical discussion about the offside rule.

In many cases of quotation the resources applied to the original utterance—in terms of notation, Γ_1—will be the same as the resources applied to the quotation utterance——in terms of notation, Γ_2. However, this will not be the case when the original utterance was in a different language to the quoting utterance. In this case it may even be that the agent of the quoting utterance does not have access to resources Γ_1 needed to process the original utterance – they may have benefitted from a translation at the time. Similarly the addressee of the quoting utterance may also not have access to Γ_1. Note also, that Γ_2 need not be the same resources that are needed to process the rest of the sentence *Anna said...* if the actual quotation is in a different language or dialect. We can switch resources in the middle of a sentence in a similar way that speakers of more than one language can in code-switching.

The choice of Γ_1 and Γ_2 has consequences for the kind of similarity measure, μ, that can be used. When $\Gamma_1 = \Gamma_2$ we can choose μ to be a strict measure which requires that the resources assign exactly the same sign type to both the original and the quoting utterance. But even here we will often want a looser notion of similarity. For example, a demonstration of an utterance editing out disfluencies we may want to count as a quotation, at least some of the time. Different choices of synonymous words may for some purposes count as quotation for others not. μ may be more or less strict depending on the purposes at hand, an extreme case being (12):

(12)   [An article about an orphaned walrus arriving in a new zoo:] During [the orphan walrus'] first look at a walrus, he was like, 'What's that?' (New York Times, 22/01/2014)

When Γ_1 and Γ_2 are resources for different languages the most we can hope for μ only requires that the contents required by the two sign types are identical. However, as is well known, identity of content between expressions in two languages can be too much to hope for. Suppose the original utterance quoted in (13a) had been the Swedish utterance (13b).

(13) a. Max said 'Bo departed'.

    b. Bo har åkt ("Bo has gone/went (by vehicle)")

There are three things which could be different about the content of (13b) which could make it different from the content of the quotation in (13a). Firstly, it is in the perfect tense. There appear to be cases where a Swedish perfect is appropriately rendered in English by a simple past, but this is not always true. Secondly, Swedish like many other languages makes a distinction between travelling by vehicle and walking. Finally, the verb which serves for *leave* here has the broader meeting *went* and could be used to describe a situation where Bo has arrived. All of these subtle differences are familiar to anybody who has been involved in translation. This has as a consequence that μ must say something about similarity of non-identical contents. Our aim here is not to go into detail characterizing similarity measures that can be used, but rather to make the point that there cannot be a single similarity measure for all instances of quotation. The similarity measures, like the grammatical resources involved, will vary from one quotation event to another.

    Finally, we need to allow for the possibility that the types compared by the similarity measure μ are not necessarily the types assigned by the resources Γ_1 and Γ_2 but subtypes of those types. This is to allow for cases where the demonstration in the quoting event is meant to show a non-linguistic aspect of the original event in addition to the linguistic type, for example, that the speaker of the original event had a cold, was drunk or had a particular expression on their face while making the utterance. This information could be added to the grammatical types to form subtypes and similarity computed between these subtypes. Note that it is important here that we construe linguistic types as types of events so that we can, for example, not only talk of types of events which are utterances of the phoneme /æ/ but also types of events which are utterances of the phoneme /æ/ with a blocked nose. While the blocked nose may not be "linguistically significant" in the standard sense, it can be significant for the content of the quotation.

2.3 Quotative Operators

There are a number of words that require direct quotation as their complements. In English the marker *like* and the verb *go* have a certain usage which requires a direct quotation as in (14a) and (14b) and does not allow an indirect quotation, as exemplified in (14c) and (14d). Such constructions do not even require a speech act to be demonstrated as Partee's example (1c) shows and also (14e). Such constructions exist in many, if not all, languages although they tend to be restricted to an informal spoken register. One might cite French *faire* and Swedish *typ/ba* as in (14f).

(14) a. I asked her if she wanted to read my paper and she was like "Are you crazy?"

    b. I asked her if she wanted to read my paper and she went "Yuck!"
    c. *I asked her if she wanted to read my paper and she was like whether I was crazy
    d. *I asked her if she wanted to read my paper and she went that she didn't, in no uncertain terms

    e. I asked her if she wanted to read my paper and she was like [expression of disgust, slowly shaking head]

    f. först sa han typ, ("first he said like")
        hej, hur gammal är du ? ("hi, how old are you?")
        sen är du såblond ? ("then, are you so blond?")
        ja ba, nae inte precis, ("I just, no not exactly")
        han typ är du singel ? ("he like, are you single?")
            `http://nellierosengren.bloggsida.se/okategoriserad/13`
                            (accessed 6th Dec. 2013)

Given the possibility of non-speech act demonstrations in these examples one might come to the conclusion that direct quotations simply refer to the demonstration and lose the content of the original event, if it was a speech act. There is much discussion in the literature which shows this not to be true, including the examples (2) discussed by Partee. These examples involve predicates that can combine with direct quotation, as well as indirect quotation, as in (15). Indeed, there is a clear relationship between the indirect and direct quotation complements in (15a,b)—both support the inference (15d):

(15)   a. Zohar asked whether she snored.
       b. Zohar asked 'naxarti?'.
       c. Zohar asked 'naxarti?', a Hebrew sentence often uttered by people who have just woken up.
       d. Zohar asked a question, a question about herself.

The examples in (2) and in (15) indicate that, in one way or another, the complement of a *direct* quotation behaves as a bona fide semantic entity, giving rise to nominal and verbal anaphora, as well as to propositional ('the opposite') or interrogative inference. These facts seem to hold even of constructions which require a direct quotation as in (16).

(16)   I asked her if she wanted to read my paper and she was like "Are you crazy?" But I'm not. I'm perfectly sane.

Our earlier claim that language is sign-based and feature-based, will play an important role in identifying different aspects of an utterance and enabling an utterance to be a demonstration of a prior utterance while at the same time making available the normal content of the utterance.

3 A Dialogical Perspective on Semantics

In this section we introduce the tools underlying our view of dialogical perspective on grammar and semantics. We introduce Type Theory with Records (TTR), which enables us both to develop a grammatical ontology, a semantic ontology, and a theory of interaction.

3.1 Grounding and Clarification Interaction Conditions

In dialogue all words and phrases are, in principle, subject to clarification interaction (see (Ginzburg and Cooper, 2004; Purver, 2006; Ginzburg, 2012; Cooper, 2013a))

(17) a. A: Most researchers acknowledge that pragmatics is important.
  b. B: What do you mean by 'most'?
  c. B: acknowledge?/important?/pragmatics?
  d. B: is? Since when?

This is a conventionalized process: as corpus studies have shown, there are a very small number of possible construals of CRs—as request for confirmation, repetition, intended content resolution (see Purver et al (2001); Rodriguez and Schlangen (2004)). There are thus in the aftermath of an utterance (and indeed while it is ongoing) two essential branches:

– **Grounding**: the utterance is understood and signalled as such (Clark, 1996), its content added to the common ground, uptake occurs.
  Clarification Interaction: some aspect of the utterance causes a problem; this triggers an exchange to repair the problem.

Hence, we need an entity off of which both 'shared understanding' update AND clarification potential (*Grounding and Clarification Interaction conditions*) can be read. In order to understand how this can be formalized we turn to sketch the framework that provides the logical underpinnings of KoS, Type Theory with Records (TTR).

3.2 Developing a semantic and grammatical ontology in Type Theory with Records (TTR)

An utterance of a proper name like *Bo* is a sign, that is, it is an event of a certain type relating a certain kind of utterance to a content. We can think of a sign essentially as a pair of an event e and a content c. (We will argue later that more things are involved.) One way of thinking of this pair is as the set $\{e, c\}$. This set, however, does not *label* its elements in any way. Given an arbitrary sign modelled in this way it may not be obvious which element is playing which role. For this reason we use *records*, which include not only the objects in a set but also labels for those objects. For now, we will use the labels 's-event' (for "speech event") and 'cont'. (for "content"). Each label can label only one object in a record. Formally, we can model records as sets of ordered pairs consisting of a label and an object as in (18a). We will represent this as in (18b) in order to emphasize the relationship with feature structures in linguistics and to make larger structures more readable.

(18) a. $\{\langle \text{s-event}, e\rangle, \langle \text{cont}, c\rangle\}$
  b. $\begin{bmatrix} \text{s-event} & = & e \\ \text{cont} & = & c \end{bmatrix}$

Part of what makes TTR a type theory is that each object a belongs to a type T, in symbols, $a : T$. In fact in TTR objects may belong to many types, even infinitely many types. Thus, for example, we may say that e belongs to the type *SEvent* (for "speech event") and c belongs to the type *Cont* (for "content"). Types are regarded as mathematical entities in their own right within the semantic universe. They are not just collections of objects. ':', the "of-type relation", represents a relation which holds between objects and types. This means that types are intensional in the sense that two distinct types can be related to the same set of objects by the of-type relation. We also have record types which are labelled sets of types, that is, sets of ordered pairs of labels and types such that each label in the record type is associated with exactly one type. The type of signs as we have so far discussed them is thus (19a). As with records we represent this in the, for linguists, more convenient form (19b).

(19) a. $\{\langle \text{s-event}, \textit{SEvent} \rangle, \langle \text{cont}, \textit{Cont} \rangle\}$

b. $\begin{bmatrix} \text{s-event} & : & \textit{SEvent} \\ \text{cont} & : & \textit{Cont} \end{bmatrix}$

The of-type relation between records and record types is defined such that a record, r is of a record type T just in case for each labelled type in T r contains a labelled object of that type:[7]

(20) The record

$$\begin{bmatrix} l_1 & = & a_1 \\ l_2 & = & a_2 \\ \ldots & & \\ l_n & = & a_n \end{bmatrix} \text{ is of type: } \begin{bmatrix} l_1 & : & T_1 \\ l_2 & : & T_2 \\ \ldots & & \\ l_n & : & T_n \end{bmatrix}$$

iff $a_1 : T_1, a_2 : T_2, \ldots, a_n : T_n$

The type *SEvent* may be considered to be a record type. Its exact nature is a matter of the analysis you wish to propose. One candidate is (21), although one can argue that there are speech events which do not belong to this type.

(21) $\begin{bmatrix} \text{e-loc} & : & \textit{Loc} \\ \text{sp} & : & \textit{Ind} \\ \text{au} & : & \textit{Ind} \\ \text{e} & : & \textit{Phon} \\ c_{\text{loc}} & : & \text{loc(e,e-loc)} \\ c_{\text{sp}} & : & \text{speaker(e,sp)} \\ c_{\text{au}} & : & \text{audience(e,au)} \end{bmatrix}$

This means that a speech event would be modelled as a record of this type, that is, a record containing a spatial location labelled 'e-loc' ("event location"), two individuals ("speaker" and "audience") and an utterance event (labelled 'e') which is a phonological event, an event of type *Phon(ology)*. The remaining fields in the type place constraints on the relations that hold between these objects. Predicates like 'loc', 'speaker' and 'audience' are *type constructors*. When supplied with arguments they yield a type (normally a type of situation) which intuitively is to be thought of a providing a proof of a given fact. Thus, if e is an event and l is a location then loc(e, l)

---

[7] Note that r may contain more objects with labels not mentioned in the type.

is intuitively to be construed as a type of situation such that any situation of the type shows that e is located at l. In other words $s : \text{loc}(e, l)$ if and only if s constitutes a proof that the e's location is l. Before describing what a *sign* is and how utterances enter into context, we briefly describe the view of propositions we assume.

Building on a conception articulated 30 years earlier by Austin (Austin (1950)), Barwise and Etchemendy (1987) developed a theory of propositions in which a proposition is a structured object $prop(s, \sigma)$, individuated in terms of a situation s and a situation type σ.

TTR offers a straightforward way for us to model propositions using records. A proposition is a record of the form in (22a). The type of propositions is the record type (22b) and their truth can be defined, as in (22c):

(22)   a. $\begin{bmatrix} \text{sit} = r_0 \\ \text{sit-type} = p_0 \end{bmatrix}$

      b. $\text{Prop} = \begin{bmatrix} \text{sit} : \text{Record} \\ \text{sit-type} : \text{RecType} \end{bmatrix}$

      c.   A proposition $\begin{bmatrix} \text{sit} = r_0 \\ \text{sit-type} = p_0 \end{bmatrix}$ is true iff $r_0 : p_0$

In (23) we exemplify a sign or utterance event type in the format of the grammatical framework HPSG: this has fields for *phon*, (syntactic) category, contextual parameters, and content—the contextual parameters include minimally those discussed earlier in our explication of the type *SEvent* above:[8]

(23) $\begin{bmatrix} \text{PHON} : \texttt{is georges here} \\ \text{CAT} = \text{V[+fin,+root]} : \text{syncat} \\ \text{CONSTITS} = \left\{ \text{is, georges, here, is georges here} \right\} : \text{set(sign)} \\ \text{C-PARAMS} : \begin{bmatrix} \text{spkr: IND} \\ \text{addr: IND} \\ \text{c1 : address(s,a)} \\ \text{s0: SIT} \\ \text{l: LOC} \\ \text{g: IND} \\ \text{c3: Named(g,'georges')} \end{bmatrix} \\ \text{CONT} = \text{Ask(spkr,addr, ?} \begin{bmatrix} \text{sit} = \text{s0} \\ \text{sit-type} = \begin{bmatrix} \text{c : In(l,g)} \end{bmatrix} \end{bmatrix}) : \text{IllocProp} \end{bmatrix}$

---

[8] The convention is that a field $\begin{bmatrix} \ell = a : T \end{bmatrix}$ is used as a convenient notation for $\begin{bmatrix} \ell : T_a \end{bmatrix}$ where T_a is the type T restricted to a singleton type whose only witness is a. This is why the apparent asymmetry between the 'phon' and 'c-params' fields on the one hand, and the 'cat' and 'cont' fields is merely notational.

3.3 Grounding and Clarification Interaction conditions in KoS

The original conception of Austinian propositions was that s is a situation deictically indicated by a speaker making an assertion. Ginzburg (2012) extends this idea to the notion of *locutionary propositions* to characterize the communicative process, here the propositions individuated in terms of an utterance event u_0 as well as its grammatical type T_{u_0}. In terms of an interlocuter's dialogue gameboard, the issue can be formulated as follows: what information needs to be associated with *Pending*, that component of the dialogue gameboard which keeps track of utterances undergoing integration into the dialogue gameboard, to enable the formulation of grounding conditions/clarification interaction potential? The requisite information needs to be such that it enables the original speaker to interpret and recognize the coherence of the range of possible clarification queries that the original addressee might make. Ginzburg (2012) offers detailed arguments on this issue, including considerations of the phonological/syntactic parallelism exhibited between clarification requests and their antecedents and the existence of clarification requests whose function is to request repetition of (parts of) an utterance. Taken together with the obvious need for *Pending* to include values for the contextual parameters specified by the utterance type, Ginzburg concludes that the type of *Pending* combines tokens of the utterance, its parts, and of the constituents of the content with the utterance type associated with the utterance. An entity that fits this specification is the *locutionary proposition* defined by the utterance. In (24) we exemplify a locutionary proposition, whose utterance type we introduced earlier. The content of the utterance is given by the path *sit.cont*:

(24)

$$
\text{sit} = \begin{bmatrix}
\text{PHON} = & \text{izjorjhiya} \\
\text{CAT} = & \text{V[+fin,+root]} \\
\text{CONSTITS} = & \left\{ \begin{array}{l} \text{u1(iz),u2(jorj),} \\ \text{u3(hiya), u4(izjorjhiya)} \end{array} \right\} \\
\text{C-PARAMS} = & \begin{bmatrix} \text{spkr} & = & \text{A} \\ \text{addr} & = & \text{B} \\ \text{s0} & = & \text{sit1} \\ \text{l} & = & \text{l0} \\ \text{g} & = & \text{g0} \\ \text{c3} & = & \text{pr1} \end{bmatrix} \\
\text{cont} = & \text{Ask(A,B, ?} \begin{bmatrix} \text{sit} & = & \text{sit1} \\ \text{sit-type} & = & \left[\text{c : In(l0,g0)} \right] \end{bmatrix} \text{)}
\end{bmatrix}
$$

$$
\text{sit-type} = \begin{bmatrix}
\text{PHON : is georges here} \\
\text{CAT} = \text{V[+fin,+root] : syncat} \\
\text{CONSTITS} = \left\{ \text{is, georges, here, is georges here} \right\} : \text{set(sign)} \\
\text{C-PARAMS :} \begin{bmatrix} \text{spkr: IND} \\ \text{addr: IND} \\ \text{c1 : address(s,a)} \\ \text{s0: SIT} \\ \text{l: LOC} \\ \text{g: IND} \\ \text{c3: Named(g,'georges')} \end{bmatrix} \\
\text{CONT} = \text{Ask(spkr,addr, ?} \begin{bmatrix} \text{sit = s0} \\ \text{sit-type} = \left[\text{c : In(l,g)} \right] \end{bmatrix} \text{) : IllocProp}
\end{bmatrix}
$$

In the immediate aftermath of a speech event u, *Pending* gets updated with a record of type *locutionary proposition* (LocProp). Here T_u is a grammatical type for classifying u that emerges during the process of parsing u. Here we identify this with a *sign* in the sense of Head Driven Phrase Structure Grammar (HPSG), as we exemplified above in (23). The relationship between u and T_u—describable in terms of the proposition p_u given in (25a)—can be utilized in providing an analysis of grounding/CRification conditions, as shown in 25b:

(25)　a.　$p_u = \begin{bmatrix} \text{sit = u} \\ \text{sit-type} = T_u \end{bmatrix}$

　　　b.　Grounding/clarification interaction conditions:

　　　　　1. Grounding: p_u is true: the utterance type fully classifies the utterance token.

2. clarification interaction: p_u is false, either because T_u is weak (e.g., incomplete word recognition) or because u is incompletely specified (e.g., incomplete contextual resolution—problems with reference resolution or sense disambiguation).

It is useful to conceive of the integration of an utterance in an information state as a potentially cyclic process. Instantiation of some, perhaps all, contextual parameters will occur as soon as an utterance has taken place, assuming T_u is uniquely specified; if this is not the case, then clarification interaction can occur on that level. Parameter instantiation can also take place subsequently, as when more information is provided as a consequence of clarification interaction. Given this, utterance integration can be broken into three components:[9]

(26) a. **Pending update**: in the immediate aftermath of a speech event u, *Pending* gets updated with a record of the form $\begin{bmatrix} \text{sit} = \text{u} \\ \text{sit-type} = \text{T}_u \end{bmatrix}$

  b. **Contextual extension**: If T_u is uniquely specified, try to instantiate the contextual parameters of T_u relative to the context provided by the dialogue gameboard: find a record w that extends u and such that w contains a subrecord of the c-param anchoring intended by u's speaker; integrate w into *MaxPending*:
  $MaxPending := \begin{bmatrix} \text{sit} & = w \\ \text{sit-type} & = \text{T}_u \end{bmatrix}$

  c. **Move update/Pending downdate**: if *MaxPending* is true, update Moves, so that *LatestMove*:= *MaxPending*, downdate *MaxPending* from *Pending*.

4 Back to Quotation

The ontology we have developed for dialogue processing can now be utilized for analyzing the quotation data from earlier sections, with pure quotation denoting signs and direct quotations denoting locutionary propositions. Quotative predicates select for clauses restricted to denote either abstract entities *simpliciter* or locutionary propositions. We could formulate our formal account in a number of ways. One would involve type shifting—postulating mappings between utterance types triggered by quotation. Such an approach is highly consonant with the view of grammar developed as a dynamic resource available to an agent to create a language on the fly (Cooper and Ranta (2008); Cooper (2012)). Another possibility, somewhat more conservative in a way, and in line with Bonami and Godard (2008), is to appeal to one or more quotative *constructions*. This approach, emphasizing that quotation fits into a traditional view of grammar, is the format we adopt here, though it would be straightforward to adopt the former approach.

---

[9] *Moves* is the component of the dialogue gameboard where grounded utterances are represented.

4.1 Pure Quotation

The simplest construction we could associate with pure quotation is one in which it denotes the phonological type of its unquoted correlate, as in (27):

(27)   *pure-quot-ph* =
$\begin{bmatrix} \text{PHON : T.phon} \\ \text{CONT = T.phon : phoncat} \end{bmatrix}$

$\begin{bmatrix} \text{T : Sign} \\ \text{HEAD-DTR : T} \end{bmatrix}$

Note that the pure quotation type from (27) will lose the information about the original content of the sign. Moreover, we argued in section 2 that, in fact, a pure quotation utterance concerns potentially the entire information about the sign, not merely its phonological form. This can be easily fixed by using the construction type in (28), wherein the content of the quoted phrase is the whole sign-type, rather than that in (27). One additional aspect of this construction concerns its contextual parameters. We specify that these all get absorbed and the construction, hence, has no contextual parameters:

(28)   *pure-quot-ph* =
$\begin{bmatrix} \text{PHON: T.phon} \\ \text{C-PARAMS : []} \\ \text{CONT= T : RecType} \end{bmatrix}$

$\begin{bmatrix} \text{T : Sign} \\ \text{HEAD-DTR : T} \end{bmatrix}$

4.2 Direct Quotation

Recall that we proposed that in dialogue an utterance $u0$ is represented in an interlocutor's DGB by means of a locutionary proposition, a proposition whose situational component is the utterance event $u0$ and whose type component is a grammatical type T_{u0,Γ_0} assigned to the utterance by a given grammatical resource Γ_0. Not surprisingly, our analysis of direct quotation assigns to such clauses as their content a locutionary

proposition, with the situational component being the (original) utterance event and the type component a grammatical type T_{u0,Γ_0} assigned to the utterance by a grammatical resource associated with the person reporting the speech.

Whereas with pure quotation we associated solely the grammatical type, we include here also the utterance event which, in conjunction with the grammatical type, enables the utterance's content to be read off, as we explained in section 3.3. This is key both in capturing the similarity with indirect quotation (see data above examples (15)) and Partee–inspired data such as examples (2) above. Beyond this, though we emphasize that by postulating locutionary propositions as the denotations of direct quotation, we are offering a direct answer to the question we had posed in section 1, namely *Why is direct quotation so easy*? The answer is that this is because it exploits mechanisms in place for dialogue processing.[10]

By analyzing direct quotation complements as denoting *locutionary propositions*, which include as a component a sign we can capture simultaneously the type and token aspects of such complements. Indeed the fact that a single predication can address both type and token aspects, as in (29):

(29) a. 'Was I snoring' was asked by Bill and is a frequently used interrogative clause.

  b. 'Am I snoring?' asked Bill, a sentence frequently uttered by men who don't think they snore. It is usually answered by 'You were before you woke up.'

The construction we associate with this type is a fairly minor variant on the one we posited for pure quotation. The crucial difference is that although the contextual parameters of the original type have been projected away, a new one is projected to include the utterance event. To this we need to add constraints that relate the utterance event and the grammatical type appearing in the construction. As we discussed in section 2, what we need to enforce is *similarity* between the original utterance u and the type T used to quote it—we capture this in terms of a *similarity measure* μ and a set of grammatical resources Γ, the relationships between which we explicate below:[11,12]

---

[10] Of course there is a bit of a chicken/egg issue here, given that direct quotation is one of the mechanisms that gets language off the ground.

[11] In order to accommodate quotations of non-linguistic events, we would simply need to modify our characterization of the type of the head daughter as being grammatical to a somewhat wider class of types. Such types would certainly have a field for *phon* and might also have a field for *cont*.

[12] An anonymous reviewer for the *Journal of Logic, Language, and Information* asks whether by invoking similarity measures we are not rendering our proposal to be unfalsifiable. S/he writes 'Suppose that John gave a lecture, and somebody asks Peter about the lecture, how it was. Peter: John said: 'blablabla'. Can 'blablabla' stand for the content of John's lecture? If not, how is it blocked in the model?' On our account the truth of Peter's report requires an utterance of 'blablabla' to be similar to the sequence of utterances that made up John's lecture, relative to a contextually given similarity measure. A number of similarity measures we have mentioned above would not accommodate such similarity, e.g., one based on identity of content and *a fortiori* on identity of form. In fact, a measure of similarity that would make Peter's report true would need to be very coarse grained, essentially one that made all utterances of a given language similar, as for instance when one hears a lecture in a language one does not understand or when one wishes to convey that, for all intents and purposes what someone has said is contentless.

(30) a. *direct-quot-ph* =

$$
\begin{bmatrix}
\text{PHON} = \text{T.phon} : \text{phoncat} \\
\text{CAT} = \text{T.cat} : \text{syncat} \\
\text{C-PARAMS} : \begin{bmatrix} u : \text{Rec} \\ \mu : \text{SimilarityMeasure} \\ \Gamma : \text{set(GrammType)} \\ c1 : \text{Similar}(T, u, \mu, \Gamma) \end{bmatrix} \\
\text{CONT} = \begin{bmatrix} \text{sit} = u \\ \text{sit-type} = T \end{bmatrix} : \text{LocProp}
\end{bmatrix}
$$

$$
\begin{bmatrix}
T : \text{GrammType} \\
\text{HEAD-DTR} : T
\end{bmatrix}
$$

We turn to explicate briefly the relationship among the contextual parameters of a direct quotation:

(31) a. Let Γ be a set of grammatical resources. We will not specify this formally here. If u is a speech event and Γ a set of grammatical resources we use $\Gamma(u)$ to represent the sign type assigned to u by the Γ (ignoring the possibility that several alternative sign types may be provided for u).

b. A similarity measure on types is a pair of a threshold θ and a binary function μ_0 on types which returns a real number for any pair of types and whose maximum value is returned for any pair of identical types and for which any value greater than θ indicates similarity.[13]

c. If e is a speech event, T a grammatical type, μ a similarity measure and Γ a set of grammatical resources, the type

   $\text{Similar}(T, u, \mu, \Gamma)$

   is non-empty iff there exists a grammatical type T_1 such that $T_1 \sqsubseteq \Gamma_1(e_1)$ $e : T$, and $\mu_0(T_1, T) \geq \theta$

4.3 Quotative Predicates

One crucial point the data we discussed earlier establishes is that there are predicates ('go', 'be like', 'faire', 'typ') that select only for direct quotation complements. Another is that there are predicates that select simultaneously for direct and indirect quotation complements, with mutual entailments between corresponding com-

---

[13] Additional axioms need to be met but they need not concern us here.

plements (see (15 above).[14] We specify direct quotation selection via the semantic type we postulated for them—LocProp, whereas predicates selecting for both types will be specified via the abstract entity associated with an utterance of the direct complement—*question* ('ask'), *proposition* ('claim'), *outcome* ('order') etc.

We illustrate this with three lexical entries: one for a quotative predicate like 'be like', one for 'ask' taking a direct quotation complement, and one for an indirect quotation complement of 'ask'. In the two first cases the complement selected is a locutionary proposition, representing an utterance whose speaker is identical to the entity associated with the subject of the predicate. The *cont* value of the predicate is identified with the value of the utterance component of the locutionary proposition. For 'be like' this can be any kind of illocutionary act,[15] whereas for direct 'ask' this has to be a query. In the final indirect case of 'ask', one simply unifies in the question–denoting complement daughter and the referent of the subject daughter:

(32) a.
$$
\begin{bmatrix}
\text{PHON} : \textit{be like} \\
\text{CAT} = \text{v[+fin]} : \text{PoS} \\
\text{SPR-DTR} : np \land \begin{bmatrix} \text{content} = x : \text{Ind} \end{bmatrix} \\
\text{COMP-DTR} : \begin{bmatrix} \text{CONT} = p : \text{LocProp} \\ p.\text{sit-type.ctxt-params.spkr} = x : \text{IND} \\ i = p.\text{sit.cont} : \text{IllocProp} \end{bmatrix} \\
\text{CONT} = i : \text{IllocProp}
\end{bmatrix}
$$

b.
$$
\begin{bmatrix}
\text{phon} : \textit{ask} \\
\text{CAT} = \text{v[+fin]} : \text{PoS} \\
\text{SPR-DTR} : np \land \begin{bmatrix} \text{content} = x : \text{Ind} \end{bmatrix} \\
\text{COMP-DTR} : \begin{bmatrix} \text{CONT} = p : \text{LocProp} \\ p.\text{sit-type.ctxt-params.spkr} = x : \text{IND} \\ q : \text{Question} \\ i = \text{Ask}(x,q) : \text{IllocProp} \end{bmatrix} \\
\text{CONT} = i : \text{IllocProp}
\end{bmatrix}
$$

---

[14] And to this one should add, as noted by Bonami and Godard (2008), the existence of predicates *incompatible* with direct quotation complements. But we will not attempt to capture this pattern here, which extends to verbs of cognition—an analogy much emphasized by Recanati.

(i) Bill thought to himself 'What a guy I am'.

(ii) #Bill knew 'What a guy I am'. (cf. We all know what an impressive candidate Bo is.)

[15] And we could extend this somewhat to allow for speech events like, say, grunts that might not be illocutionary in a strict sense.

c.
$$\begin{bmatrix} \text{phon} : \textit{ask} \\ \text{CAT} = \text{v[+fin]} : \text{PoS} \\ \text{SPR-DTR} : \textit{np} \wedge \begin{bmatrix} \text{content} = \text{x} : \text{Ind} \end{bmatrix} \\ \text{COMP-DTR} : \begin{bmatrix} \text{CONT}= \text{q} : \text{Question} \end{bmatrix} \\ \text{CONT} = \text{Ask(x,q)} : \text{Prop} \end{bmatrix}$$

5 Previous Formal Grammatical Literature on Quotation

In this section we discuss some recent formal grammatical accounts of quotation whose insights we draw on.

Bonami and Godard We have already mentioned Bonami and Godard's paper (Bonami and Godard (2008)) several times. This paper makes several important contributions. Descriptively, it highlights the canonical nature of quotative complements, thereby providing arguments for their within–grammar nature; it points out the need for the grammar to accommodate 'agrammatical' quotative complements. Formally, it offers an account for a quotative complement as a sign–denoting construction. And it points to intrinsic difficulties that standard typed feature structure based grammar, of which their account is self-consciously one, has with quotation. One fundamental problem faced by Bonami and Godard's account involves direct quotation, their intended quarry: whereas a sign–based construction is desirable for dealing with pure quotation, it cannot work for direct quotation since what one needs there in order to evaluate similarity is, in addition, the (source) utterance *token*. But this latter is not made available in standard typed feature structure based grammar.

Potts In the concluding section of his paper, Potts (2007) writes (*a propos* of a remark by Bart Geurts) that 'Quotation is a hugely important matter for linguistic theory. It forces us to enrich our stock of basic entities.' We agree with the first sentence, but disagree with the second sentence, albeit recognizing its subjectivity. Potts introduces a standard grammar which he enriches with a type for expressions whose denotations he identifies with triples of phonological types, syntactic types, and logical expressions. These *ad hoc entities* are used as denotations for pure quotation. In contrast, our approach simply reuses entities that are assumed to exist for independent reasons, namely signs (utterance types). For dealing with direct quotation Potts introduces a two dimensional semantics involving the regular denotations and an additional dimension dubbed the 'speech report dimension'. Our account of direct quotation utilizes locutionary propositions, independently introduced as the contextual representation of utterances.

Geurts and Maier The general framework in which Geurts and Maier(GM) work in is the 'binding theory' of presupposition (BTP, see e.g., Geurts (1999)), an offshoot

of Discourse Representation Theory (Kamp and Reyle (1993)).[16] The general intention is to bring to bear BTP tools to capture the context dependency of a quotation analogously to anaphoric pronouns, definite noun phrases, and other presupposition-inducing expressions.

On GM's account, the semantic effect of quotation is to shift the ordinary meaning of an expression α to a new meaning that involves mentioning α. If the quotation is a 'pure' one, it denotes a linguistic expression and its semantic type may differ from that of the quoted expression. For example, in (33) quotation maps the semantic type of the adjective 'monosyllabic' to that of singular terms, and the resulting expression denotes the word 'monosyllabic':

(33)   'Monosyllabic' is not monosyllabic. (GM's (31))

In other cases, e.g. mixed and direct quotation, the semantic type of the quotation is identical to that of the quoted expression. For example, on the GM analysis, (34a) denotes the DRS in (34b), with the crucial innovation the postulation of the discourse referent Q which denotes (or purports to denote) the same kind of object the word 'police' would have denoted. The semantic content expressed by the quotation involves reference to the expression 'POlice'; it might be, e.g., 'whatever it is Tony meant when he used the expression 'POlice''. In this sense 'POlice' is used and mentioned at the same time: the mentioning becomes part of the expression's content;

(34)   a. Tony says the 'POlice' are on their way. (GM's (42))
       b. [y: Tony(y), e: say_e(y,[x e' Q: $E_{e'}$(x,Q,POlice),
           z: Q(z), e": onTheirWay$_{e''}$(z)])] (GM's (43))
       c. [y: Tony(y), e Q: $E_{e'}$(x,Q,POlice), z: Q(z),
           say_e(y,[x [e": onTheirWay$_{e''}$(z)]])] (GM's (44))

On GM's account there are at least two ways in which mixed quotations depend on the context. A mixed quotation always has a source, which has two coordinates: a speaker and a speech event. The speaker coordinate may be a concrete individual or it may be institutional or generic. Similarly, the speech-event coordinate may be a concrete event or it may be habitual or generic. (34b) illustrates this in that x, e' are discourse referents that need to be bound via presuppositional binding—in this case (34c) emerges with x identified with Tony and e' with the saying event. Hence, (34a) gets a reading paraphrasable as *there is an occasion e on which Tony used the expression 'POlice' to express the property Q, and that on the same occasion Tony said that the Q are on their way*. But for a variety of other cases, discussed by GM, different binding possibilities emerge, mediated by binding operators such as quantification, negation, and focus.

In contrast to Potts' two dimensional account, which GM critique, their own account is one dimensional and captures elegantly and precisely various interactions between variable/presupposition binding and quotation, which in a larger study we would try to integrate into our account.

---

[16] For a more detailed discussion of GM's approach, also developed in Maier (2007) and subsequently, see De Brabanter (2010).

Nonetheless, there are a number of intrinsic shortcomings. For a start, given the unrelated ways of dealing with pure and direct quotation, it is not clear how copredication cases like (29) can be handled. An additional problem, relating to the framework in which the account is couched concerns selection—there is no obvious way in DRT and its close variants to specify the type which a direct quotation predicate such as 'be like' selects for.[17]. Similarly, the account of pure quotation appeals to denotations as expressions, whereas in fact we have argued that the requisite denotations need to be signs. With respect to direct quotation, given the lack of a rich type theory, it cannot formulate similarity as a condition relating source and quotation, as we do here. Nor can the account deal with diagonalization, as we do here by postulating grammatical resources as contextual parameters.

6 Conclusions

Direct and pure quotation have often been viewed as mysterious, outside the pale in most formal grammars or requiring unorthodox semantic/logical treatment, despite the fact that direct quotation is typologically universal and ontogenetically prior to indirect quotation. We have suggested that once one assumes a dialogical perspective on language such as provided by KoS, formalized in a rich type theory like TTR, much of the mystery evaporates. In particular, one can utilize as denotations for quotative constructions entities that are independently motivated for dialogue processing—utterance types and locutionary propositions, austinian propositions about speech events. The "diagonalizing property" of quotation, its ability to accommodate the unwashed and the alien, viz. ungrammatical and foreign utterances, is unproblematic on a view of language in flux, where there is no overarching notion of grammar, but recourse is made to multiple grammatical resources, depending on the purpose and setting. Our account of direct quotation invoking similarity and grammatical resources as contextual parameters provides the basis for an account of fluctuating standards of faithfulness between quotative and original utterance.

What we have provided here is of course a sketch: a detailed grammar needs to be developed and extended to mixed and scare quotation. We hope we have at least motivated the desirability of a more detailed picture in these terms.

---

[17] See Asher (2011) for such a type theoretic framework, though it is not integrated with the discourse semantics, in contrast to TTR/KoS.

References

Asher N (2011) Lexical meaning in context: A web of words. Cambridge University Press

Austin J (1962) How to Do Things with Words. Oxford University Press, ed. by J. O. Urmson

Austin JL (1950) Truth. In: Urmson J, Warnock GJ (eds) Philosophical Papers, Oxford University Press

Barwise J, Etchemendy J (1987) The Liar. Oxford University Press, New York

Barwise J, Perry J (1983) Situations and Attitudes. Bradford Books, MIT Press, Cambridge, Mass.

Bonami O, Godard D (2008) On the syntax of direct quotation in French. In: Proceedings of the 15th international conference on head-driven phrase structure grammar, pp 358–377

Brennan SE, Clark HH (1996) Conceptual pacts and lexical choice in conversation. Journal of Experimental Psychology: Learning, Memory and Cognition 22:482–493

Cappelen H, Lepore E (2012) Quotation. Stanford Encyclopedia of Philosophy

Clark H (1996) Using Language. Cambridge University Press, Cambridge

Clark H, Gerrig R (1990) Quotations as demonstrations. Language 66(4):764–805

Cooper R (2012) Type theory and semantics in flux. In: Kempson R, Asher N, Fernando T (eds) Handbook of the Philosophy of Science, vol 14: Philosophy of Linguistics, Elsevier, Amsterdam

Cooper R (2013a) Clarification and generalized quantifiers. Dialogue and Discourse 4:1–25

Cooper R (2013b) Type theory, interaction and the perception of linguistic and musical events. In: Orwin M, Howes C, Kempson R (eds) Language, Music and Interaction, Communication, Mind and Language, vol 3, College Publications

Cooper R, Larsson S (2009) Compositional and ontological semantics in learning from corrective feedback and explicit definition. In: Edlund J, Gustafson J, Hjalmarsson A, Skantze G (eds) Proceedings of DiaHolmia: 2009 Workshop on the Semantics and Pragmatics of Dialogue, Department of Speech, Music and Hearing, KTH, pp 59–66

Cooper R, Ranta A (2008) Natural Languages as Collections of Resources. In: Cooper R, Kempson R (eds) Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution, Communication, Mind and Language, vol 1, College Publications, London, pp 109–120

de Cornulier B (1978) L'incise, la classe des verbes parenthétiques et le signe mimique. Cahier de linguistique (8):53–95

Davidson D (1979) Quotation. Theory and Decision 11(1):27–40

De Brabanter P (2010) The semantics and pragmatics of hybrid quotations. Language and Linguistics Compass 4(2):107–120

Everett D (2012) Quotatives in wari' and piraha: different strategies for reporting the contents of other minds. Slides from Bochum Quotation Workshop

Geurts B (1999) Presuppositions and Pronouns. Elsevier

Geurts B, Maier E (2005) Quotation in context. Belgian Journal of Linguistics

   17:109–28

Ginzburg J (2012) The Interactive Stance: Meaning for Conversation. Oxford University Press, Oxford

Ginzburg J, Cooper R (2004) Clarification, ellipsis, and the nature of contextual updates. Linguistics and Philosophy 27(3):297–366

Ginzburg J, Moradlou S (2013) The earliest utterances in dialogue: towards a formal theory of parent/child talk in interaction. In: Fernández R, Isard A (eds) Proceedings of SemDial 2013 (DialDam), University of Amsterdam

Healey P (1997) Expertise or expertese?: The emergence of task-oriented sublanguages. In: Shafto M, Langley P (eds) Proceedings of the 19th Annual Conference of the Cognitive Science Society, pp 301–306

Healey PGT, Swoboda N, Umata I, King J (2007) Graphical Language Games: Interactional Constraints on Representational Form. Cognitive Science 31(2):285–309

Kamp H, Reyle U (1993) From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory. No. 42 in Studies in Linguistics and Philosophy, Kluwer, Dordrecht

Larsson S, Cooper R (2009) Towards a formal view of corrective feedback. In: Alishahi A, Poibeau T, Villavicencio A (eds) Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition, EACL, pp 1–9

Maier E (2007) Mixed quotation: between use and mention. Proceedings of LENLS

Martin-Löf P (1984) Intuitionistic Type Theory. Bibliopolis, Naples

Montague R (1974) Formal Philosophy: Selected Papers of Richard Montague. Yale University Press, New Haven, ed. and with an introduction by Richmond H. Thomason

Nordström B, Petersson K, Smith JM (1990) Programming in Martin-Löf's Type Theory, International Series of Monographs on Computer Science, vol 7. Clarendon Press, Oxford

Pagin P, Westerståhl D (2010) Pure quotation and general compositionality. Linguistics and Philosophy 33:381–415

Partee B (1973) The syntax and semantics of quotation. A festschrift for Morris Halle pp 410–418

Postal PM (2004) Skeptical linguistic essays. Oxford University Press Oxford

Potts C (2007) The dimensions of quotation. In: Jacobson P, Barker C (eds) Direct Compositionality, pp 405–431

Purver M (2006) Clarie: Handling clarification requests in a dialogue system. Research on Language and Computation

Purver M, Ginzburg J, Healey P (2001) On the means for clarification in dialogue. In: van Kuppevelt J, Smith R (eds) Current and New Directions in Discourse and Dialogue, Kluwer, pp 235–256

Quine WVO (1960) Word and object. MIT press, Cambridge, Mass

Recanati F (2010) Truth-Conditional Pragmatics. Clarendon Press Oxford

Rodriguez K, Schlangen D (2004) Form, intonation and function of clarification requests in german task-oriented spoken dialogues. In: Ginzburg J, Vallduvi E (eds) Proceedings of Catalog'04, The 8th Workshop on the Semantics and Pragmatics of Dialogue, Universitat Pompeu Fabra, Barcelona

Sag IA, Wasow T, Bender EM (2003) Syntactic Theory: A Formal Introduc-

tion, 2nd edn. CSLI Publications, Stanford, URL `http://cslipublications.stanford.edu/site/1575864002.html`

de Saussure F (1916) Cours de linguistique générale. Payot, Lausanne and Paris, edited by Charles Bally and Albert Séchehaye

Searle JR (1969) Speech Acts: an Essay in the Philosophy of Language. Cambridge University Press

Sperber D, Wilson D (1986) Relevance: Communication and Cognition. Blackwell

Werning M (2012) Making quotation transparent. Slides from Bochum Quotation Workshop