

In Peter Bosch, David Gabelaia, & Jerome Lang (eds.) Logic, Language, and Computation. 7th International Tbilisi Symposium on Logic, Language, and Computation, Tbilisi, Georgia, Oct. 2007, Revised Selected Papers. Springer. Lecture Notes in Artificial Intelligence. 2009

## Processing Definite Determiners. Formal Semantics Meets Experimental Results

Peter Bosch

Institute of Cognitive Science  
University of Osnabrück, D-49069 Osnabrück, Germany  
pbosch@uos.de

**Abstract** Experiments on the online processing of linguistic utterances provide information about language processing in the first instance, and only indirectly about linguistic knowledge, while it has been linguistic knowledge, and not linguistic processing, that has been the subject matter of theoretical linguistics. So how can such evidence be relevant to theoretical linguistics? Or how can linguistic theory inform a theory of language processing? – This issue is discussed here with respect to the processing and the formal semantics of the English definite determiner. I argue that the meaning of the definite determiner, as it shows up in experiments on online comprehension, can actually be accounted for in an incremental variant of current formal semantics.

**Keywords:** definite determiner, domain restrictions, formal semantics, incremental processing, psycholinguistics, eye-tracking,

### 1 Linguistic knowledge and language processing

Theoretical linguistics investigates the native speaker's implicit *knowledge of the language*. This includes not only phonology, morphology, syntax, and compositional semantics, but also the systematic and equally implicit knowledge the speaker has of *using* linguistic expressions appropriately in a context or situation. Linguistic knowledge in this sense is characterized in abstract algebraic terms, very much as in the theory of formal languages. Questions relating to how linguistic knowledge is *implemented* in human behaviour or in the human brain, on the other hand, are not part of theoretical linguistics, but belong to a theory of linguistic processing, i.e., to neurolinguistics or psycholinguistics.

What, then, is the relation between linguistic knowledge and linguistic processing? I don't think that we have a very good answer yet. But one fairly simple way of relating linguistic knowledge to linguistic processing is in assuming that the linguistic knowledge provides *constraints* on processes of comprehension and production. In some cases, all other things being equal, these constraints may provide for empirical predictions regarding the time course and output of linguistic processing. Linguistic theory thus may inform a theory of language processing, and observations

about language processing may inform linguistic theory, i.e., support or disconfirm its predictions.

In the following I will look at some experimental results concerning the processing of definite determiners and attempt to relate these results to what theoretical linguistics has to say about the definite determiner. In Section 3, I will discuss a proposal for how current semantic formalism may be used to describe incremental comprehension processes. Also here the issue is the meaning and processing of the definite determiner.

## **2 Some findings from experimental work: Immediacy, Incrementality, and Crossmodality**

In this section I want to review some experimental results about linguistic comprehension and, on the side, as it were, introduce three processing properties that are now widely accepted as properties of human natural language comprehension and that constitute a certain challenge to linguistic theory: immediacy, incrementality, and crossmodality.

By *immediacy* I mean the observation that all information that becomes available to the language processor is in principle used immediately in the comprehension process and thus may show immediate effects. By *incrementality* I mean the observation that information that has already been acquired in the ongoing comprehension process is used to control later processing steps. And by *crossmodality* I mean the observation that linguistic processing happens in tandem with non-linguistic processing steps, which may guide or be guided by the linguistic processing. These ideas about linguistic processing have entered psycholinguistics in the 1980s (cf. Marslen-Wilson & Tyler 1980, Garfield 1987, Altmann & Steedman 1988) and are now shared by a broad community of researchers (cf. Trueswell & Tanenhaus 2005, Hagoort & van Berkum 2007).

### **2.1 Visual World**

The experimental work I want to focus on is close to what has become known as the *Visual World* paradigm (Tanenhaus e.a. 1995). Experimental subjects wear a device on their heads that makes a video recording, showing exactly what they are looking at. They listen to spoken instructions or stories and carry out various tasks. The eye-tracking provides evidence of the cognitive activity of subjects that can be correlated with the linguistic input. A methodologically essential point is that eye movement in this set-up is spontaneous and not under the subject's conscious control. Subjects' reflection or intuition on meaning do not interfere.

In a typical Visual World experiment about lexical access for instance (e.g., Allopenna e.a. 1998) participants view a panel with four drawings of simple objects, such as, for instance, a beetle, a beaker, a speaker, and a dolphin, and they listen to instructions as in (1).

- (1) Pick up the {beetle, beaker, speaker, dolphin}.

Participants focus the target object more frequently than its competitors already before the target noun is completed, unless the word is similar to a word that would be the default name for one of the competing objects in the display. For instance., when the instruction uses the word *beetle* and both a beetle and a beaker are present, participants' focussing frequency for both objects is initially similar and the difference starts to show only after the word's offset. In brief: focussing reacts as soon as there is enough information.

We conducted an experiment in a slightly more natural setting, where subjects were not choosing from a small set of separately depicted referents, but where they listened to short stories while they were viewing a picture of a related semi-natural scene (Karabanov e.a. 2007). While subjects heard a story as in (2), they would be shown a picture as in Figure 1.

- (2) Heute ist Markt im Dorf. Die Marktfrau streitet sich mit dem Arbeiter. Sie sagt jetzt gerade, dass er ihr nun das neue Fahrrad zurückgeben soll, das er sich geliehen hat.  
'It's market day in the village. The market woman is quibbling with the worker. She's just saying that he should give the new bike back that he has borrowed.'

Subjects' focussing probability for the referents of the referential expressions in the story (full lexical NPs as well as pronouns in our example) regularly started to increase immediately, already while the referential expression was heard, and reached a peak within about one second.



**Figure 1** Display from Karabanov e.a. (2007)

This is not what linguistic theory would make us expect. If we took theoretical linguistic accounts straightforwardly as accounts of processing, what should happen would be rather something like a purely linguistic bottom-up process: lexical access would be performed word by word and the morphological, syntactic, and semantic properties of each lexical item would be retrieved from the mental lexicon. Once the end of a sentence is reached, parsing could start and the string of expressions would be assigned a constituent structure, which would determine, among other

things, which expressions may count as referential constituents, so that on the basis of their lexical meaning and anaphoric status subjects could then assign to these expressions referents in the picture. The last step would then be reflected in the focussing frequency for these referents.

What we find experimentally is quite different: subjects process incoming referential expressions immediately with regard to the referential domain, here the visual domain. This also holds for pronouns, which are referentially resolved immediately when they are heard (cf. Eisenband e.a. 2000). This requires not only that the linguistic information is used as soon as it is available, but also that the visual input, and possibly other relevant sources of knowledge, are used immediately in the comprehension process.

The various sources of non-linguistic and contextual knowledge, together with what has already been understood, may occasionally be strong enough to predict referents of expressions that have not yet occurred in the acoustic input – just as in everyday life where we are often able to complete a sentence that somebody else started. In Example (2) for instance we found that the focussing probability for the worker in the display started to increase already at the time the word *streitet* was heard – presumably, because subjects "understand" at this point that the market woman is said to quibble with somebody; and since there is only one other person in the picture, the expected prepositional object expression is likely to refer to him. The German verb *streitet* 'quibbles' is commonly used in constructions like (*sich*) *mit jemandem (über etwas) streiten* 'to quibble with somebody (over something)', where the verb may either be construed reflexively or not. In either case linguistic knowledge predicts a referential expression for the person to quibble with to occur soon after the verb. The early increase of focussing on the worker would indicate that already when the verb is understood the corresponding interpretation is expected. These observations about anticipation are well supported by earlier experiments by Altmann & Kamide (1999) and Boland (2005).

## 2.2 Determiner gender

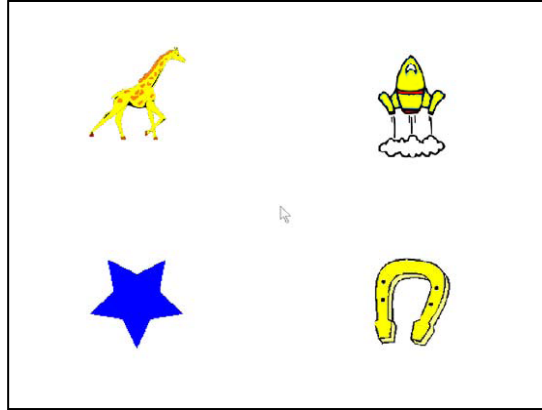
The example just discussed shows that at least some grammatical and semantic information from the lexicon is used immediately in conjunction with visual information to predict referential interpretations for expressions still to come. But one may wonder if this applies also to more abstract grammatical features, such as gender or number, or definiteness. We will look at an experiment about determiner gender first.

### 2.2.1 German determiner gender

Hartmann (2004) investigated the role of determiner gender in a Visual World experiment. Subjects were viewing displays as in Figure 2 while they heard instructions as in (3).

- (3) Klicken Sie auf { das<sub>[neut]</sub> gelbe<sub>[neut]</sub> Hufeisen<sub>[neut]</sub> / die<sub>[fem]</sub> gelbe<sub>[fem]</sub>  
Giraffe<sub>[fem]</sub> / die<sub>[fem]</sub> gelbe<sub>[fem]</sub> Rakete<sub>[fem]</sub> / den<sub>[masc]</sub> blauen<sub>[masc]</sub>  
Stern<sub>[masc]</sub> }

'Click on {the yellow horseshoe / the yellow giraffe / the yellow rocket / the blue star}'



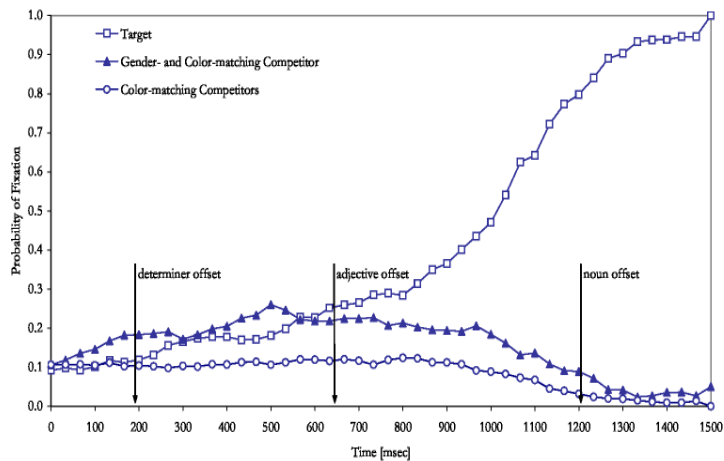
**Figure 2** Display with a yellow giraffe, a yellow rocket, a blue star, and a yellow horseshoe

German nouns come in one of three genders, masculine, feminine, or neuter, which represent a purely formal and semantically unpredictable noun classification, except when the genders correspond to the natural gender of persons. Determiners as well as adjectives agree in gender with their noun heads. Thus in instructions like (3), presented with Figure 2, already the determiner would exclude some of the objects as potential referents for the unfolding Determiner Phrase (DP), and the adjective with the additional information on colour would lead to further exclusions. For the four instructions in (3) the referential options would hypothetically develop as follows:

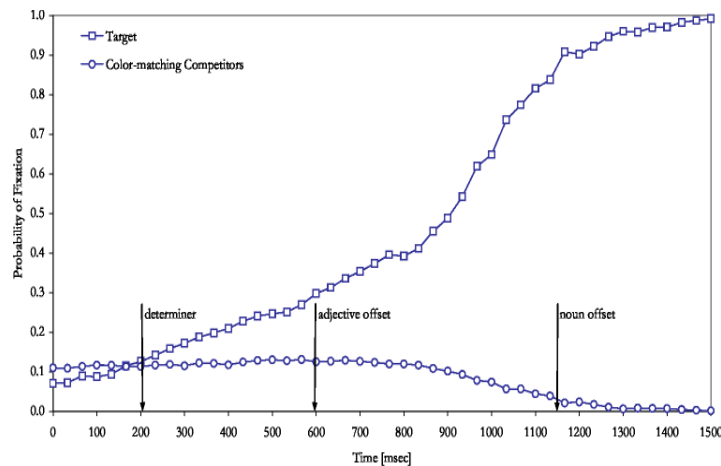
- (a) Klicken Sie auf die<sub>[fem]</sub> gelbe<sub>[fem]</sub> Giraffe<sub>[fem]</sub>  
*die* – excludes star and horseshoe; *gelbe* – leads to no further exclusions;  
*Giraffe* – establishes the referent.
- (b) Klicken Sie auf die<sub>[fem]</sub> gelbe<sub>[fem]</sub> Rakete<sub>[fem]</sub>  
*die* – excludes star and horseshoe; *gelbe* – leads to no further exclusions;  
*Rakete* – establishes the referent.
- (c) Klicken Sie auf den<sub>[masc]</sub> blauen<sub>[masc]</sub> Stern<sub>[masc]</sub>  
*den* – excludes all competitors; unique reference is established before the  
adjective or the noun have occurred
- (d) Klicken Sie auf das<sub>[neut]</sub> gelbe<sub>[neut]</sub> Hufeisen<sub>[neut]</sub>  
*das* – excludes all competitors; unique reference is established before the  
adjective or the noun have occurred

Hartmann's results support this hypothetical selection process. Figure 3 shows the focussing probabilities over time for a condition in which there was one competitor with the same colour and gender as the target, plus two other competitors that are not of the same gender but of the same colour as the target (the graph shows the average focussings for the latter two referents). There is no significant difference in focussing probability for different referents during the hearing of the determiner. This would in fact not even be possible, because it takes about 200 ms to initiate a

saccade and the determiners took just under 200 ms. Very soon after the determiner offset, and before the adjective could have had any influence on the saccades, at around 300 ms from the determiner onset (and 100 ms after adjective onset) we can observe a clear increase in focussing frequency for both the target and the gender-congruent referent, not though for the other competitors. The final differentiation, between the target on the one hand, and the gender- and colour-matching competitor on the other, becomes significant in the focussing probabilities only after the noun is recognized, at about 200 ms after the noun onset.



**Figure 3** Focussing probability over time for *den blauen Stern* in view of a display with one gender- and colour-matching competitor (e.g., *den blauen Hut*) and two colour-matching competitors (e.g., *die blaue Rakete*, *das blaue Hufeisen*)



**Figure 4** Focussing probability over time for *den blauen Stern* when the display shows no gender-matching competitors (e.g., *die blaue Rakete*, *das blaue Hufeisen*, *die blaue Giraffe*)

In a condition where there are no gender-matching competitors (like (c) and (d) above) subjects seem to have decided on the target already at the

time of processing the determiner. A difference in focussing frequency shows up around the time of determiner offset Figure 4<sup>1</sup>.

### 2.2.2 *Determiner meaning*

What do these experiments tell us about the meaning of the determiner? It may seem odd to link the gender feature to questions of meaning. After all German gender has no semantic content and merely reflects an abstract classification of nouns. But we saw that the determiner may select a referent for the DP already when no other part of the DP is known, and referent selection is a matter of semantics, if anything is. Still, linguistic theory would not attribute a referential function to determiners. What the experiment shows, however, is that – given suitable conditions – determiners can sometimes do the full job that referential expressions do.

This points to a gap in the theory: as long as we cannot say what those "suitable conditions" are, we cannot link our linguistic generalization to our observations. A linguistic theory that bluntly says that determiners are not referential expressions is obviously wrong in view of what we saw in Hartmann's experiment. – What are we to do? Withdraw our statement about referential expressions and determiners?

Let us think about where the gap is that needs to be bridged. Frege models the denotation of the determiner as a partial function, defined for a domain in which the NP denotes a singleton. In a modern formulation from Heim & Kratzer (1998:75,85) this read as in (4):

$$(4) \quad \lambda f: f \in D_{\langle e, t \rangle} \& \exists! [x f(x)]. \iota y [f(y)=1] \langle\langle e, t \rangle, e \rangle$$

In plain English: for any predicate  $f$  that denotes a singleton, the unique thing  $y$  that is  $f$ . The idea of making the definition of the function depend on properties of the domain to which it is applied, i.e., letting the use of the determiner presuppose that these domain properties are satisfied, opens up the option of taking context into account. Accordingly, the notion of referentiality can be relativized to assumptions about the context. In (4) the only assumption about the context is in the domain condition, i.e., that the denotation of  $f$  is a singleton. We shall see later that the use of domain conditions can be broadened and may provide one way of doing something about the gap between theory and observation. We shall return to this point.

---

<sup>1</sup> Dahan e.a. (2000) investigated the same question as Hartmann for French definite determiners, but used instructions with nouns following directly their determiners. They could establish clear evidence for gender priming by showing that nouns that were initially phonologically identical with the target noun but of a different gender were not activated. But they did not find a difference in focussing frequency between phonologically unrelated nouns of different gender, at least not within the first 200 ms after noun onset. At the time when the difference becomes significant, however, the effect may be attributed either to the preceding determiner or to the noun itself. Hartmann's experiment resolved this indeterminacy by inserting between the determiner and the noun an adjective that adds no disambiguating information but pulls the effect of the determiner and the noun apart.

For the moment let's just suppose that (4) is the meaning of the definite determiner and that it is actually part of the determiner's entry in the mental lexicon. As long as  $f$ , i.e. the denotation of the nominal, is not known, the determiner thus *could not* have any processing effect. Hartmann's experiment seems to demonstrate the contrary. – But then one of our assumptions may be wrong: who says that the denotation of the following nominal is not known? It is clearly not known from the linguistic input at this time, but utterance comprehension is *crossmodal* and takes in also the visual information. The instruction *Klicken Sie auf...* 'click on...', uttered in view of a display as in Figure 2, limits the choice of referents for the following determiner phrase to four objects, and linguistic experience provides default nouns (i.e. basic level common nouns that are most frequently used for naming the relevant objects) for these objects. For a display as in Figure 2 and an instruction like (c) or (d) only one of these nouns happens to be gender-congruent with the determiner. – So the knowledge active in the processor after hearing the determiner *den* is this:

- (i) an entry in the mental lexicon that makes the noun *Stern*<sub>[masc]</sub> the default description for  $\llbracket \text{star} \rrbracket$  instances,
- (ii) the identification of exactly one display object as an instance of  $\llbracket \text{star} \rrbracket$ ,
- (iii) the subsumption of the remaining display objects under concepts with non-gender-congruent default nouns, such as *Giraffe*<sub>[fem]</sub>, *Rakete*<sub>[fem]</sub>, and *Hufeisen*<sub>[neut]</sub>
- (iv) a lexical entry for the denotation  $\llbracket \text{den} \rrbracket$  that also includes gender information:  $\lambda f: f \in D \langle e, t \rangle \& \exists ! x (f(x) \& g(x, \text{masc})) . \iota y f(y)$ <sup>2</sup>

None of (i) – (iii) is strictly speaking *linguistic knowledge*. (ii) and (iii) represent purely contingent contextual knowledge, and (i) is knowledge about preferences of lexical usage in a language community, i.e., with no change in the semantics of *Stern*, there could still be another noun that German speakers would prefer, or that comes to their minds first, when they recognize the object in the display. So the only piece of linguistic knowledge proper that is involved in the effect of determiner gender on referent identification is the lexical entry in (iv).

There is still a certain complication that comes from the syncretism of German determiner forms. The form *den* in its function as a determiner, is ambiguous between four grammatical words, with any of the following features values: [num:sg; gend:m; case:acc], [num:pl; gend:n; case:dat], [num:pl; gend:m; case:dat], [num:pl; gend:f; case:dat]. In addition we have its use as a demonstrative pronoun and as a relative pronoun; in either case the form is unambiguous: [num:sg; gend:m; case:acc]. Since in the regular understanding of *Klicken Sie auf...* only the accusative would be permitted in the continuation and we are left with a three-way ambiguity between determiner, demonstrative pronoun, and relative pronoun. Since also the

---

<sup>2</sup> It would be a little sloppy to attribute gender to anything other than a linguistic expression. So let's take  $g$  as a function that assigns to a thing,  $x$ , the gender of the noun that is preferentially used for a description of  $x$ . But note that this brings in *linguistic experience*: The function  $g$  is given not as part of *knowledge of the language* but can be acquired only from linguistic experience.

relative pronoun is unlikely in this position, the realistic choice is between definite determiner and demonstrative pronoun. Since the features relevant for referent selection are identical in both cases, i.e., [num:sg; gend:m], the ambiguity that is left after applying all constraints that we find in linguistic knowledge does not affect processing.

In the case of the neuter accusative determiner *das* the syncretism is again limited to the ambiguity between the singular determiner, the singular demonstrative and the singular relative pronoun, so that we have the same situation as with *den*.

In the third case, the case of the accusative determiner *die*, the situation seems worse to start with, *die* boosting 12 grammatical words only in its accusative "reading". Nine of these are plural forms, however, and as such they are made less likely by the display properties: there is no plurality of objects in the display, at least not at the conceptual level of basic level common noun denotations. Still, in principle, we could have the unlikely instruction *Klicken Sie auf die Objekte im Bild, eins nach dem anderen* 'click on the objects in the display, one after the other'. Ignoring this option, we are down to the same three-way ambiguity again as already with *den* and *das*.

The point I've been trying to make is simple: in order to explain the observed focussing behaviour we require assumptions about *linguistic knowledge*, in the case at hand morphological, lexical, syntactic, and semantic knowledge, *plus* knowledge from *linguistic experience*, like naming preferences and the knowledge of basic level common nouns, *plus* knowledge of the *reference situation*, here the visual display. The formulation for the determiner denotation in (iii) includes the relevant linguistic knowledge, and with the function *g* it also includes a way of taking linguistic experience for the relevant domain into account. Still: (iii) is a representation of linguistic knowledge and is valid for the determiner *den* in any context, *pace* the ambiguity caused by its syncretism.

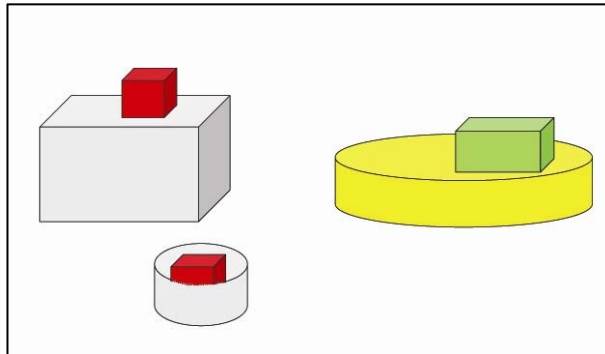
### 3 Definiteness

Let us suppose that our rough model for the role of gender information in the definite determiner, as part of the domain condition of the determiner, is correct in the sense that it helps to isolate the contribution of linguistic knowledge to the effect that Hartmann observed. Would the role of the *definiteness* feature in the determiner condition do an equally useful job in explaining how *this* feature contributes to the processing of definite DPs?

Suppose subjects are viewing a display as in Figure 5 while they hear an instruction like (5), with a prosody that preserves the structural ambiguity.

(5) Put the red cube on the block on the disk.

The experiment would include the possibility to manipulate the display, e.g., by mouse action. What do we expect subjects to do? How will their focussing probability develop over the time they listen to the instruction?



**Figure 5** A large grey block with a red cube on top, a yellow disk with a green block on top, and a red cube inside a round grey box.

### 3.1 A problem for formal semantics

Let us start from the theory side and attempt a conventional formal semantic derivation for a sentence like (5). There are two alternative constituent structures as in (6) and (7).

- (6) put [the [red cube]] [on [the [block [on [the disk]]]]]  
 (7) put [the [red [cube [on [the block]]]]] [on [the disk]]

The two constituent structures would result in formal representations that require the definite determiner function  $\lambda f \in D_{\langle e,t \rangle} \& \exists ! x f(x)=1. \lambda y [f(y)=1]$  to be applied to the function  $\lambda z. \text{red\_cube}(z)$  in (6) and to  $\lambda z. \text{block}(z)$  in (7), neither of which, with respect to the world depicted in Figure 5, satisfies the domain condition for  $f$ : the denotation of *red cube* is not a singleton – the display shows two red cubes – and for (7), there is no singleton denotation for *block* – the display shows at least two blocks. The structure in (6) seems to disqualify also intuitively, as the DP *the red cube* cannot unambiguously be assigned a referent in our display. The alternative DP, *the red cube on the block* in (7), seems intuitively fine and could be assigned an unambiguous referent. Still, also in this case the formal derivation is blocked, because the domain condition is not satisfied.

Supposing that our intuition is supported in an experiment yet to be carried out and subjects choose the interpretation corresponding to (7). How could we capture this theoretically? There are two solutions I am aware of. The first is by Nicholas Haddock (1987) and the second by Cecile Meier (2003).

The core idea of Haddock, applied to our example, is that, although the expression *the block* as such would not be unambiguous in the world of the display, listeners of (5), by the time they have reached the expression *the block*, would also know that the block must be a block on which there is a red cube. If this latter condition is added, the display in Figure 5 yields the needed singleton denotation. Haddock presents his idea in the framework of a constraint satisfaction approach and uses a store for collecting constraints, so that, in collecting constraints for the referent of the entire DP, he has already collected the constraints that the referent

must be unique, must be red, a cube, must be on something, and that that something must be described as a singleton. At this point the set of potential referents, given the display, is narrowed down to the disk and the big grey block. The final word, *block*, would provide the last piece of information and the constraint set is satisfied by the red cube on top of the big grey block. – There is a problem however that does not immediately show up when constraints are collected: if the uniqueness constraint on *block* is taken seriously, the singleton reference can only be established, intuitively speaking, by saying that it is the block on top of which there is the red cube. But what red cube? Well, presumably the one that is on top of the block, etc. – This is not entirely convincing.

The other proposal by Meier (2003) reverts to the idea that not all definites need to be used referentially, but some, in particular embedded definites as in our case *the block*, may be predicational. Meier uses an additional lexical denotation for the predicative use of the determiner that turns a property into a unique property and does not *presuppose* a unique block, but *entails* the existence of such a block. – There are indeed cases that may call for such a move, in particular when the determiner does not involve a specific but rather an arbitrary reference (or, perhaps, is not referential at all), as, e.g., in *She stood by the window* – which is not felicitously continued by *The other window was closed* (cf. Carlson e.a. 2006, Cieschinger 2006). Carlson speaks here of "weak definites", similarly Poesio (1994). In the case at hand, however, I believe that we are concerned with regular definites. Note, in particular, that the reference of *the block* may be resumed anaphorically – which should not be possible if the expression was used as a weak definite: *Pick up the red cube on the block. It must be cleared so that we can put the green block there.*

### 3.2 Incremental construction

Although I don't find Haddock's solution fully convincing as it stands, his idea of an incremental construction is also pursued here, albeit with the help of conventional formal semantic representations, as we used them in modelling Hartmann's observations. Let me illustrate the method by walking step by step through an utterance of (5), i.e., *Put the red cube on the block on the disk.*

The lexicon should provide us for *put* with a denotation like (8).

$$(8) \quad \lambda x_e [\lambda y_e \lambda z_{\langle e,t \rangle} \cdot \text{PUT}(x,y,z)] \langle e, \langle e, \langle \langle e,t \rangle, t \rangle \rangle \rangle$$

Assuming that in view of the experimental context the listener expects an instruction and that *put* is the first word in the instruction, an interpretation as an imperative is reasonable. We will ignore the deontic element here and simply insert the listener, *l*, as the subject:

$$(9) \quad [\lambda y_e \lambda z_{\langle e,t \rangle} \cdot \text{PUT}(l,y,z)] \langle e, \langle \langle e,t \rangle, t \rangle \rangle$$

The next word in the instruction is *the*, with a lexical denotation as in (4), i.e.,  $\lambda f: f \in D_{\langle e,t \rangle} \& \exists! [x f(x)] \cdot \lambda y [f(y)=1] \langle \langle e,t \rangle, e \rangle$ , where  $f \in D_{\langle e,t \rangle} \& \exists! [x f(x)]$  is a condition on the domain of *f*: the function is defined only for predicates

with a singleton denotation. The determiner thus imposes a first constraint on the denotation of the direct object expression, represented in (9) by the variable  $y$ . The string *put the* thus leads to the representation in (10). We are not claiming, though, that *put the* in any sense figures as a constituent. (10) represents only an intermediate state of the processor, which shows that the processor expects a predicate denotation to complete the information on the direct object denotation.

$$(10) \quad \lambda f: f_{\langle e,t \rangle} \& \exists! x(f(x)). \lambda z. \text{PUT}(l, iy.f(y), z)_{\langle \langle e,t \rangle, \langle \langle e,t \rangle, t \rangle \rangle}$$

The following word, *red*, with the lexical denotation  $\lambda g. \text{RED}(g)_{\langle \langle e,t \rangle, \langle e,t \rangle \rangle}$  provides such information by a constraint on a predicate  $g$ , i.e., a nominal still to come. It is added to the domain condition for  $f$ : the denotation of the direct object must not only be unique but must also be a red something.

$$(11) \quad \lambda g_{\langle e,t \rangle} \lambda f: f_{\langle e,t \rangle} \& \exists! x(f(x) \& g(x) \& \text{RED}(g)). \\ \lambda z. \text{PUT}(l, iy.f(y), z)_{\langle \langle e,t \rangle, \langle \langle e,t \rangle, t \rangle \rangle}$$

With the next word, *cube*, with the denotation  $\lambda x. \text{CUBE}(x)_{\langle e,t \rangle}$ , the direct object DP could be completed. The domain of  $f$  is now restricted so that  $f$  must be of the type  $\langle e,t \rangle$  and its denotation must be a singleton red cube.

$$(12) \quad \lambda f: f_{\langle e,t \rangle} \& \exists! x(f(x) \& \text{RED}(\text{CUBE}(x))). \lambda z. \text{PUT}(l, iy.f(y), z)_{\langle \langle e,t \rangle, t \rangle}$$

But the display does not offer a unique red cube, there are two. So the DP node cannot be closed and the processor must wait for further information that could help decide between the two red cubes in the display.

The principle is probably clear by now. The representation specifies at each point in the comprehension process the "meaning" of what has already been processed by assigning a denotation to the string *as it is understood in the utterance context*, including all information available. Constituent structure enters only via the semantic type of the intermediate representation, by specifying the denotations still required in order to arrive at a complete constituent. The semantic type of (12), e.g., is  $\langle \langle e,t \rangle, t \rangle$ . Thus what is still required to complete the instruction is an  $\langle e,t \rangle$  type denotation to fill in the third argument slot of the verb *put*. The first argument has already been fixed to the listener by the constant  $l$  and the second argument is specified as  $iy.f(y)$ , where  $f$  is restricted with respect to its domain: it must be a unique thing that is red and a cube.

The next word, *on*, with the lexical denotation  $\lambda x \lambda y. \text{ON}(y,x)_{\langle \langle e \rangle, \langle e,t \rangle \rangle}$ , starts a further constraint on the singleton that the processor is looking for by adding to the domain restriction for  $f$  and by adding the expectation that the next expression would provide a value for  $v$ , i.e., the location of  $x$ .

$$(13) \quad \lambda v. \lambda f: f_{\langle e,t \rangle} \& \exists! x(f(x) \& \text{RED}(\text{CUBE}(x)) \& \text{ON}(x,v)). \\ \lambda z. \text{PUT}(l, iy.f(y), z)_{\langle \langle e \rangle, \langle \langle e,t \rangle, t \rangle \rangle}$$

But this representation as it stands does not make explicit that the value for  $v$  must be an entity that has a red cube on it. The latter constraint is inferred from the domain restriction for  $f$ , and if our representation is to

predict how incoming information is used, the constraint must be added as a restriction on  $v$  as in (14).

$$(14) \quad \lambda v: \exists x(\text{RED}(\text{CUBE}(x)) \& \text{ON}(x, v)). \\ \lambda f: f_{\langle e, t \rangle} \& \exists! x(f(x) \& \text{RED}(\text{CUBE}(x)) \& \text{ON}(x, v)). \\ \lambda z. \text{PUT}(l, ty.f(y), z)_{\langle \langle e \rangle, \langle \langle e, t \rangle, t \rangle \rangle}$$

The following determiner, *the*, starts a DP for the position of the prepositional object variable  $v$  in (14) and adds a uniqueness constraint for the NP denotation to come. The processor now expects an  $\langle e, t \rangle$ -type denotation to complete the information on the thing on which there is a red cube.

$$(15) \quad \lambda g: g_{\langle e, t \rangle} \& \exists! w.g(w) \& \exists x(\text{RED}(\text{CUBE}(x)) \& \text{ON}(x, w.g(w))). \\ \lambda f: f_{\langle e, t \rangle} \& \exists! x(f(x) \& \text{RED}(\text{CUBE}(x)) \& \text{ON}(x, w.g(w))). \\ \lambda z. \text{PUT}(l, ty.f(y), z)_{\langle \langle e, t \rangle, \langle \langle e, t \rangle, t \rangle \rangle}$$

The next word, *block*, completes the DP, and the uniqueness condition on  $g$  is satisfied by the display: there is exactly one block in the display on which there is a red cube. Thus the processor can move via (16) to (17).

$$(16) \quad \lambda f: f_{\langle e, t \rangle} \& \exists! x(f(x) \& \text{RED}(\text{CUBE}(x)) \& \text{ON}(x, u. \exists v[\text{BLOCK}(u) \\ \& \text{ON}(v, u) \& \text{RED}(\text{CUBE}(v))]) . \lambda z. \text{PUT}(l, ty.f(y), z)_{\langle \langle e, t \rangle, t \rangle}$$

$$(17) \quad \lambda z. \text{PUT}(l, ty. \text{RED}(\text{CUBE}(y) \\ \& \text{ON}(y, u. \exists v[\text{BLOCK}(u) \& \text{ON}(v, u) \& \text{RED}(\text{CUBE}(v))]), z)_{\langle \langle e, t \rangle, t \rangle}$$

The trick that has given us a unique block is that the notion that *contextual denotation* for *the block* is not  $u. \text{BLOCK}(u)$  but rather  $u. \exists v[\text{BLOCK}(u) \& \text{ON}(v, u) \text{ RED}(\text{CUBE}(v))]$ . The latter has an unambiguous interpretation in the display, the former would not.

But there is also an alternative model for the unique interpretation of *the block* as follows. When the potential ambiguity of *the red cube* in (5) causes the parser not to close the current DP node, but to attach the following information to the NP, the reference resolution for the DP becomes the current task on the agenda of the comprehension process. Hence the following input must, if possible, be used immediately for this task, thus narrowing the attention to the two red cubes as alternative referents for the DP being built. If this hypothesis is correct, then already with the occurrence of the preposition *on* a decision for one of the two red cubes is preconfigured, and within the domain of attention thus created, the denotation of *the block* is indeed unique.

A Visual World experiment by Chambers e.a. (1998) strongly suggests the latter model. Using a display that included, among other things, either two containers (a can and a bowl) or just one container (a can), subjects heard instructions like *Put the cube inside the can*. In the first condition, focussing frequency for the can started to increase over the focussing frequency of the bowl between 100 and 200 ms after the onset of the noun *can*. In the condition with only the can and no other container, the saccades to the can went up already about 100 ms before the offset of *inside*, i.e., well before the determiner or the noun *can* was heard.

If the latter model is correct, we must take the preposition *on*, and the relation that it denotes, more seriously and allow  $\lambda x \lambda y \text{ON}(y,x)_{\langle\langle e \rangle, \langle e, t \rangle\rangle}$  to impose constraints on its argument domains.  $\llbracket \textit{on} \rrbracket$  would require a *location object*, i.e., the value of  $x$  (which, together with the relation  $\llbracket \textit{on} \rrbracket$  yields a location property) and a *located object*, the value of  $y$ . The domain for the block is thus limited to objects suited as *location objects*.

I believe that either model makes sense and that experimental work will have to decide between them. We predict on either model that there will be a significant increase in the focussing frequency for both red cubes, starting with the hearing of the word *red*. If then, starting with the word *on*, the focussing frequency for the red cube in the round box would drop and at the same time focussings would increase for the other red cube on top of the grey block, this would support the second of the above models. If on the other hand the differentiation between the focussings to the two cubes only starts with the word *block*, this would support the first model.

#### 4 What's new, then?

The modelling of the comprehension process proposed in this paper uses the apparatus of current formal semantics, including lexical entries and constituent structure, without modification. The representations are re-interpreted though as *representations of states of the processor*, and include all relevant information that is cross-modally available. Semantic types are procedurally re-interpreted as representing expectations of the processor. The feature that supports this re-interpretation is the lexically given domain restrictions of the various functions used: knowledge of the context enters as knowledge about the domain of our functions.

The central idea, then, is to change as little as possible in theories of linguistic knowledge and use linguistic knowledge to *constrain the theory of language processing*. Knowledge of the language is thus respected as an object of study in its own right. We don't want to deal with processing matters in the theory of linguistic knowledge, and we don't want to mistake processor properties for properties of language. - This is a strictly modular approach to linguistic processing, but it assumes strong interaction between modules.

#### Acknowledgements

The experiments reviewed in this paper were conducted at our institute in Osnabrück, involving work by Nadine Hartmann, Anke Karabanov, Graham Katz, and Peter König. I am very grateful for their cooperation. I am also grateful for the invitation to present these ideas at the 7th International Symposium on Language, Logic and Information, 1-5 October, 2007, in Tbilisi, Georgia, and at the Summer School on Formal Methods in Philosophy and Linguistics, 19.-31.08, 2007 in Tartu, Estonia and for the comments that I received. I also had very helpful comments from Stefan Evert and from two anonymous reviewers for the current volume. I regret that, for mere lack of space, I could not take up, or even acknowledge, all issues raised in these comments.

## References

1. Allopenna, P., Magnuson, J. S., & Tanenhaus, M. K. (1998) Tracking the time course of spoken word recognition using eye-movements. *Journal of Memory and Language* 38:419–439.
2. Altmann, G.T.M. & Kamide, Y. (1999) Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73: 247–264.
3. Altmann, G.T.M. & Steedman, M.J. (1988) Interaction with context during human sentence processing. *Cognition* 30:191-238.
4. Boland, J.E.(2005) Visual arguments. *Cognition* 95:237–274
5. Chambers, C.G., Tanenhaus, M.K., Eberhard, K.M., Carlson, G.N. & Filip, H. (1998). Words and worlds: The construction of context for definite reference. In M.A. Gernsbacher & S.J. Derry (Eds.) *Proc. of the 20th Ann. Conf. of the Cognitive Science Society*. Mahwah, NJ: Erlbaum. pp. 220-225
6. Cieschinger, M. (2006) *Constraints on the Contraction of Preposition and Definite Article in German*. BSc Thesis, Univ. of Osnabrück, CogSci. [http://www.cogsci.uos.de/~CL/download/BSc\\_thesis\\_Cieschinger.pdf](http://www.cogsci.uos.de/~CL/download/BSc_thesis_Cieschinger.pdf)
7. Carlson, G., R. Sussman, N. Klein, and M. Tanenhaus (2006). Weak definite NPs. In C. Davis, A. R. Deal and Y. Zabbal (eds.). *Proceedings of NELS 36*. UMass/Amherst. GLSA/Chicago
8. Dahan, D., Swingley, D., Tanenhaus, M.K., & Magnuson, J.S. (2000). Linguistic Gender and Spoken-Word Recognition in French. *J.Mem&Lang* 42:465-480.
9. Eisenband, A.J., Brown-Schmidt, S., & Trueswell, J (2000): The Rapid Use of Gender Information. *Cognition* 76, B13-B26
10. Garfield, J.L. (ed) (1987) *Modularity in Knowledge Representation and Natural Language Understanding*. MIT Press.
11. Haddock, N.J. (1987): Incremental interpretation and combinatorial categorial grammar. *Proceedings of IJCAI 1987*. Vol.2:661-663
12. Hagoort, P. & van Berkum, J. (2007) Beyond the sentence given. *Philosophical Transactions of the Royal Society* 362: 801-811.
13. Hartmann, N. (2004) *Processing Grammatical Gender in German - An Eye-Tracking Study on Spoken-Word Recognition*. BSc Thesis. Univ. of Osnabrück, Cognitive Science. [http://www.cogsci.uni-osnabrueck.de/~CL/download/Hartmann\\_Gram Gender.pdf](http://www.cogsci.uni-osnabrueck.de/~CL/download/Hartmann_Gram Gender.pdf)
14. Heim, I. & Kratzer, A. (1997). *Semantics in Generative Grammar. An Introduction*. Blackwell. Oxford.
15. Karabanov, A., Bosch, P. & König, P. (2007) Eye Tracking as a Tool to Investigate the Comprehension of Referential Expressions. In S. Featherston & W. Sternefeld (eds.): *Roots. Linguistics in Search of its Evidential Base*. Mouton de Gruyter. Berlin. pp.207-226
16. Marslen-Wilson, W.D., & Tyler, L. K. (1980) The temporal structure of spoken language understanding. *Cognition* 8:1-71
17. Meier, C. (2003) Embedded Definites. In van Rooy, R. (ed.) *Proceedings of the Fourteenth Amsterdam Colloquium*. Amsterdam. ILLC. pp163-168
18. Poesio, M. (1994). Weak Definites. In: *Proceedings of the Fourth Conference on Semantic and Linguistic Theory*. SALT-4.
19. Tanenhaus, M.K., Spivey-Knowlton, M.-J., Eberhard, K.M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language-comprehension. *Science* 268:1632-1634.
20. Trueswell, J.C. & Tanenhaus, M.K. (eds.) (2005) *Approaches to Studying World-Situated Language Use*. MIT Press.