

Procedural Meaning and the Semantics/ Pragmatics Interface

ANNE BEZUIDENHOUT

1 Introduction

Relevance theorists have argued that we must distinguish between words that encode concepts and those that encode procedures. The latter encode instructions that constrain the inferential phase of verbal communication. (This is explained in more detail below). This raises the question as to how we are to understand the notion of procedural encoding. I will argue that the notion of a procedural unit is something that has a place in an account of language *use*, and hence it belongs to a theory of pragmatic performance and not to a theory of semantic competence.

A very strong statement of this claim would be that the phrase ‘procedural semantics’ is a contradiction in terms. This paper examines an argument that purports to demonstrate this. The conclusion of the argument is that if one is interested in a theory of language production and understanding (in other words, in a theory of linguistic performance), and one appeals to the notion of procedural knowledge, then one cannot regard such procedural knowledge as embodied in semantic rules without turning it into something with conceptual content, and thus turning it into something that loses its procedural character. In other words, one cannot simultaneously treat something as procedural and as semantic. Something that lies on the procedural side of the procedural/declarative divide is something inherently

pragmatic and belongs to a performance system, and is distinct from the knowledge that is constitutive of a speaker-hearer's semantic competence.

By stating the conclusion in this way, it seems that one must accept that the following distinctions line up with one another:

Semantic/pragmatic
Declarative/procedural
Competence/performance

In the course of critically evaluating the above-mentioned argument against procedural semantics, I will ask whether the parallelism between these three distinctions is ultimately sustainable.

2 The Conceptual/ Procedural Contrast

I begin with a description of the contrast between procedural and conceptual encoding. The vast majority of lexical items have conceptual meaning, including common nouns ('chair', 'water', etc.),¹ verbs ('consider', 'leap', etc.), adjectives (e.g., 'red', 'slow', etc.), adverbs ('sadly', 'quickly', etc.), and prepositions ('behind', 'under', etc.). That is to say, these items encode concepts. If a lexical item has conceptual meaning, this meaning can potentially contribute to the truth-conditional content of an utterance containing that lexical item. However, having conceptual meaning doesn't guarantee that the item will be truth-conditionally relevant, since items with conceptual meanings sometimes play a non-truth-conditional role. Consider the following:

(1) Sadly, John's mother died last night.

(2) John looked sadly at the mess his dog had made.

'Sadly' has some conceptual content, but in (1) it is functioning as a sentence adverbial, and hence its conceptual content does not contribute to the truth-conditional content of the utterance, or at least to the truth-conditional content of the primary proposition expressed by the utterance. Rather, it contributes to a higher-level proposition that expresses the speaker's attitude towards the proposition that John's mother died the night prior to the

¹I set to one side the Putnam/Kripke view about natural kind terms, such as 'tiger' and 'water'. On this view, the concepts associated with such terms are merely part of the stereotypes associated with these terms. These stereotypes are not extension-determining meanings. Instead, what fixes the extension of a natural kind term *N* is a causal-historical chain of uses of *N* that can be traced back to some event at which *N* was first introduced into the language. The extension of *N* includes all and only those objects whose natures are relevantly similar to the natures of those objects demonstratively or otherwise picked out during this introductory event. Since it is up to science to discover these natures, we must defer to the experts to determine which objects belong to the extensions of our terms. This linguistic division of labor means that we can use these terms even though our associated concepts are not uniquely identifying.

time of utterance. The speaker could have conveyed his attitude to this proposition using non-linguistic means, say by uttering the sentence 'John's mother died last night' in a sad tone of voice. The use of 'sadly' in (1) is in contrast to its use in (2), where 'sadly' *does* contribute conceptual content to the proposition expressed.

In contrast to items such as common nouns, verbs, etc., relevance theorists have claimed that words such as 'but', 'however', and inferential 'so' encode procedures. These items indicate something about the context in which the utterances of which they are a part are to be processed. They guide the hearer towards intended contextual effects,² and hence reduce the overall effort required to process the discourse.³ Consider utterances of sentences such as:

(3) Kathleen is 37 years old but still attractive.

(4) It is winter in Genoa but summer in Johannesburg.

The use of 'but' in (3) is sometimes called the denial of expectation use. See Lakoff (1971). This use of 'but' presupposes that 37-year-old women are generally unattractive. It is possible that the speaker shares this prejudice against 37-year-old women.⁴ However, even if the speaker rejects this presupposition, she must be assuming that this assumption is widely held by members of a certain cultural/social community. For instance, (3) could be uttered in response to what the speaker perceives to be Hollywood's prejudice against older women.

The 'but' in (4) signals a contrast between two sets of implications. It signals that whatever propositions the listener was prepared to infer from the utterance of the first conjunct, he should infer a parallel but opposite set of propositions from the utterance of the second conjunct. For instance, if the listener was prepared to infer from the proposition expressed by the

²There are three kinds of contextual effects that result from processing utterances. There might be contextual implications (i.e., ones that follow from what is expressed by the utterance together with assumptions already in the context, but that do not follow from either of these components alone). Second, an existing assumption might be strengthened. Third, an existing assumption might be contradicted.

³RT is often accused of emphasizing language understanding and ignoring language production. But the reduction in overall effort mentioned in the text applies equally to speaker and to hearer. If the speaker can rely on her hearer to pick up on her linguistic clues as to how to process her utterances, she can spare herself the effort of explicitly articulating everything she means to convey. Levinson (2000), in the course of defending his default theory of generalized conversational implicature, makes the point that articulation is costly.

⁴Thomas (1995: 57) cites a reaction by the actress Kathleen Turner when she received a movie script that was written for a female character '37 years old but still attractive'. Turner circled the word 'but' in the script and returned it to the movie studio with a note saying 'Try again'. She clearly attributed the presupposition mentioned in the text to the writer(s) of the script and perhaps also to the movie producers and studio executives.

utterance of the first conjunct that anyone traveling to Genoa should pack warm clothes, then he should infer from the proposition expressed by the utterance of the second conjunct that anyone traveling to Johannesburg should *not* pack warm clothes. In addition to the uses illustrated here, there are undoubtedly further uses of ‘but’.

Lexical items such as inferential ‘so’, ‘since’, ‘as’ and ‘because’, and words and phrases such as ‘moreover’ and ‘after all’ are also said to encode procedures. Consider:

- (5) John wasn’t at the party last night. So he must have stayed home.
- (6) He was hungry, so he went to McDonalds.
- (7) As/since/because John was hungry, he went to McDonalds.
- (8) As/since John isn’t here, he must still be in his meeting.
- (9) John owes me money. Moreover, he owes me a lot.
- (10) Have another drink. After all, it’s your birthday.

‘So’ in (5) indicates that the utterance that follows should be processed as a conclusion. It differs from the ‘so’ in (6), which indicates that John’s going to McDonald’s is a causal consequence of his being hungry. ‘As’, ‘since’ and ‘because’ may signal that what follows is a cause, as in (7), or a reason, as in (8). ‘Moreover’ in (9) signals that what follows is an elaboration, and ‘after all’ in (10) signals that what follows offers justification or support.

Many other lexical items and linguistic constructions, too numerous to survey here, have been said to encode procedural meanings. For discussion see Andersen & Fretheim (2000); Blakemore (1987, 1988); Blass (1990); Fretheim (1998a,b); Giora (1997); Iten (1997, 1998, 2002); Nicolle (1995, 1998); Wilson (1998); Wilson & Sperber (1993).

It is also possible to have items that encode *both* concepts *and* procedures. For instance, consider:

- (11) He [gesturing to someone] is the man I saw stealing the car.
- (12) The man from the IRS called. He will be here tomorrow.

The use of the pronoun ‘he’ carries a presupposition of weak familiarity, and hence its use signals to the listener that he should identify the referent of the pronoun with an individual that is salient/accessible in the context. The individual may be one that is salient in the physical context, as in (11), or from the prior linguistic context, as in (12). However, ‘he’ also carries some (admittedly rather minimal) conceptual/descriptive meaning. It carries person, number, gender, and animacy information (third person singular animate male). This conceptual content may or may not be truth-conditionally relevant. Direct reference theorists, for example, assume that singular referring expressions, including pronouns, contribute nothing but their referents to the propositions expressed with their help. Hence the gender, number and person information encoded by a pronoun such as ‘he’ would not be part of the proposition expressed by its means, but would

simply constrain the search for an appropriate referent. However, if there are attributive uses of such pronouns, as argued by Bezuidenhout (1997), then such conceptual meaning may become truth-conditionally relevant. An example of such an attributive use is: 'He [pointing at a large footprint in the sand] must be a giant'. This arguably expresses the proposition that the male who made the footprint must be a giant.

Wilson and Sperber (1993) argue for the need to distinguish procedural and conceptual encoding, and they suggest at least two ways of doing so:

- (a) Conceptual meaning is available to consciousness, whereas procedurally encoded information is not (1993: 16).
- (b) Conceptual meaning is compositional, whereas procedurally encoded information is not (1993: 18).

However, both these criteria are problematic. Criterion (a) is problematic because even if we concede, as I think we should, that procedurally encoded information is not available to consciousness,⁵ it is not clear that conceptual meanings as conceived of in RT are available to consciousness either. According to RT, conceptual meanings are the lexical concepts that are (potential) constituents of the LF-representations that are the output of linguistic decoding processes. These representations in turn are inputs into the pragmatic processes of enrichment and loosening that result in *ad hoc* concepts. The latter are constituents of the propositional forms (explicatures) of utterances. See Carston (1997, 2002). At best it would be these *ad hoc* concepts (the pragmatically adjusted lexical concepts) that are available to conscious awareness. Thus it is not clear that the consciousness test helps to make a principled distinction between conceptual and procedural encoding.

Criterion (b) is also problematic. It is suggested in the course of a discussion of illocutionary adverbials. Wilson and Sperber claim that these adverbials have conceptual meanings and yet are non-truth-conditional, in the sense of not contributing content to the proposition expressed by an utterance (i.e. to its base-level explicature). Instead, these adverbials contribute content to higher-level explicatures. They indicate something about the speaker's attitude towards the expressed content or about the manner in which the speaker is expressing this content. Consider:

⁵Wilson and Sperber think that procedurally encoded information is unavailable to consciousness because we do not have direct access to 'the inferential computations used in comprehension' (1993: 16). This suggests that they are identifying procedural information with *inferential computations*. This is different from the view I will defend in the following sections, according to which procedural meanings are *causal dispositions*. For present purposes this difference between Wilson and Sperber and myself is not relevant, since I would agree that procedural information is not available to conscious awareness.

- (13) Frankly, he is someone I never liked.
- (14) Confidentially, he will not get the job.
- (15) Unfortunately, I can't help you.

In (13) the speaker expresses the proposition that the male individual in question is someone the speaker never liked, but she also conveys the proposition that this information is being given in an open manner, that reveals her true opinions. (14) would be analyzed in a similar manner. In (15), the speaker expresses the proposition that she can't help the hearer, but she also conveys the proposition that this fact is something she regrets. In defense of the view that such sentence adverbials should be given a conceptual rather than a procedural analysis, Wilson and Sperber point out that such adverbials can be part of semantically complex phrases, which must be compositionally understood, as shown by (16) – (18) below:

- (16) Frankly speaking, he is someone I never liked.
- (17) Speaking frankly, though not as frankly as I would like, he is someone I never liked.
- (18) In total, absolute confidence, he will not get the job.

However, the lexical items that are treated in RT as encoding procedural information can also occur as a part of semantically complex phrases. Consider the following:

- (19) She is interested in our job. But, and this is a big but, we'll have to offer her much more money.
- (20) Photoshop allows us to digitally manipulate photographic images. So, although perhaps not obviously so, photography has become more like painting.

I do not think that the occurrence of these items in semantically complex phrases should deter us from a procedural analysis of these terms. But it does suggest that the compositionality test is not a reliable one for distinguishing the procedural from the conceptual.

In this paper I will not be concerned with deciding which items are to be analyzed in procedural terms and which are to be analyzed in conceptual terms (which to a large extent is what discussions of the conceptual/procedural distinction have been focused on). Rather, I will assume that there are some items that are correctly analyzed in procedural terms, and then ask what it means to say that such terms encode procedures.

3 Rules, Representations and the Procedural/ Declarative Distinction

So, what does it mean to say that lexical items of the sort mentioned above encode procedural information, or as Blakemore (1992: 151) puts it, 'encode instructions for processing propositional representations'? It is natural

to assume that it means that the entries for these items in an ideal speaker-hearer's mental lexicon contain these instructions. In other words, to say that the lexical entry for item *X* contains procedural information is to say that there is a rule or instruction 'written' in the entry for *X* that specifies that a certain procedure must be followed if certain conditions are fulfilled. Such a rule would be what AI researchers call a *production rule*, which is a condition-action pair, $\langle C, A \rangle$. It is a rule that says that if a certain condition *C* is fulfilled, then a certain action *A* must be performed.

Talk of rules being written under entries is a metaphorical way of speaking. The idea is that the mental lexicon uses a proprietary code for representing semantic information. This code could be *Mentalese*, which is the system of representation that Fodor (1975) claims is the language of thought. Hence these rules will be written in *Mentalese*. When the speaker-hearer accesses an item in the lexicon, he or she is able to access the information associated with that item, presumably in some sort of 'look-up' process that allows the entry for that item to be searched for relevant information. If the item encodes procedural information, the look-up process will access a rule. If the item encodes conceptual/descriptive information, the look-up process will access a concept.⁶

If the look-up process accesses a rule, and the condition of the rule is met, a certain action will be performed. Here the action in question is some sort of symbol manipulation (i.e. manipulation of *Mentalese* representations). In other words, given certain representations as input, certain representations will be produced as output. Figure 3.1 illustrates this representationalist picture of how (from the perspective of language comprehension) conceptual and procedural information is embodied at the sub-personal, computational level. On this representationalist conception, concepts and procedures are just two sorts of meanings, and both should be regarded as a part of our semantic competence. With these assumptions in place, it is now possible to state the argument against procedural meaning.

⁶Some philosophers equate concepts with rules. The concept *red* is a rule that specifies the conditions under which it is appropriate to apply the word 'red' (or whatever word in your language is associated with that concept – 'rouge' or 'rooi' or whatever). This would be to reject the conceptual/procedural distinction as understood within relevance theory, since all lexically encoded meaning would collapse into procedural meaning. So, for the purposes of this paper I will assume that concepts and rules are distinct.

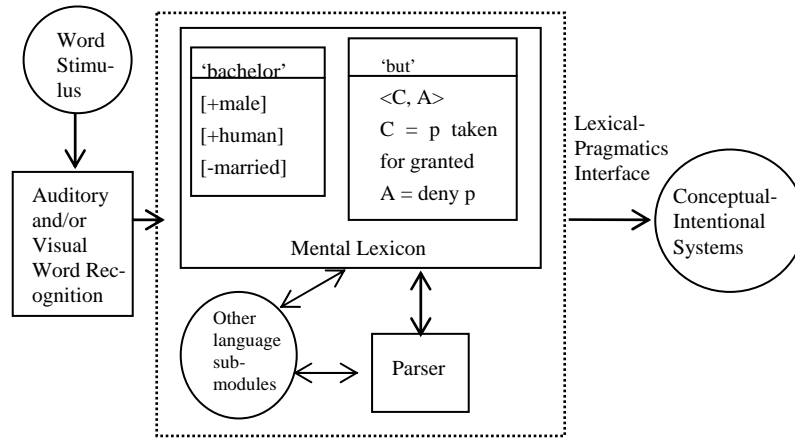


Figure 3.1: The Representationalist Picture of the Mental Lexicon

The argument against procedural meaning thus conceived is very simple to state. It is a version of an argument that is used in a slightly different context by Dennett (1981), and by Kripke (1982) in his discussion of Wittgenstein on rule following. If procedural rules are treated as items that are represented in Mentalese and as items that are listed in lexical entries (which in turn are simply lists of representations – data structures – that can be searched), then we face a form of the rule following paradox. Rules themselves will simply be more symbols along with the symbols they are meant to manipulate. But then it looks as though we'll have to posit another set of rules for instructing us on how to use the original rules. Now, if these new rules are themselves items that are represented in Mentalese, the same problem arises all over again. Clearly we are off on an infinite regress here. Wittgenstein thought that if this regress of rules is to be avoided, there must be some point at which the rules give out. The whole system must therefore be grounded in something other than rules. For Wittgenstein, rules are grounded in actions or practices.

Wittgenstein would not have approved of the sort of computational perspective on language use that I have adopted in this paper, but his insight can be translated into one that applies within this computational perspective. The regress of rules I described can be stopped if at some point rules give way to procedures, where procedures are understood in some *non-representational* or *non-symbolic* way. We should think of procedures as part of the architecture of the performance system. In other words, procedures are embodied as part of the causal structure of the language production and comprehension system, and are not symbols or representations that

are manipulated (inferentially or otherwise) within the system. Here I am following Dennett (1981: 50), who contrasts two ways of storing information: storing it in propositional or coded form and storing it tacitly, in the organization of the representational system.

Dennett makes this distinction in the course of a discussion of how our beliefs are stored. We potentially believe an indefinite number of things, but in the course of our lives will only generate some finite number of propositions by means of an extrapolator-deducer mechanism, whose operations Dennett describes as follows (1981: 45-46, his emphases):

It has the capacity to extract axioms from the core when the situation demands it, and deduce further consequences. To do this, it needs to have an information store of its own, containing information about what items it would be appropriate at any time to retrieve from the core, ... Now how will the extrapolator-deducer mechanism store its information? In its own core library of brain-writing sentences? If it has a core library, it will also need an extrapolator-deducer mechanism to act as librarian, and what of *its* information store? Recalling Lewis Carroll's argument in 'What the Tortoise Said to Achilles', we can see that an extrapolator-deducer will be hamstrung by a vicious regress if it must always rely on linguistically stored beliefs which it must retrieve and analyze about what can be deduced from what...

The conclusion is that writing – for instance, brain writing – is a *dependent* form of information storage. The brain must store at least some of its information in a manner not capturable by a brain-writing model.

One might call the alternative conception of procedures that I am advocating the *dispositionalist* conception, in contrast to the *representationalist* or *rule-based* conception of procedures that I have argued leads to an infinite regress of rules. If we accept this dispositionalist conception of procedures, must we reject the relevance theory distinction between conceptual and procedural encoding? Not at all. However, we will need to understand the distinction in a different way from the one assumed at the beginning of this section. The conceptual/procedural distinction is not a difference in the type of information encoded by a lexical item (concept vs. rule). Rather, the distinction is one regarding how such knowledge is embodied. Conceptual meanings (concepts) are embodied as mental representations/symbols. Procedures are embodied as causal dispositions, as ways in which the system acts on representations.

Blakemore (1992: 148-151) contrasts what she calls a representational with a procedural analysis of expressions like 'so', 'after all' etc. In her discussion, 'representational' is functioning as a synonym for 'conceptual' so that the contrast Blakemore is discussing is just the concep-

tual/procedural distinction discussed in Section 2. Blakemore attributes a 'representational' (i.e. conceptual) analysis of items like 'so' to Grice. Grice (1989: 362) claims that one can make what he calls a ground floor statement and simultaneously perform a higher-order speech act of commenting in a certain way on the lower-order speech act. Consider 'Our computer is down. So I cannot help you'. On Grice's analysis, this simply asserts that the speaker's group's computer is not functioning and that the speaker cannot help the hearer. But the 'so' communicates the higher-order proposition that the speaker's being unable to help is explained by the computer's being down (i.e. that being unable to help is a causal consequence of the computer's being down). Blakemore construes this analysis of 'so' as a 'representational' one, since she says it amounts to claiming that 'so' has conceptual content (because 'so' communicates a higher-order proposition). Blakemore contrasts this with her own procedural account of 'so', according to which 'so' encodes instructions for processing propositional representations. So Blakemore's use of 'representational' is different from the one in the previous paragraph, which applies not to a particular type of analysis of lexical items such as 'so', but to a particular understanding of what it is for an item to lexically encode a procedure. Blakemore does not address the latter issue. Blakemore (1992: 150) says that when a speaker uses 'so' *the speaker is instructing the hearer* to interpret what follows as a conclusion. But she does not offer any details of a sub-personal, computational account to explain how this personal-level claim can be true. In contrast, my main concern is to offer a computational account that explains how the conceptual/procedural distinction is psychologically realized.

I have argued that procedures must be thought of as part of the causal structure of the *language performance system* (viz., the language production and comprehension system). This is because, on the comprehension side, the sort of procedural knowledge we are concerned with is knowledge of what are appropriate contexts for the interpretation of utterances and knowledge of how to connect utterances so as to produce adequate contextual effects for no gratuitous processing effort. On the production side, it is knowledge of how to guide the hearer to these contexts and how best to organize the discourse so as to make one's communicative intentions clear. So we have moved quite far from the representationalist idea that conceptual and procedural meanings are both stored in the mental lexicon. Procedures are strictly not a part of the lexicon at all, but of the performance system, although presumably the performance system is able to use information from various sources, including conceptual information from the mental lexicon.

Perhaps we've moved too quickly. Perhaps we need to backtrack and assume that procedures and concepts *are* both stored in the lexicon, the

former tacitly and the latter explicitly. This would be tantamount to thinking of procedural information as part of the causal structure of the *mental lexicon*, not of the *performance system*. Now I agree that the mental lexicon must have certain information tacitly stored. As the quote from Dennett (1981: 45-46) suggests, *any* cognitive (sub-)system will need to have some sort of extrapolator-deducer mechanism, which is able to manipulate the information that is symbolically represented in that system. If this mechanism is not to be 'hamstrung by a vicious regress', its operation will necessitate a distinction between information that is represented in symbolic form and information that is hard-wired into the architecture of the (sub)-system. This applies just as much to the mental lexicon as it does to the performance system.

Clearly, we must recognize the existence of certain purely lexical processes, e.g. the hypothesized 'look-up' processes mentioned earlier. The lexicon is after all just a store of information and there must therefore be purely lexical processes involving the storage, cataloguing, and retrieval of such information.⁷ Moreover, there must be rules governing such processes, and these rules cannot be (or cannot all be) propositionally represented, or else we face a regress of rules. So there must be information tacitly stored in the causal structure of the mental lexicon. But even though we must recognize that some information is tacitly stored in the lexicon, this information is not the procedural information that was described in the previous section.

We must also recognize that there are processes such as sense extension, lexical disambiguation, type coercion, conceptual blending, pragmatic enrichment and loosening, etc. See Asher and Lascarides (1996); Carston (1997, 2002); Copestake and Briscoe (1996); Coulson (2001); Fauconnier (1985, 1997); Fauconnier and Sweetser (1996); Pustejovsky (1995, 1998); Pustejovsky and Boguraev (1996); Sperber and Wilson (1995). Some people might regard these as lexical processes. However, all of these are processes in which lexical and pragmatic information interact, and so they cannot be thought of as *purely* lexical. Hence the procedures governing these processes cannot be purely lexical either.

⁷What information is lexically stored is a matter of some dispute. Some theorists, such as MacDonald et al. (1994), argue that a great deal of information is included in the lexicon. Besides the usual semantic, lexical category and thematic role information that one would expect in a lexical entry, they also think that there is information about possible argument structures and possible phrase structures (X-bar structures) that the item could be a part of. (1994: 688, Figure 3). Moreover, they claim that 'each component of a lexical entry carries information about its frequency of occurrence in the language' (1994: 685).

For example, Asher and Lascarides (1996) show how discourse level information can interact with lexical information in the process of lexical disambiguation. Asher and Lascarides claim that there are various rhetorical relations that affect the structure of discourse. The rhetorical relations they have in mind are ones such as: *Narration*, *Elaboration*, *Explanation*, *Background*, *Evidence*, *Consequence* and *Contrast*. Consider the following two discourses:

- (1) The judge asked where the defendant was. His barrister apologized, and said he was at the pub across the street. The court bailiff found him slumped underneath *the bar*.
- (2) The judge asked where the defendant was. His barrister apologized, and said he was at the pub across the street. But in fact the court bailiff found him slumped underneath *the bar*.

Asher and Lascarides argue that the ambiguous term ‘the bar’ will be resolved in favor of the pub meaning in case (1), because this is a narrative story, and if we see the sentences in the discourse as connected via the relation *Narration*, we will assume that the events are described in their temporal sequence. On the other hand, the use of ‘but’ in (2) signals a contrast, and so the events described will be related via the rhetorical relation *Contrast*. In this case the term ‘the bar’ will be interpreted in its courtroom sense.

These are cases where discourse knowledge helps in lexical disambiguation, so the influence is from pragmatics to lexical semantics. But sometimes the influence is in the opposite direction. Knowledge of semantics determines what sort of discourse relation one will infer. Consider:

- (3) Max fell. John pushed him
- (4) Max enjoyed a large meal. He ate salmon and devoured lots of cheese.

In the case of (3), one’s semantic knowledge that pushing can cause the thing pushed to move will suggest that the second sentence in (3) is to be understood as an explanation for the occurrence reported by the first. In (4), one’s semantic knowledge that enjoying a meal involves eating food and that salmon and cheese are types of food will suggest that the second sentence is an elaboration on the event reported by the first sentence.⁸

⁸Asher and Lascarides (1996, 1998) offer analyses of discourse within a framework they call Segmented Discourse Representation Theory (SDRT). In SDRT, discourse contexts are represented as recursive, relational structures (SDRSs), which involve discourse representation structures (DRSs) that represent the contents of clauses, but that also represent discourse relations like *Parallel*, *Narration* and *Background*. These relations hold between DRSs and they indicate how the discourse hangs together. So discourse relations that are given a procedural treatment in RT are declaratively represented in SDRT. This declarative approach is not something that Asher and Lascarides attempt to justify.

The main point here is that because processes such as sense extension and lexical disambiguation involve the *interaction* of lexical and pragmatic information, the procedures governing them cannot be thought of as purely lexical. Hence these procedures cannot be thought of as part of the causal architecture of the mental lexicon. Rather, they must be thought of as part of the causal architecture of the performance system more generally. What is true of the procedures governing processes such as sense extension is equally true of the procedures that were described in Section 2, which govern the inferential phase of verbal communication. That the operation of these procedures is not confined to the mental lexicon should be even clearer, for these procedures are not used only to interpret single lexical items, or even single phrases or sentences.⁹ They play a role also in the recovery of implicatures and help the hearer to construct discourse-level representations. Hence there is even more reason than in the case of the procedures involved in sense extension, lexical disambiguation, etc. to see these procedures as non-lexical or at least not purely lexical.

So I conclude that conceptual and procedural information are not two types of information both of which are stored in the lexicon, one type explicitly and the other tacitly. Rather, procedural information isn't properly speaking a part of the lexicon at all, but part of the causal architecture of the performance system more generally.

Thus far the picture that emerges is that lexical-conceptual meanings are embodied in a different way from the procedural knowledge that is used in inferential communication. To mark this fact that one might invoke Chomsky's distinction between competence and performance. The conceptual meanings that are associated with lexical items constitute an ideal speaker-hearer's *semantic competence*, whereas the ability to use those concepts constitutes the ideal speaker-hearer's procedural knowledge, and this procedural knowledge is what drives language production and understanding (*viz. linguistic performance*). So it appears then that the conceptual/procedural distinction lines up with the competence/performance distinction as well as with the semantic/pragmatic distinction. The notion of a procedural unit is something that has a place in an account of language *use*, and hence it belongs to a theory of pragmatic performance and not to a theory of semantic competence.

⁹Sometimes they are used in this way. For example, indexicals such as 'he' encode procedural information, which is used in context to determine the referents of these indexicals. Hence this procedural information is used in the recovery of *explicatures*. On the other hand, lexical items such as 'but' encode information that is used in the recovery of *implicatures*.

4 Objections to the Dispositional Analysis and Replies

One might try to resist the causal/dispositionalist understanding of the procedural element in language by insisting that the procedural knowledge associated with lexical items such as ‘but’, ‘however’, etc. *is* rule-based. After all, we can formulate the procedure associated with ‘but’ as a rule or instruction that has propositional content. Hence to say that we have the procedural knowledge in question is to say that we have mentally represented what the rule specifies (the condition-action pair). If we have mentally represented this condition-action pair, then we should say that this rule is part of our *semantic competence*, and that it is knowledge that is accessed in the course of language production and understanding.

In reply, we need to distinguish what belongs to our *theory* of language understanding and production (viz. our theory of linguistic performance), and what belongs to the performance system itself. As theorists we surely *can* give propositional expression to the procedural knowledge embodied in the linguistic performance system. But if we are to avoid the regress of rules mentioned in the previous section, we cannot conceive of these rules as implemented in the linguistic system in the form of mentally represented rules that are entered in the mental lexicon in the same way that conceptual meanings are.

A second objection to the dispositionalist view is that even if items like ‘but’, ‘so’, and ‘after all’ encode procedural knowledge, this knowledge will still have to be exercised in particular conversational contexts. This suggests that there must be some sort of ‘trigger condition’ associated with these lexical items that indicates when it is appropriate to exercise the dispositions they encode. Surely such trigger conditions will have to be explicitly represented, so that after all procedural knowledge will have to rest on something declarative. However, like the first objection this confuses what is part of our theory of the performance system and what is part of the system itself. We as theorists can explicitly state the conditions under which a certain disposition will be exercised. But all that is necessary is that there be some situation that triggers the exercise of the disposition in the speaker-hearer. The speaker-hearer need not represent this situation, either explicitly or tacitly.

A third objection to the dispositional view of procedures is a version of an objection that Chomsky (1996) raises against Dummett and Kenny. Chomsky attributes to these philosophers the view that knowledge of language is an ability, and hence that knowing a language is just like knowing how to ride a bike. Linguistic knowledge for these philosophers is knowledge-how rather than propositional knowledge-that. Chomsky objects that abilities or any kind of know-how must be grounded in something non-

dispositional. He writes: 'knowing-how involves a crucial cognitive element, some internal representation of a system of knowledge' (Chomsky, 1996: 565). So, applying this to the dispositionalist understanding of procedures, one might argue that procedural knowledge presupposes declarative knowledge, and that underlying any procedure is some mentally represented system of propositional knowledge. This would appear to reverse the order of dependence suggested by the regress of rules argument laid out in the previous section. That argument purported to show that propositional knowledge must rest ultimately on something non-representational/procedural.

I believe that the regress of rules argument trumps any claim that propositional knowledge is more basic than procedural knowledge. But in any case, I think this objection from Chomsky can be defused, because there is an understanding of Chomsky's point about linguistic abilities according to which it is *not* a challenge to the conclusion of the regress of rules argument. Chomsky's point can be seen as a particular application of a more general argument that has been made in the philosophy of science about dispositional properties, such as the property of being soluble. It has been argued that solubility, say of table salt, cannot be reduced to facts about the behavior of table salt in various types of situations. It is of course true that if table salt is placed in water it will dissolve, and if it is placed in gravity it will dissolve, and so on. But these dispositional facts about table salt must supervene on something more basic. In particular, there must be facts about the microstructure of table salt that explain these dispositional facts.

In a similar vein, Chomsky can be understood as saying that linguistic abilities rest on and are explained by something more basic. Presumably the more basic level here is the computational level. Chomsky describes this as the level at which we can talk of an 'internal representation of a system of knowledge'. This phrase of Chomsky's may seem to commit us to the idea that all linguistic knowledge is encoded in mental symbols. However, the phrase 'internal representation' can be treated as a term covering both explicit and tacit ways of storing information. (See Dennett 1981: 50). Not all information is explicitly mentally represented (i.e. by means of mental symbols). Some is stored tacitly, in the organization of the representational system. On this way of understanding Chomsky, his critique of Dummett and Kenny on linguistic abilities does not contain the seeds of a challenge to the dispositional understanding of procedures, given that we construe the dispositional view as claiming that at least some linguistic information must be tacitly stored, in the organizational structure of the performance system.¹⁰

¹⁰This does mean stretching the word 'representation' so that it covers both mental symbols and causal structures, and requires a rephrasing of the distinction between what I earlier called

A fourth objection is that I've misapplied the regress of rules argument. One might argue that the most that argument shows is that *some* procedural information must be tacitly represented in the causal architecture of the system, not that *all* of it must be. For all the regress argument shows, much of our procedural knowledge is stored in the mental lexicon (along with lexical concepts) in the form of explicit rules. Well, yes, there could be propositionally stated rules in the mental lexicon. But then these fall on the conceptual side of the conceptual/procedural divide, and they are not properly speaking procedural in the sense I intend. They stand in need of interpretation or application just as much as non-propositional lexical concepts do, and in that sense such rules do not drive interpretation. They are things being manipulated, not things doing the manipulating.

A fifth objection is that my view commits me to the claim that there are no procedural symbols in the language of thought. Natural language words that have conceptual meanings correspond to lexical concepts, and these lexical concepts are in effect (a subset of) symbols of Mentalese. So by denying that words such as 'but' correspond to concepts, I am committed to saying that there are no analogs in Mentalese to words such as 'but'. There is no mental symbol in Mentalese that plays the same role that 'but' plays in English. But this is implausible, since one can reason in Mentalese just as one can in English, and surely such non-linguistic reasoning needs markers of structure (like 'so' and 'but') just as reasoning expressed in natural language does? In response one could simply deny that this is so. Procedural markers such as 'but' have evolved to help hearers identify discourse structure, but an individual thinker does not need such aids in order to know how his own reasoning is structured. One possible retort is that a thinker can engage in a mental dialogue with himself, taking first one side of an issue then another, and surely such mental dialogue would stand in need of structural markers? I do not deny that we often rehearse arguments in our minds. But such conscious thought processes are carried out in natural language – English if one is a monolingual speaker of English. And of course such conscious thinking can make use of words such as 'but' and 'so', since these are English words.

But the sort of pragmatic processes that are involved in language comprehension and production are not available to conscious awareness, and these unconscious 'inferential' processes involving the manipulation of Mentalese symbols do not stand in need of symbols that are analogues to natural language structural markers. If there were such symbols, they would just be more mental symbols in need of manipulation, and so the regress of

the representational and dispositional accounts of procedural knowledge, since even the dispositional view counts as representational in this stretched sense.

rules argument would apply straightforwardly to show that truly procedural elements cannot be embodied as mental symbols but must be part of the causal architecture of the system doing the reasoning.

A sixth objection is as follows. I have argued that procedural information is not part of the mental lexicon, since it is not stored in the mental lexicon either explicitly (in the form of a propositionally represented rule) or tacitly (in the causal structure of the mental lexicon). It is, I have claimed, stored tacitly as part of the causal structure of the performance system more generally. (This system may of course, as part of its functioning, need to access information stored in the mental lexicon). One might object: 'How could 'but' and its ilk not be a part of the lexicon when these are words in the language?'. Obviously I would not deny that 'but' is a word of the language. But I do deny that every word of the language is necessarily associated with an entry in the mental lexicon. Some words trigger procedures and some words are associated with concepts (and some words have dual aspects and so both trigger procedures as well as are associated with concepts). Only words associated with concepts have entries in the mental lexicon. One might respond that so long as I concede that 'but' is a word of the language, I must be committed to there being some place where a mental representation of that word is stored, and where else would that be but the mental lexicon? Yes, I concede there must be a word store, and for all I've said so far, there could be a single store for all words. What I am denying is only that all words have *entries* in the lexicon. In other words, on my picture, even if 'but' (or a representation of it) is stored in the mental lexicon, activating that word will not give access to a data structure. Instead, a disposition to process data structures in a certain way will be triggered.

This leads to a seventh objection. I've said that the role of some words is to trigger procedures, not to give access to concepts in the mental lexicon. So I am committed to there being two classes of words whose roles in processing are very different. But, the objection goes, it is implausible to claim that the roles of words are divided in this way. Implausible as it may seem to common sense, there nevertheless is a fair amount of empirical evidence that supports this view. Psycholinguists have long been interested in a distinction between content words and function words (or open-class versus closed-class words). This distinction does not perfectly line up with the conceptual/procedural distinction, but there is a good deal of overlap. Many of the words regarded by psycholinguists as function words (e.g. determiners, sentence connectives etc.) are given a procedural treatment

within RT.¹¹ The point is that there is evidence that content and function words are not accessed or processed in the same way. See Bradley (1978); Bradley et al. (1980); Friederici (1985); Garrett (1975, 1980, 1981); Koriat and Greenberg (1994, 1996); Joshi (1985); Park et al. (2002); Rapp and Caramazza (1997); Stemberger (1982).¹²

Not all of this work can be discussed here. I will briefly mention some of the findings of these researchers. There is evidence that function and content words participate in different sorts of speech errors. See Garrett (1975, 1980, 1981); Stemberger (1982). Bilingual code-switching data, discussed by Joshi (1985), also indicates that function and content words behave differently. Bradley (1978) found that in lexical decision tasks, where participants must decide whether a stimulus is a word or a non-word, there is a frequency effect for content words but not for function words. That is, there is a significant difference in reaction time on the decision task between high- and low-frequency content words but this difference is not significant for function words. Bradley also found that content and function words behave differently in tests of interference effects in judgments of non-words. Participants are slower to reject non-words when these begin with strings that correspond to real words. (e.g., people are slower to reject TOASTLE than to reject POASTLE). Bradley found this interference effect for content words, but not for function words. Bradley et al. (1980: 278) use such findings to hypothesize that the organization of word-retrieval mechanisms is different for function and content words. This in turn is used to explain the particular sorts of language performance deficits found in people suffering from Broca's aphasia. (1980: 281). Broca's aphasics have problems with the functional aspects of language, both in language production and comprehension, although their ability to use content words seems

¹¹One way in which the content/function distinction fails to line up with the conceptual/procedural distinction is that prepositions such as 'down', 'under' etc. are treated by psychologists as function words, whereas these would fall on the conceptual side of the conceptual/procedural divide, since they clearly encode (fairly rich) conceptual content. Prepositions such as 'down' are more like words such as 'loves', which stand for real world relations, than they are like 'but' or 'after all', which indicate something about the information structure of discourse. Friederici (1985) is aware of this problem, and she is careful to distinguish between what she calls lexical prepositions and obligatory propositions. These are distinguished precisely by the fact that the former have conceptual content, whereas the latter play a role in identifying the structure of sentences (and presumably also the structure of whole discourses).

¹²There is also work that suggests that content and function words are stored and accessed similarly. See Schmauder (1996); Schmauder et al. (2000). However, Schmauder and her collaborators included prepositions such as 'down' as examples of function words. But as mentioned in note 11, prepositions such as these encode concepts not procedures. Hence the work by Schmauder and her colleagues does not give a pure picture of how truly procedural words such as 'but' are stored and accessed.

relatively spared. This pattern of dissociation would be explained if content and function words are associated with two distinct recognition routines, one of which is impaired in Broca's aphasics.

Friederici (1985: 135) notes that several attempts to replicate Bradley's frequency effects have been unsuccessful. Nevertheless, Friederici accepts Bradley's claim that there is a processing distinction between function and content words. Friederici thinks that this distinction will not be revealed in the sorts of tasks that Bradley used, in which words were presented in isolation. Instead, Friederici presented function and content words in semantically constraining versus semantically neutral sentential contexts. Participants were given a booklet of target words (one word per page) and then they heard a series of sentences. For each auditory presentation (one corresponding to each target word) participants were instructed to signal as soon as they heard the target word. Friederici found that for normal (i.e. non-impaired) listeners, there was a significant context effect for content words, but not for function words. In other words, for content words, listeners were quicker to identify target words in semantically constraining as opposed to semantically neutral contexts. There was no such benefit for function words. Friederici takes this as evidence that content and function words operate at different processing levels. Function words play a role at the structural level in determining the frame within which content words will be processed, and hence they are relatively insensitive to semantic context.

This conclusion is bolstered by the performance of a group of agrammatics (e.g. those suffering from Broca's aphasia) on Friederici's word monitoring task. Like normal listeners, agrammatics showed a context effect for content but not for function words. But unlike normal listeners, their monitoring times for function words were slower than their monitoring times for content words. Friederici takes this to support the claim by Bradley et al. (1980) that in Broca's aphasics, mechanisms for the retrieval and processing of function words are impaired.

Koriat and Greenberg (1994, 1996) found that in a letter detection task, participants were more likely to skip letters in function words than in content words. Moreover, this is not attributable to the fact that function words are more frequent and/or predictable, so that they tend to be skipped over in reading. Koriat and Greenberg were able to rule out explanations in terms of frequency by using a feature of Hebrew. In Hebrew, function morphemes can appear as prefixes. By using content words that had the same initial letters as words with function prefixes, but where the letters were part of the word stem rather than a prefix, Koriat and Greenberg were able to show that the missing letter effect is greater for function words than for content words. That is, letters are skipped more often when they are part of function prefixes than when those same letters are part of stems of content words.

(These content words were matched for length and frequency with the prefixed words). In English, some words can play a function role in one context and a content role in another. For instance, compare the role of 'on' in 'on his way' with its role in 'on-switch'. Koriat and Greenberg have found that letter skipping is more likely in contexts in which such words are playing a function role.

Koriat and Greenberg argue that their findings support a structural view of text processing. The role of function words is to set a structural frame into which concepts accessed by content words will subsequently be slotted. They argue that once function words have played their role, their representations fade into the background, which explains the missing letter effect. Letter detection is harder for words with degraded representations. They also suggest that the structure revealed by function words is a tool that helps focus the meanings (the concepts) that are accessed via content words. These conceptual meanings are the primary objects of attention and memory (1996: 1194). This is compatible with the view argued for in this paper, where words that encode procedures are treated differently from words that encode concepts. Procedural information is stored as part of the causal architecture of the performance system, and so is indeed a 'tool'. It is what drives interpretation rather than being an object of interpretation.

I would add that my dispositional account of procedural encoding bolsters Koriat and Greenberg's explanation for their missing letter effect. If words like 'but' encode procedures rather than concepts, and procedures do not correspond to propositional representations, then the only representation associated with a word such as 'but' is a word-level representation. Thus, once processing has proceeded to the deeper, conceptual level, the word-level representation will drop out of the picture, and so presumably will fade. Schmauder et al. (2000) found evidence in the eye movement record during reading that readers are more likely to reread function words than content words. This would be explained if the only representation for a function word is the word-level one. Such representations would need to be continually refreshed, since they are not maintained and reinforced by a connection to a conceptual representation that is the object of attention and memory, as is a content word representation.

5 Objections to the Parallelism of the Semantic/ Pragmatic, Conceptual/ Procedural and Competence/ Performance Distinctions

One might object that by locating procedural encoding in the pragmatic sphere and locating conceptual encoding in the lexical-semantic sphere my

view commits me to two separate modules, a pragmatics module and a semantics one. However, a commitment to a pragmatics module leads to difficulties with another claim of mine, namely the claim that procedural information is part of the linguistic performance system, whereas conceptual information is part of our linguistic competence. Prince (1985, 1988, 1997) argues that there is a special purpose linguistic-pragmatics module, which interfaces between the language system proper and the system that controls discourse-level processes. This pragmatics module has its own proprietary store of pragmatic knowledge that is declaratively represented and that constitutes what Prince calls the speaker-hearer's *discourse competence*. If we accept the idea of a pragmatic discourse competence, then we will have to concede that the parallelism between the declarative/procedural, competence/performance, and semantics/pragmatics distinctions fails.

But it is by no means clear that we should accept Prince's arguments for the existence of a special purpose linguistic-pragmatics module. Prince argues that in order to explain how people know which linguistic forms are the appropriate ones to use in certain contexts, they must be supposed to have some sort of pragmatic competence. We cannot explain how people are able to make these form-function choices simply by appeal to their knowledge of meaning plus general conversational principles, since cross-linguistic evidence shows that very different linguistic forms serve the same function. Which forms serve which functions is something arbitrary and language-specific, and hence must be specially learned.

For example, in English we use It-clefts to focus certain content, and to put other content into the background. Consider:

(1) The Israelis claimed that it was they who found Eichmann.

(1) presupposes that someone found Eichmann, and so can only be felicitously used in a context where it is shared knowledge that Eichmann has been found. On the other hand, what is in focus position is new information. However, in Yiddish, a completely different syntactic structure is used to focus content:

(2) ... dos hobn zey gefunen aykhmanen

... this have they found Eichmann

This *dos*-construction is used to functionally focus the pronoun 'zey' ('they'). As Prince says: 'What marks the special function of this construction is the presence of a sentence-initial *dos* which is not an argument of the verb.' (Prince, 1988: 169) It seems that in Yiddish this construction is contrasted with another construction that is syntactically parallel except that it has a sentence-initial 'es' ('it'). This *es*-construction has a different discourse function. It is used when 'the fewest assumptions about shared knowledge are warranted' (1988: 169).

Prince's main point is that there is nothing iconic about any of these constructions (in either English or Yiddish) that would indicate what function they have. Neither can their function be predicted from knowledge of meaning together with knowledge of general conversational principles, such as the Gricean maxims. These constructions must be specially learned. For example, English speakers must learn principles such as 'If you want to focus content, use the It-cleft construction.' Hence the knowledge as to which forms to use for which functions must be declaratively represented and be part of a speaker-hearer's linguistic competence. Moreover, since this is knowledge of language *use* it is pragmatic knowledge, and so is a part of *pragmatic* competence (rather than semantic competence).

However, the same sorts of arguments that were used in Section 3 apply here and show that even if some form-function knowledge is encoded in the linguistic-pragmatics module in the form of principles or rules, not all such knowledge can be represented in this way, since this would lead to a regress of rules. Some such knowledge must be procedural, i.e. tacitly represented in the causal structure of the linguistic-pragmatics module, rather than explicitly represented in whatever proprietary mental code is used by this pragmatics module.

Moreover, it is not clear that Prince has established that *any* form-function knowledge (i.e. knowledge of which form to use for which function) is declarative knowledge that is represented in the proprietary code of some special purpose linguistics-pragmatics module. First, since the sort of knowledge that Prince is concerned about is very similar to the procedural knowledge that relevance theorists and others have described (see Blake-more 1992: 142-145), it is more parsimonious to reject the idea of a special-purpose linguistic-pragmatics module and instead see the special knowledge that Prince is talking about as knowledge that is tacitly stored in the performance system. So there *is* something special about the knowledge that Prince is pointing us to, but what is special about it is that it is procedurally encoded in the performance system, not that it is declaratively represented in a special-purpose linguistics-pragmatics module.

Second, form-function knowledge is not something that we need to invoke just to explain a small set of facts, such as facts about how different languages handle the focusing of content. Probably no pair of languages solves all form-function problems in exactly the same way. For instance, not all languages have definite and indefinite articles (e.g. Russian and Chinese do not). If we assume that the function of the definite article in English is to signal that the following information is already familiar in the context, whereas the indefinite article signals that the following information is new, then Russian and Chinese speakers must be able to signal this familiar/new contrast in some other way. And of course they are able to signal this in a

different way. Similarly, Chinese has no distinction between mass and count nouns. So the mass/count distinction must be signaled in a different way than it is in English. And of course it is. Examples like this can be multiplied. The point is that learning a language is acquiring a system of linguistic forms that will constrain in various ways how one is able to convey certain information. It is not as though one learns various forms and then has to learn which of these forms one should use to perform which function. One learns the form *as* the way to perform a certain function. So form-function knowledge cannot belong to a module interfacing between the language system proper and the system handling discourse processes. It doesn't make sense to think that we can process a certain form and only then get to wonder about what function it is intended to perform. Once again, it seems more parsimonious to conceive of procedural/functional elements as being hardwired into the performance system as ways of processing forms, rather than as declarative rules specifying which forms have which functions.

So, I do not think that we should accept Prince's claim that there is a special purpose linguistic-pragmatics module, and so we do not as yet have a reason to give up the parallelism between the declarative/procedural, competence/performance, and semantics/pragmatics distinctions.

However, the objection can be pushed from the opposite direction. Instead of trying to argue that procedural knowledge is declaratively represented in a pragmatics module (and so is a part of our 'discourse competence'), one can try to argue that procedural knowledge is part of our semantic competence, my assertions to the contrary notwithstanding. After all, one might protest, it is difficult to accept that knowledge of the meaning of 'table' is part of our linguistic competence, whereas the knowledge associated with 'but' is not. Chierchia raised this line of criticism against my view in his comments at the workshop where my paper was originally presented.

Chierchia writes: 'there are a lot of aspects of language that are best understood in terms of what one would want to call pretheoretically "a procedure" and yet we want to say that they are robustly part of our competence. A case in point (of the many that can be made) is that of negative polarity items (NPIs).' (Chierchia, p.c.). NPIs are items such as 'any', which are only grammatical in downward entailing contexts, such as negative contexts or the antecedents of conditionals. NPIs are ungrammatical in positive contexts. This pattern of licensed uses is illustrated for 'any' below:

- (3) There isn't any food in the fridge.
- (4) If there is any food in the fridge, we can eat at home tonight.
- (5) *There is any food in the fridge.

Chierchia goes on to spell out what the ‘procedural’ meaning of ‘any’ might be, and offers a very compelling and elegant hypothesis in this regard. This claim of Chierchia’s is something that does not challenge my views and that I can (and do) accept. After all, it is a working assumption of my view that some words encode concepts and some encode procedures. That ‘any’ belongs to the latter class is something I can readily agree to.

Chierchia’s real challenge comes when he goes on to list some characteristics of the behavior of NPIs. Chierchia (p.c) writes:

The behaviour of NPIs has the following characteristics:

- a. It has to do with the distribution of a certain class of morphemes
- b. It shows up, in some form, in every language
- c. It is subject to parametric variations (e.g. in how large is the class of licensors in a particular language, whether Free Choice uses are admitted etc.)
- d. It is acquired early by native learners
- e. It interacts systematically with other aspects of core grammar (like locality on extraction, agreement, etc. – cf. the discussion on so called ‘Negative Concord’)

The properties in [a. – e.] are typical of core grammar, competence based phenomena (if anything is).

I am unable to completely address Chierchia’s challenge (especially points c. and e. above), as I am not a trained linguist. However, I believe the main thrust of the challenge can be met. It is compatible with my view to accept that procedural elements such as NPIs show up in some form in all languages, are subject to parametric variations, and are acquired early by native learners. Even though procedural items on my view correspond to causal dispositions rather than to mental symbols, they must be learned. Treating procedural knowledge as part of the causal architecture of the performance system does not preclude one from claiming that learners of all languages acquire such knowledge early. In fact, given the crucial structural role played by procedural elements in language production and comprehension, these causal dispositions would have to be in place early in language development.

Moreover, my view of procedural information is not incompatible with ideas of innateness and parametric variation across languages. It could be that aspects of the causal architecture of the language performance system are innate, and that the fine-grained aspects of this causal structure will only be set once the language learner is exposed to particular linguistic input.

This sort of fine-tuning of the performance system could be a matter of parametric variation across language learners, depending on what sorts of linguistic environments they are initially exposed to. (However, my view is incompatible with Chomsky's ideas about innateness, inasmuch as he claims that our innate grammatical competence is completely independent of our communicative system. I claim, on the contrary, that aspects of the performance system are also a part of our innate language endowment. Hence, language development involves a significant interplay between core grammar and the performance system).

The first point Chierchia raises (that the behavior of NPIs has to do with the distribution of a certain class of morphemes) I take to be asserting that procedural items have a certain pattern of use and that this pattern is systematic or rule-governed. I can accept that procedural items have systematic uses, and would appeal once again to the regress of rules argument to establish that talk of rules must ultimately be cashed out in terms of causal dispositions. (See also the fourth objection and my reply in Section 4 above. This is an objection that Chierchia raises against my use of the regress of rules argument).

The fifth point Chierchia raises is the most tricky for me to handle, since I do not fully understand it. It appears to be saying that procedural elements must be part of a language user's semantic competence (and thus be represented as concepts in the mental lexicon) because procedural elements interact with other aspects of core grammar. A further assumption seems to be that these 'other aspects' are themselves represented as mental symbols in the lexicon. However, at least some of the 'other aspects' that Chierchia lists seem to belong to the procedural side of the conceptual/procedural divide (e.g. agreement phenomena, such as number, gender and case markings). Hence it begs the question to assert that these procedural elements are all represented as symbols in the mental lexicon, rather than embodied in the causal architecture of the performance system. But even if these 'other aspects' are all symbolically represented in the mental lexicon, their interaction with procedural elements that are tacitly represented in the causal architecture of the performance system is not ruled out. In fact, interaction is what procedural elements are all about. In the course of the exercise of these causal dispositions, concepts in the lexicon will be accessed (i.e. interacted with). Thus I do not agree that Chierchia's challenge shows that procedural knowledge (e.g. that embodied in NPIs) is part of our semantic competence.

The final objections that I will discuss are ones implicit in the writings of relevance theorists. We saw in Section 2 that relevance theorists claim that procedural items (e.g. determiners such as 'the' or discourse markers such as 'but') encode instructions that constrain the inferential phase of

comprehension, and are involved in the derivation of both explicatures and implicatures. Thus procedural items guide interpretive processes, such as the pragmatic enrichment of semantically underspecified terms, and the interpretation of metaphorical and ironic utterances. But they are not themselves subject to processes such as pragmatic enrichment. As Vicente (2003) puts it: ‘discourse markers, ... etc. code instructions that constrain the inferential phase of comprehension at both the explicit and implicit levels, but are themselves impervious to contextual contamination, *which places them neatly within the domain of linguistic semantics.*’ (My emphases). The objection to my view is implicit in the italicized phrase. Since procedural items are context-independent, Vicente claims they belong in the domain of linguistic semantics. This contradicts my claim that they belong to the domain of pragmatics.

However, it does not follow from the fact that an item is not subject to inferential modulation or contextual contamination that the item belongs in the domain of linguistic semantics. I accept that conceptual meanings belong to the domain of linguistic semantics and are context-invariant.¹³ But I deny that all things that are context-invariant belong to the domain linguistic semantics. On the contrary, procedural knowledge as understood in this paper is also something context-invariant (in the synchronic sense – see note 13). It constitutes the stable mental architecture that enables language users to focus on the contentful aspects of verbal communication. To use language one must rely not only on certain invariant meanings but also on certain stable mental structures. The latter are what ground our procedural abilities.

A second objection is implicit in Carston (2002: 11). Carston notes that while some have argued that the semantics/ pragmatics distinction lines up with Chomsky's competence/ performance distinction, this is not the position adopted within relevance theory. Instead, according to Carston, the semantics/ pragmatics distinction corresponds to the distinction between the decoding and inferential phases of communication. For instance, to comprehend a speaker's utterance, the hearer must first decode the words the speaker uses and then use the decoded information, along with non-linguistic (e.g. encyclopedic) information, to infer the pragmatically enriched meaning that the speaker intended to convey. What is recovered via decoding belongs to semantics, whereas the pragmatically inferred meanings belong to pragmatics. My main concern has been the distinction between information that is conceptually and procedurally *encoded*. Thus, it

¹³They are context-invariant in the synchronic sense that they do not change across contexts at a time. Of course, over time an individual's lexical concepts may be changed (e.g. refined or broadened).

seems, I must concede that procedural information lies on the decoding side of the decoding/ inference divide, and hence on the side of semantics rather than pragmatics.

In response I would question the idea that the decoding/ inference distinction lines up with the semantics/ pragmatics distinction. In particular, I would challenge the idea that decoding processes are all alike. Decoding does not always give access to a conceptual entry in the mental lexicon. Sometimes, decoding triggers a procedure. In Section 4, I cited empirical evidence in support of the claim that items that encode concepts and items that encode procedures are not accessed or processed in the same way. If decoding is not a unitary process, then we are not forced to give all decoded information a uniform, semantic treatment.

An objection that combines aspects of the previous two is implicit in Blakemore's writings. After giving some examples involving lexical items such as 'so' and 'moreover', Blakemore (1992: 46) writes:

However, while it is true that the meanings of these expressions play a role in the way that contextual assumptions are used in the interpretation of the utterances that contain them, there is no reason for saying that this role is itself pragmatically determined. The knowledge of what these words mean is linguistic knowledge, and, as we have seen, this is quite different from the non-linguistic knowledge that a hearer brings to bear on the interpretation of an utterance. In other words, the meanings of these expressions can be represented in pragmatics only at the expense of the modular view of cognition...

The objection would be that by placing procedural knowledge in pragmatics, I have failed to acknowledge its essentially linguistic character and its difference from the sort of non-linguistic, encyclopedic knowledge that is used in utterance interpretation. Presumably the idea is that all encoded information (even that associated with procedural items) belongs to the language module and is accessed via a process of linguistic decoding. Non-linguistic, encyclopedic knowledge in contrast is not confined to a module, but is part of the central system and is processed inferentially.

I would counter this objection by pointing out that on my view procedural knowledge *is* distinguished from encyclopedic knowledge, even though both fall on the side of pragmatics. The latter is conceptual knowledge, whereas the former is tacitly stored in the causal architecture of the performance system. Moreover, my view doesn't threaten the idea of a language module. On the contrary, I have been assuming that there is a language module and that the concepts entered in the mental lexicon are part of this module. And I agree that these concepts are accessed via a decoding procedure. But I disagree that all decoding processes access concepts. In some cases they trigger procedures, and these procedures are not strictly

part of the language system. Their role is to guide an interaction between something that belongs to the language system (lexical concepts) and something that lies outside that system (encyclopedic and other non-linguistic knowledge).

I conclude that there are as yet no compelling reasons for giving up the parallelisms that were postulated in the introduction. That is to say, we have as yet no reason to give up on the claim that procedural knowledge belongs to the language performance system and is pragmatic, whereas lexical conceptual knowledge is declaratively represented and constitutes a speaker's semantic competence. So there is not anything simultaneously procedural and semantic.¹⁴

References

- Andersen, G., and Fretheim, T. eds. 2000. *Pragmatic Markers and Propositional Attitude*. Amsterdam: John Benjamins.
- Asher, N., and Lascarides, A. 1996. Lexical disambiguation in a discourse context. *Lexical Semantics: The Problem of Polysemy*, eds. J. Pustejovsky and B. Boguraev, 69-108. Oxford: Clarendon Press.
- Asher, N., and Lascarides, A. 1998. The semantics and pragmatics of presupposition. *Journal of Semantics* 153: 239-99.
- Bezuidenhout, A. 1997. Pragmatics, Semantic Underdetermination and the Referential/Attributive Distinction. *Mind* 106: 375-409.
- Blakemore, D. 1987. *Semantic Constraints on Relevance*. Oxford: Blackwell.
- Blakemore, D. 1988. 'So' as a constraint on relevance. *Mental Representations: The Interface Between Language and Reality*, ed. R. Kempson, 183-95. Cambridge: Cambridge University Press.
- Blakemore, D. 1992. *Understanding Utterances*. Oxford: Blackwell.
- Blass, R. 1990. *Relevance Relations in Discourse: A Study with Special Reference to Sissala*. Cambridge: Cambridge University Press.

¹⁴I would like to thank Claudia Bianchi and Carlo Penco for inviting me to participate in the Workshop on Context 2002 and to thank Claudia for encouraging me to publish this paper. Thanks also to Genarro Chierchia, who was assigned as the commentator on my paper at the workshop, probably because I earlier indicated that I would be discussing Genarro's work on dynamic semantics. However, I took an early fork in the road and it led somewhere completely different. But Genarro rose to the occasion and provided penetrating criticisms at the workshop, as well as extensive written comments later. I know that I will not have answered his criticisms to his satisfaction, but I've done my best. Thanks also to my colleagues Raffaella De Rosa and Otavio Bueno, who let me do a trial run of my workshop presentation and gave me valuable feedback. Finally, special thanks go to Begonia Vicente, who read an early version of my manuscript and subjected it to rigorous scrutiny. Although I'm sure the final product will still dissatisfy her, the paper is much improved due to her attentions.

- Bradley, D. 1978. *Computational Distinctions of Vocabulary Type*. Doctoral dissertation, MIT.
- Bradley, D., Garrett, M. and Zurif, E. 1980. Syntactic deficits in Broca's aphasics. *Biological Studies of Mental Processes*, ed. D. Caplan, 269-86. Cambridge, MA: MIT Press.
- Carston, R. 1997. Enrichment and loosening: Complementary processes in deriving the proposition expressed? *Linguistische Berichte* 8: 103-27.
- Carston, R. 2002. *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.
- Chomsky, N. 1996. Language and problems of knowledge. *The Philosophy of Language* 3rd ed., ed. A. Martinich, 558-77. Oxford: Oxford University Press.
- Copetake, A., and Briscoe, T. 1996. Semi-productive polysemy and sense extension. *Lexical Semantics: The Problem of Polysemy*, ed. J. Pustejovsky and B. Boguraev, 15-67. Oxford: Clarendon Press.
- Coulson, S. 2001. *Semantic Leaps: Frame-Shifting and Conceptual Blending in Meaning Construction*. Cambridge: Cambridge University Press.
- Dennett, D. 1981. Brain writing and mind reading, *Brainstorms*, 39-50. Brighton: Harvester Press.
- Fauconnier, G. 1985. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge, MA: MIT Press.
- Fauconnier, G. 1997. *Mappings in Thought and Language*. Cambridge: Cambridge University Press.
- Fauconnier, G., and Sweetser, E. eds. 1996. *Spaces, Worlds, and Grammar*. Chicago: University of Chicago Press.
- Fodor, J. 1975. *The Language of Thought*. New York: Crowell.
- Fretheim, T. 1998a. Intonation and the procedural encoding of attributed thoughts: The case of Norwegian negative interrogatives. *Current Issues in Relevance Theory*, ed. V. Rouchota and A. H. Jucker, 205-36. Amsterdam: John Benjamins.
- Fretheim, T. 1998b. A relevance-theoretic account of *if* and *in case*. *Proceedings of the 14th Eastern States Conference on Linguistics*, ed. J. Austin and A. Lawson, 58-69. Ithaca: CLC Publications.
- Friederici, A. 1985. Levels of processing and vocabulary types: Evidence from on-line comprehension in normals and agrammatics. *Cognition* 19: 133-66.
- Garrett, M. 1975. The analysis of sentence production. *The Psychology of Learning and Motivation: Advances in Research and Theory*, vol. 9, ed. G. Bower, 133-77. New York: Academic Press.
- Garrett, M. 1980. Levels of processing in sentence production. *Language Production*, ed. B. Butterworth, 177-220. London: Academic Press.
- Garrett, M. 1981. Production of speech: Observations from normal and pathological use. *Normality and Pathology in Cognitive Functions*, ed. A. Ellis, 19-76. London: Academic Press.

- Giora, R. 1997. Discourse coherence and theory of relevance: Stumbling blocks in search of a unified theory. *Journal of Pragmatics* 27: 17-34.
- Grice, P. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Iten, C. 1997. Because and although: A case of duality? *UCL Working Papers in Linguistics* 9: 55-76.
- Iten, C. 1998. The meaning of although: A relevance theoretic account. *UCL Working Papers in Linguistics* 10: 81-108.
- Iten, C. 2002. Even if and even: The case for an inferential scalar account. *UCL Working Papers in Linguistics* 14: 119-156.
- Koriat, A. and Greenberg, S. 1994. The extraction of phrase structure during reading: Evidence from letter detection errors. *Psychonomic Bulletin and Review* 1: 345-56.
- Koriat, A. and Greenberg, S. 1996. The enhancement effect in letter detection: Further evidence for the structural model of reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22: 1184-95.
- Kripke, S. 1982. *Wittgenstein on Rules and Private Languages*. Oxford: Basil Blackwell.
- Lakoff, R. 1971. Ifs, ands and buts about conjunction. *Studies in Linguistic Semantics*, ed. C. J. Fillmore and D. T. Langendoen, 115-50. New York: Holt Rinehart and Winston.
- Levinson, S. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.
- MacDonald, M., Pearlmutter, N., and Seidenberg, M. 1994. Lexical nature of syntactic ambiguity resolution. *Psychological Review* 101: 676-703.
- Nicolle, S. 1995. Conceptual and procedural encoding: Criteria for the identification of linguistically encoded procedural information. Paper presented at the University of Hertfordshire Relevance Theory Workshop, Hatfield Peverel, UK.
- Nicolle, S. 1998. *Be going to* and *will*: a monosemous account. *English Language Linguistics* 2: 223-43.
- Park, G., McNeil, M. and Doyle, P. 2002. Lexical access rate of closed-class elements during auditory sentence comprehension in adults with aphasia. *Aphasiology* 16: 801-814.
- Prince, E. 1985. Fancy syntax and 'shared knowledge'. *Journal of Pragmatics* 9: 65-81.
- Prince, E. 1988. Discourse analysis: a part of the study of linguistic competence. *Linguistics: The Cambridge Survey* vol. 2, ed. F. J. Newmeyer, 164-82. Cambridge: Cambridge University Press.
- Prince, E. 1997. On the functions of left dislocation in English discourse. *Directions in Functional Linguistics*, ed. A. Kamio, 117-43. Amsterdam: John Benjamins.
- Pustejovsky, J. 1995. *The Generative Lexicon*. Cambridge, MA, USA: MIT Press.

- Pustejovsky, J. 1998. The Semantics of Lexical Underspecification. *Folia Linguistica* 32: 323-47.
- Pustejovsky, J., and Boguraev, B. eds. 1996. *Lexical Semantics: The Problem of Polysemy*. Oxford: Clarendon Press.
- Rapp, B. and Caramazza, A. 1997. The modality-specific organization of grammatical categories: Evidence from impaired spoken and written sentence production. *Brain and Language* 56: 248-86.
- Schmauder, R. 1996. Ability to stand alone and processing of open-class and closed-class words: Isolation versus context'. *Journal of Psycholinguistic Research* 25: 443-81.
- Schmauder, R., Morris, R. and Poynor, D. 2000. Lexical processing and text integration of function and content words: Evidence from priming and eye fixations. *Memory and Cognition* 28: 1098-1108.
- Sperber, D., and Wilson, D. 1995. *Relevance: Communication and Cognition* 2nd ed. Oxford: Blackwell.
- Thomas, J. 1995. *Meaning in Interaction*. London: Longman.
- Vicente, B. 2003. Meaning in relevance theory and the semantics/ pragmatics distinction. ms.
- Wilson, D. 1998. Discourse, coherence and relevance: A reply to Rachel Giora. *Journal of Pragmatics* 29: 57-74.
- Wilson, D., and Sperber, D. 1993. Linguistic form and relevance. *Lingua* 90: 1-25.