

Agreement Patterns in Shona Locative Inversion

Benjamin Baxter

Abstract

Certain Bantu languages such as Shona and Lubukusu display what appears to be an optional agreement strategy between either the logical subject appearing to the right of the verb, or the preposed locative. Such optionality is troublesome for a principled theory of agreement. Following work by Diercks (2011), the present investigation asserts that what appears optional on the surface is actually derived from one of two underlying structures. Predictions made by this theory are clearly demonstrated to be correct and the resulting analysis shows that despite their surface behaviour, languages such as Shona and Lubukusu in fact adhere to the single agreement strategy found in a large majority of languages of agreement with a c-commanding NP.

Keywords: Bantu, agreement, locative, inversion

1 Introduction

The behaviour of subject agreement in Indo-European and Bantu languages shows a clear divergence in instances instance of locative inversion. Whereas in Indo-European Languages, agreement is persistently controlled by the logical subject, in Bantu languages for the most part, the verb displays agreement with the constituent at the front of the sentence. The disparity is easily demonstrated by the examples in (1) and (2).

- (1) a. In the swamp was/*were found a child.
b. In the swamps was/*were found a child.
c. In the swamp *was/were found the children.
d. In the swamps *was/were found the children.

- (2) a. Ku-mu-dzi ku-nabwera a-lendo
17-3-village 17-came 2-visitor
"To the village came visitors."

b. *Ku-mu-dzi pa-nabwera a-lendo
17-3-village 16-came 2-visitor
c. *Ku-mu-dzi a-na-bwer-a a-lendo
17-3 -village 2-came 2-visitors

(Bresnan & Kanerva, 1989, p.9)

The difference in agreement patterns above has led to the identification of a certain level of parameterization of the agreement function in Universal Grammar. The parameters are outlined by Baker (2008), as below:

- (3) The Case-Dependency of Agreement Parameter:
F agrees with DP/NP only if F values the case feature of DP/NP or vice versa.
(4) The Direction of Agreement Parameter:
F agrees with DP/NP only if DP/NP asymmetrically c-commands F.

The intention is that the agreement patterns in (1) and (2) can be accounted for by the application of (3) and (4) respectively. That is to say that in Indo-European languages, the agreement relation

between the verb and the agreeing DP is reliant on the checking of uninterpretable case features. With this as the primary restriction, the probe can look either upwards or downwards in search of its goal. In contrast, for Bantu languages, the agreement relation depends only on an NP residing in the correct structural position (SpecTP in Baker’s analysis).

This parameterization is an attractive proposal by virtue of the fact that it provides a unified analysis for the agreement patterns of Bantu and Indo-European languages in that, at a deep level of the grammar, the language is set to one of the two parameters. However, there is something troublesome in the agreement patterns of certain Bantu languages. Locative Inversion constructions in languages like Lubukusu and Shona appear to licence both upward agreement with the preposed locative and downward agreement with the postposed logical subject in a seemingly optional paradigm.

(5) Khu-si-kulu kw-a-biringikha-kho ku-m-pira [Lubukusu]
 17-7-hill 3-PST-roll-17L 3-3-ball
 “Down the hill rolled the ball.”

(6) Mu-nju mw-ola-mo ba-ba-ana [Lubukusu]
 18-home 18-PST.arrive-18L 2-2-child
 “At home arrived the children.”

(Diercks, 2011)

The above examples demonstrate an apparent optionality between the strategies in (3) and (4) in the context of locative inversion. Such observations are clearly problematic as they flout the constrained nature of the parametrization theory.

The data presented here for Lubukusu has been drawn from an analysis by Diercks (2011) which convincingly argues for an underlying structure which only surfaces as an apparent optionality, but in fact conceals two differing constructions which adhere to the parameters of Baker (2008). The Shona data however has been previously argued by Harford Perez (1983) to be an example of ‘impersonal’ concord where locatives do not actually control agreement in cases of inversion. Harford Perez’s claims are refuted here in favour of an analysis that aligns with that of Diercks.

2 Shona Syntax

2.1 Class Agreement in Shona

Shona possesses a three way distinction in locative noun class that bear slightly different semantic interpretations. Class 16 *pa-* is used in an indicative locative sense, as though the location is being pointed to. Class 17 *ku-* is taken to mean the more general status of being ‘at a location’ though it can also be used to refer to a location distant from the speaker. Finally, class 18 *mu-* is used in the sense of being contained within a certain location. These prefixes are applied as pre-prefixes to nouns which are already coded for nominal class.

(7) Pa-chi-toro a-pa
 16-7-store that-16

(8) Ku-chi-koro i-chi
 17-7-school this-7

(9) Mu-rw-izi u-mu
 18-11-river that-18

This doubling of the prefix allows for the curious optionality between class of agreement on the nominal modifier. Therefore, the modifier following the noun can agree with the locative prefix as in (7) and (9), or it can take concord of the internal nominal prefix as in (8). On consulting with a native speaker, the semantic difference between such constructions was reported to be slight. Therefore, in examples like (8), when agreeing with the locative the NP was taken to refer to things happening at the school as a physical location. Instead, when agreement aligned with the internal prefix, the meaning was taken as referring to occurrences at the school as an organisation.

As for the wider status of agreement in the language, Fortune (1955) outlines tables for each noun class and the variants of concord appearing on each verbal element. In this way, there can be shown to be seven possible agreeing constituents: *Adjective*; *demonstrative*; *number*; *quantifier*; *possessive*; *subject*; and *object*. To this list can further be added the relative marker which appears on the verb and must agree with any non- subject relative (Zentz, 2016). Of most significance to the present study is that of subject agreement.

2.2 Shona Locative Inversion

A typical Shona declarative phrase follows the canonical SV ordering. The verb in such a phrase bears a prefix that is commonly referred to as the subject marker. This prefix co-varies in ϕ -features with the subject nominal as in (10):

- (10) V-ana va-notamba pa-chikoro
 2-children 2-play 16-school
 “The children play at school.”

In such a phrase, it is not possible for the verb to agree with any other NP than the sentence initial subject. However, there exist inverted constructions in which the subject concord on the verb can vary between possible candidates. Therefore, in the Locative Inversion derived from the declarative phrase in (10), the verb can either bear concord of the preposed locative NP or postposed logical subject.

- (11) Pa-chikoro pa-notamba v-ana
 16-school 16-play 2-children
 “At school the children play.”

- (12) Pa-chikoro va-notamba v-ana
 16-school 2-play 2-children
 “At school the children play.”

This optionality between agreement presents the same problems as the Lubukusu data discussed previously. The examples in (11) and (12) appear to show a parameterization between upward and downward agreement language internally. In the course of the following sections then, it will be argued that the apparent divergence from the Upward Agreement Hypothesis in Shona in fact adheres to it once some basic syntactic processes are assumed for the derivation of Shona Locative Inversion.

3 Locative Inversion Agreement as Impersonal Concord

3.1 Subject status of locatives

The first test which Harford Perez applies to the Shona data is that of reflexivization. Assuming, as Harford Perez does, that subjects are universally available for relativization, it can be rather easily demonstrated that preposed locative NPs behave in the same way as preverbal logical subjects in their

availability to relativisation and the manner in which this takes place. This is demonstrated with reference to the example in (13).

- (13) Ku-mu-sha u-ko ku-nogara va-nhu va-zhinji ku-no mu-sika
 17-3-village that-17 17-live 2-people 2-many 17-have 3-market
 “At that village, where many people live, there is a market.”

Harford Perez notes that the locative in this example patterns with the relativization of a logical subject with “a low tone on the subject prefix of the relative verb, and no relative prefix” (1983). Whilst Harford Perez concedes that such data indicate subject-like behaviour of locative NPs, this does not rule out her differing analysis.

Data from passivisation is also included by Harford Perez with further concession to the observation that the phenomenon offers no evidence in support of locative NPs not being subjects. Indeed, passivisation is available to logical subjects, objects and locative NPs all the same.

One observation of syntactic behaviour that is deemed by Harford Perez as vindicating of her analysis is the prohibition of “conjoinability” or formation of a coordinate structure. Specifically, this is the ruling out of a structure in which a preposed locative NP is conjoined with the preverbal logical subject of an identical verb phrase. Harford Perez notes that these structures are perfectly licit for preverbal logical subjects as in (14).

- (14) Zvi-rongo ne zvi-yo zvi-noteng-es-wa mu-mariketi
 8-pots with 8-grain 8-buy-CAUS-PASS 18-market
 “Pots and grain are sold at the market.”

(Harford Perez, 1983, p. 137)

It is not possible to construct an example of a coordinate structure in which a preverbal logical subject is conjoined with a preposed locative NP. Harford Perez explains this impossibility by demonstration of a complementary distribution between preposed locative NPs and preverbal logical subjects. The asymmetry is illustrated with the following examples.

- (15) a. Mu-komana Ø-akavata
 1-boy 1-slept
 “The boy slept.”
 b. *Mu-munda m-akavata
 18-field 18-slept
- (16) a. Mu-munda u-mu m-akavata m-ombe
 18-field that-18 18-slept 10-cattle
 “In the field slept cattle.”
 b. *Mu-komana Ø-akavata m-ombe
 1-boy 1-slept 10-cattle

(Harford Perez, 1983, p.137)

Whilst Harford Perez does not attempt to construct the relevant conjoined phrases and notes simply that the inability for locatives and logical subjects to appear in the same positions as in (15) and (16) demonstrates a prohibition for conjunction. This is shown by an ungrammatical judgement for a version of (14) in which a canonical declarative phrase is conjoined with a locative inverted phrase.

- (17) a. *Zvi-rongo ne mu-mariketi zvi-noteng-es-wa zvi-yo
 8-pots with 18-market 8-buy-CAUS-PASS 8-pots
- b. *Zvi-rongo ne mu-mariketi mu-noteng-es-wa zvi-yo
 8-pots With 18-market 8-buy-CAUS-PASS 8-pots

Examples (17a) and (17b) are included to show that the subject agreement on the verb does not have an effect on the grammaticality of the phrase. This is despite the fact that Shona makes use of a conflict resolution between conjoined subjects of different nominal classes whereby the subject agreement is with the NP of higher animacy (Hawkinson & Hyman, 1974).

This asymmetry between the availability of conjunction between logical subjects and the prohibition of conjunction between logical subject and preverbal locative is taken by Harford Perez as an indication that preposed locative NPs do not behave as one would expect of a subject, and as such challenges the notion of their controlling subject concord.

However, the examples in (15) and (16) do not show complementary distribution so much as an incompatibility. In (15b), what makes the construction illicit is not the presence of the locative in a preverbal position, but rather the lack of a postposed logical subject. Indeed the example in (16a) shows that when the postposed logical subject is present, a locative NP is grammatical in preverbal position. Likewise, in (16b), the prohibition of the phrase is not due to any positional restriction, but rather that there are two logical subjects of the verb. Furthermore, in Anderson (1976) from where Harford Perez takes the conjunction diagnostic, the stipulation is that coordination well-formedness between two phrases relies on “a shared chunk of material common to them both.” This is clearly not the case in the examples above. If anything, the examples in (15) and (16) only demonstrate that locatives cannot bear the same semantic role with the verb: an issue which is definitely not in contention and is the reason for differentiating the postposed element as the *logical* subject. Furthermore, locatives are perfectly licit in coordination constructions with other locative NPs, showing that syntactically, they still pattern with logical subjects:

- (18) Ku-chikoro ne ku-mba ku-nodzidza v-ana
 17-school with 17-home 17-learn 2-children
 “At school and at home children learn.”

3.2 Distributional and Morphological Evidence

Turning to distributional and morphological behaviour, Harford Perez (1983) argues that locatives behave differently in their apparent control of verbal concord than do logical subjects. Principally, Harford Perez notes that verbal concord can appear on the verb when there is no locative NP present:

- (19) Pa-ne mu-nhu Ø-akamirira ku-ku-ona
 16-have 1-person 1-wait to-you-see
 “There is someone who has been waiting to see you.”

According to Harford Perez (1983), the example above demonstrates that locative concord is not tied to the presence of a locative NP. This motivates her analysis of the agreement as an impersonal concord. While she entertains the theory that the locative NP could have been elided just as in examples of pro-drop with the logical subject, Harford Perez argues that there is no remnant locative reading available. Whilst such a conclusion is tenable with these observations, the data in fact are contradictory to the intuitions of a native speaker informant to this study who states that the locative reading that someone has been waiting *outside* is in fact mandated. Pro-drop is a highly productive feature of Shona and can proceed with any NP when licensed by agreement morphology on the verb

(Mugari, 2013). Therefore, with the locative interpretation available, Shona can be analysed as patterning with the Chichewa data discussed earlier in allowing pro-drop of agreeing preposed locative NPs.

There is one possibility for a verbal agreement with a null expletive if one assumes the analysis of Demuth and Mmusi (1997) and Marten (2006) that the class 17 prefix is ambiguous between a locative and expletive reading. It does not follow however that all cases of locative concord are subjectless. Indeed it does not mean to say that even uses of class 17 concord is impersonal. Rather the interpretation is that when there is class 17 agreement on the verb and no locative present in the phrase, the reading can be said to be ambiguous between locative and expletive.

Harford Perez then states that there are in fact verbs which take locative concord, but do not permit the presence of a preposed locative.

- (20) Ku-nofungirwa kuti Sekuru va-ngu ibenzi
17 -suspected that uncle 2-my fool
“It is suspected that my uncle is a fool.”

(Harford Perez, 1983)

These facts however are also in alignment with predictions made by the stipulation that the class 17 prefix is ambiguous between locative and expletive reference. Examples of such constructions with all three of the locative prefixes would be needed to argue for their total independence from locative NP control. In consultation with a native speaker informant, such examples were not grammatical. The informant reported that the specific verb used in (20) was not licit with any other locative prefix. Furthermore, the informant could not come up with any verbs which allow a choice between prefixes, but prohibit the appearance of a preverbal locative. These facts are in accordance with the view of the class 17 prefix as ambiguous between locative or expletive.

In addition to locative concord without the presence of a locative NP, Harford Perez notes that locative concord is also available when the locative NP appears post-verbally.

- (21) Kw-akasvika va-nhu ku-danga
27-arrive 2-people 17-cattle.pen
“People arrived at the cattle pen.”

(Harford Perez, 1983)

Such facts are indeed curious if one assumes that the locative is in its base generated position. Under such assumptions it does become unlikely that the locative is controlling agreement on the verb. Of course, as before, it is possible to maintain that the class-17 prefix is ambiguous between locative and expletive reading, but it should be noted that such constructions were deemed grammatical by this study’s informant with any of the three locative prefixes. Therefore, phrases such as (22) are perfectly acceptable.

- (22) Pa-nodzidza v-ana pa-chikoro
16-learn 2-children 16-school
“The children learn at school.”

Whilst this behaviour does seem to license an interpretation of locative concord as impersonal concord, I will argue that the sentence final locative is not in a base generated position but is structurally higher in the phrase than the verb and is only linearised to its right. This will be taken up later.

3.3 Mis-match Between Locative and Verbal Concord

One final oddity that is mentioned in Harford Perez (1983) but is also left unexplained is the existence of an apparent mismatch between the class of prefix on the locative NP and the locative agreement prefix on the verb. Harford Perez demonstrates three examples of this.

- (23) Mu-nguva i-yoyo kw-akazvarwa mw-ana
18-9time 9-that 17-was born 1-child
“At that time was born a child”

- (24) Mu-nguva ya -ma-Dzviti p-akanga pa-si-na Ø-aimbotipwe
18-9time 9of-6-Ndebele 16-was 16-NEG-have 1-said.dry-up
“In the time of the Ndebele (incursions) there was no one to put a stop to it (say dry-up)”
(Chakaipa, 1958, cited in Harford Perez, 1983)

- (25) Kw-abudwa pa-nze nhasi
17-came.out 16-outside today
“There has been come outside today” (someone has come outside)
(Fortune, 1955, cited in Harford Perez, 1983)

The examples in (23), (24) and (25) are indeed troublesome for a theory of verbal concord controlled by locative NPs. Interestingly, Harford Perez (1983) does not incorporate such constructions into her theory despite their being perhaps the strongest evidence of all that verbal concord and locative noun class act independently of one another. Instead her position is that despite their being independent, locative NPs influence the selection of impersonal concord by “providing an obvious choice among the three locative classes available”. Still, the data here presents a problematic anomaly for a theory of agreement with the locative NP. As can again be mentioned here, the examples in (23) and (25) align with earlier identification of the class 17 prefix as expletive. (24) however shows a direct mismatch between two prefixes which are undoubtedly locative. Such constructions are certainly rare and could not be generated by the informant to this study when asked to do so. In reference to the example in (24) despite the mismatch being perceived as odd, the informant offered that the different semantics of the prefixes could be at play, specifically; the mu- prefix used to mean “in” and the pa- prefix meaning “at.” There is perhaps a possibility that the specific spatial reference allowed for a semantics to override normal agreement techniques. However, the issue must be set aside until more data can be found to confirm the existence of the phenomenon and necessitate an analysis.

The arguments presented by Harford Perez (1983) are used to argue against the subject status of proposed locative NPs. The main empirical claim however is that locative NPs do not control verbal concord in the manner of the logical subject. Such a stance is not tenable from the tests discussed above and there exists a more comprehensive analysis of verbal agreement in Diercks (2011) which can easily be applied to Shona and explain both logical subject and locative agreement equally.

4 Aligning Shona with Lubukusu

4.1 Comparisons of Shona and Lubukusu Locative Inversion

Diercks’ (2011) study of Lubukusu reveals some similarities between the language and Shona in the domain of Locative Inversion. As such, it is productive to align the two analyses. As for the characteristics of Lubukusu itself, the relevant features can be summarised here.

The crucial point with regards to the Lubukusu comparison with Shona is the apparent presence of a downward agreement relation between the verb and postposed NP constituents. This phenomenon is labeled “disjoint agreement” by Diercks and contrasts with “repeated agreement” in which both the subject marker and locative clitic agree with the preposed locative NP.

(26) Mu-mu-siiru kw-a-kwa-mo ku-mu-saal [Disjoint Agreement]
 18-3-forest 3-PST-fall-18LC 3-3-tree

“In the forest fell a tree.”

(27) Mu-mu-siiru mw-a-kwa-mo ku-mu-salaa [Repeated Agreement]
 18-3-forest 18-PST-fall-18LC 3-3-tree

“In the forest fell a tree.”

As with Shona, the apparent downward agreement observed in disjoint agreement phrases is at odds with the parameterising of agreement processes. This contradiction however is dealt with by Diercks’ analysis of a different underlying structure for the two constructions in (26) and (27). The structures are presented by Diercks as below:

(28) Disjoint Agreement Locative Inversion
 [CP LocC-V [TP Subj...[Subj V Loc]]]

(29) Repeated Agreement Locative Inversion
 [TP Loc T-V [VP Subj V Loc]]

What is crucial in the above analysis is that the upwards agreement parameter is preserved under both disjoint and repeated agreement constructions. This is shown by the necessity of the NP that triggers verbal agreement to occupy the Specifier of TP position. Whichever element occupies this position controls agreement on the verb regardless of the surface word ordering. Therefore, when the subject marker co-varies with the noun class of the postposed logical subject, this is because the subject NP remains in SpecTP whilst the locative and verb raise to SpecCP and C respectively.

Conversely, when the subject agreement is with the preposed locative NP, this is triggered by the movement of the locative to the SpecTP position, thereby blocking the logical subject from appearing there. Thus the logical subject remains VP internal. As is commonly assumed for Bantu generally (Diercks, 2012; Demuth & Harford, 1999; Ngonyani, 1998) and Shona specifically (Zeller, 2009; Aranovich, 2015) the verb consistently undergoes V-to-T movement. Therefore, with the logical subject remaining VP internal, and the verb raising to T, the word order is inverted, thus accounting for this crucial characteristic of Locative Inversion.

Such an analysis makes clear predictions which can be tested for both Lubukusu and Shona. Diercks (2011) makes use of the extraction morphology of Lubukusu to demonstrate that in disjoint agreement constructions in the context of relativisation or raising, agreement morphology on the verb must pattern with the extraction of non-subject constituents. In contrast, with repeated agreement examples of extraction, the verb patterns with typical subject extraction.

Whilst Shona and Lubukusu do not share the same methods of extraction marking, extraction marking on the verb in Shona nevertheless patterns differently with respect to subjects and non-subjects. As mentioned by Harford Perez (1983), Shona subject extraction prohibits the presence of a relative marker on the verb. However, with non- subject extraction, such agreeing relative markers are obligatory (Zentz, 2016). Therefore, if in cases of Locative Inversion with locative agreement the locative NP sits in SpecTP, we expect that relativisation with an agreeing relative marker is ruled out. This is indeed the case.

- (30) Nda-karara ku-chikoro (*kwa-) ku-nodzidza v-ana
 1-slept 17-school (*17REL-) 17-learn 2-children
 “I slept at the school at which children learn.”

Conversely, when relativisation occurs for instances of Locative Inversion in which the verb agrees with the postposed logical subject, the relative marker is presumed to be necessitated. This again is what is observed.

- (31) Nda-karara ku-chikoro *(kwa-) va-nodzidza v-ana
 1-slept 17-school *(17REL-) 2-learn 2-children
 “I slept at the school at which children learn.”

These facts clearly align with the predictions made by an analysis of the locative NP occupying SpecTP when it triggers agreement and sitting in some other position when agreement is with the postposed logical subject.

More predictions can be made with regards to the position of the logical subject. If the logical subject can be clearly diagnosed as remaining VP internal, then we should expect that in those cases, the verb cannot bear agreement with the postposed logical subject. Just such a case presents itself in the paradigm of wh-question formation. In Shona, wh-questioned subjects are prohibited from appearing preverbally except in embedded clauses (Zentz, 2016). As wh-question formation is usually tied to the application of focus, the requirement of wh-subjects to appear postverbally is therefore in line with the IAV interpretation of Hyman and Polinsky (2010) which claims that the subject remains in SpecvP. With this in mind, we can therefore predict that in cases of locative inversion in which the postposed logical subject is questioned, the verb should be prohibited from agreeing in class morphology with the wh-subject as it could not have occupied the SpecTP position at any point in the derivation. This again is exactly what is observed:

- (32) Ku-chikoro kw-akasvika v-ana va-pi?
 17-school 17-arrived 2-children 2-which
 “Which children arrived at the school?”

- (33) Ku-chikoro v-akasvika v-ana va-pi?
 17-school 2-arrived 2-children 2-which

As for the position of the verb in the respective constructions, Demuth and Harford (1999) provide a compatible analysis of relativisation in Shona. They propose that in non- subject relative clauses, the presence of a relative clitic in the complementiser head triggers raising of the verb to C and a subsequent inversion of the verb and subject. The interpretation given by Demuth and Harford (1999) is that the relative clitic does not meet minimal phonological word requirements and that the verb raises to C as a repair strategy. Conversely, in instances when there is a fully lexical complementiser used, this movement of the verb is blocked. This interpretation therefore explains the availability of a preverbal wh-subject in embedded clauses as the complementiser head is occupied by the lexical relativiser ‘kuti’ and thus blocks movement of the verb to C. Given that the analysis defended in this study posits that the preposed locative occupies the SpecCP position and the verb bearing logical subject agreement is in C, we should expect such constructions to be ruled out if a fully lexical complementiser is occupying the complementiser head. In elicitations of such constructions from an informant, the phrases pattern exactly as expected.

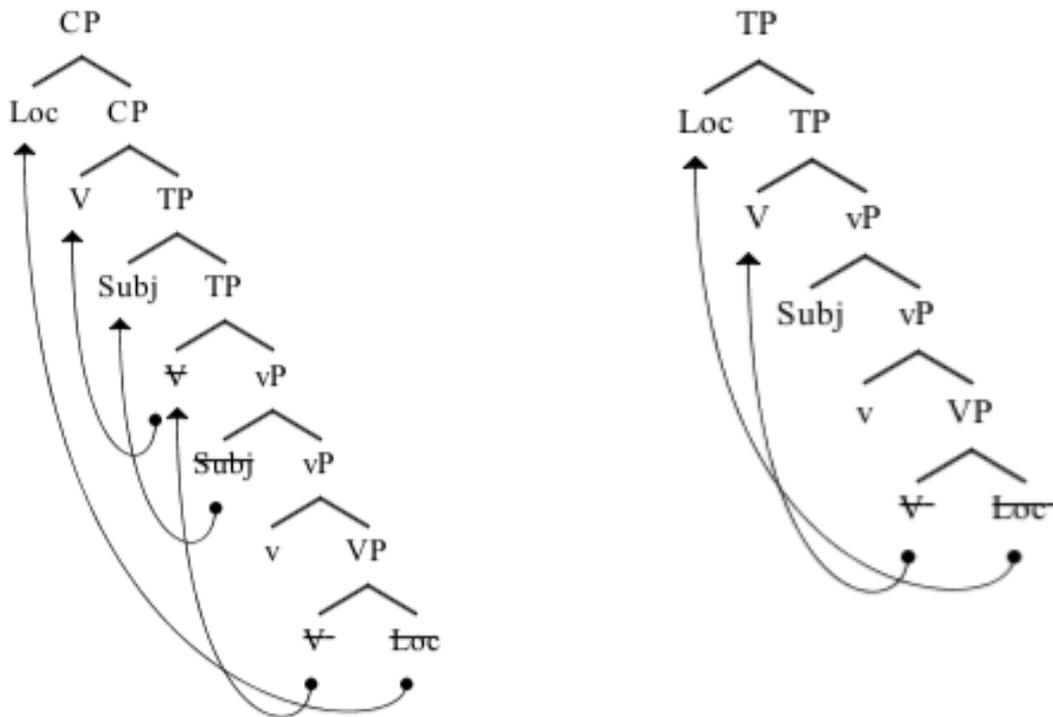
(34) Nda-karara ku-chikoro kuti v-ana va-dzidze
 1-slept 17-school that 2-children 2-learn
 "I slept at the school so that children can learn."

(35) *Nda-karara ku-chikoro kuti va-dzidze v-ana
 1-slept 17-school that 2-learn 2-children

The above observations align neatly with the theory for Lubukusu in Diercks' (2011) and are further fortified by the motivation for inversion in relativised contexts outlined in Demuth & Harford (1999). However, if the interpretation is that the inversion is the *result* of relativisation, this does raise the question of why there exist non-relativised instances of Shona Locative Inversion in which the verb agrees with the postposed logical subject but does not carry a relative marker. Nevertheless, if we must set aside the motivation for inversion of verb and logical subjects for further research, we can still accept that in Shona Locative Inversion with postposed logical subject agreement on the verb, the verb behaves as if it occupies the complementiser head. Therefore it is possible to draw up the representations of the two agreement patterns in Shona as in the diagrams below.

(36) Agreement with postposed logical subject

(37) Agreement with preposed locative



There is however one issue which is not covered by the theory in its present form. That is the ability of locative phrases to control verbal concord whether appearing sentence initially or finally. Recall that this evidence was used by Harford Perez (1983) to argue against the control of verbal agreement by the locative under the assumption that the locative is in its base generated position low down in the phrase. However, if we maintain that verbal concord is only available to constituents in SpecTP, it must be the case that what appears to be a base generated position for the locative NP is in fact a structurally higher right-dislocated position. With this in mind, there is a simple test which can be implemented in the divination of right-dislocated status: postverbal word order. In Shona, the ordering of postverbal adjuncts such as locative and temporal phrases is fixed. Therefore, in the declarative sentence in (38), the temporal adjunct follows the locative NP.

- (38) V-ana v-anodzidza pa-chikoro nhasi.
 2-children 2-learn (16-school) today

“The children are learning at school today.”

However, when verbal concord is with the locative class, we see that the locative cannot appear in its base generated position, but instead appears to the right of the temporal adjunct.

- (39) Pa-nodzidza v-ana (*pa-chikoro) nhasi pa-chikoro
 16 -learn 2-children (16-school) today 16-school

“The children are learning at school today.”

Right-dislocation is certainly a productive feature in Bantu languages (Zerbian, 2006; Halpert & Zeller, 2015) and is shown to be a possible construction for right-edge wh-question formation in Shona by Zentz (2016). Therefore, what may appear to be a structurally low phrase within the VP, is actually in a position c-commanding the verb, thus allowing for proper upward agreement and only being linearised to the right of the phrase on the surface.

4 Conclusion

The data from Shona Locative Inversion presented here pose an interesting problem for the parameterized theory of agreement as laid out by Baker (2008). The agreement relations that surface in these constructions run counter to the claims that a language makes use of one of two strategies for agreement in that Shona appears to show an optional choice between both upwards and downwards agreement. An established theory of agreement in Shona Locative Inversion by Harford Perez (1983) was argued to be both lacking in its account of observed phenomena and insufficient in its compatibility with cross-linguistic theories of agreement relations. It was shown that the agreement marker on the verb in Shona is too closely linked to the class of preposed locative to allow for any other interpretation than direct control in the manner of typical preverbal subjects. Instead, the behaviour of agreement in Shona Locative Inversion has been shown to be directly compatible with Diercks’ (2011) theory of Lubukusu agreement. With reference to distributional effects and morphological extraction marking, the locative element can be shown to behave exactly as the logical subject in its control of verbal concord. Therefore, Shona can be considered an adherent of Baker's positional agreement parameter, and that apparent divergence from this can be shown to arise from one of two possible underlying structures.

References

- Anderson, S. R. (1976). On the notion of subject in ergative languages. *Subject and topic*, 1, 23.
 Aranovich, R. (2015). Realizational morphology beyond the lexicon: synthetic and periphrastic verbs in Shona. *Language Sciences*, 51, 54-66.
 Baker, M. C. (2008). *The syntax of agreement and concord* (Vol. 115). Cambridge: Cambridge University Press.
 Birner, B. (1995). Pragmatic constraints on the verb in English inversion. *Lingua*, 97(4), 233-56.
 Bresnan, J., & Kanerva, J. M. (1989). Locative Inversion in Chicheŵa: A Case Study of Factorization in Grammar. *Linguistic Inquiry*, 20(1), 1-50.
 Bresnan, J. (1994). Locative inversion and the architecture of universal grammar. *Language*, 72-131.

- Carstens, V. (2011). Hyperactivity and hyperagreement in Bantu. *Lingua*, 121(5), 721-741.
- Cheng, L. L. S., & Downing, L. J. (2009). Where's the topic in Zulu? *The Linguistic Review*, 26(2-3), 207-238.
- De Kind, J., & Bostoen, K. (2012). The applicative in ciLubà grammar and discourse: A semantic goal analysis. *Southern African linguistics and applied language studies*, 30(1), 101-124.
- Dembetembe, N. C. (1976). The Syntax of Sentential Complements in Shona. Doctoral dissertation, School of Oriental and African Studies.
- Demuth, K., & Harford, C. (1999). Verb raising and subject inversion in comparative Bantu. *Journal of African Languages and Linguistics*, 20(1), 41-61.
- Demuth, K., & Mmusi, S. (1997). Presentational focus and thematic structure in comparative Bantu. *Journal of African Languages and Linguistics*, 18, 1-20.
- Diercks, M. (2011). The morphosyntax of Lubukusu locative inversion and the parameterization of Agree. *Lingua*, 121(5), 702-720.
- Diercks, Michael. (2012). Parameterizing Case: Evidence from Bantu. *Syntax*, 15(3), 253-286.
- Ethnologue. (2016). Lubukusu, retrieved from: <http://www.ethnologue.com/language/bxk>.
- Ethnologue. (2016). Shona, retrieved from: <http://www.ethnologue.com/language/sna>.
- Fortune, G. (1955). *An Analytical Grammar of Shona*. London: Longmans, Green and Co.
- Halpert, C. (2012). Argument licensing and agreement in Zulu. Doctoral dissertation, Massachusetts Institute of Technology.
- Halpert, C., & Zeller, J. (2015). Right dislocation and raising-to-object in Zulu. *The Linguistic Review*, 32(3), 475-513.
- Harford Perez, C. (1983). Locative pseudo-subjects in Shona. *Journal of African languages and linguistics*, 5(2), 131-156.
- Harford Perez, C. (1986). *Aspects of complementation in three Bantu languages*. Reproduced by the Indiana University Linguistics Club.
- Hawkinson, A. K., & Hyman, L. M. (1974). Hierarchies of natural topic in Shona. *Studies in African linguistics*, 5(2), 147.
- Hyman, L. M., & Polinsky, M. (2010). Focus in Aghem. In M. Zimmerman & C. Fery (Eds), *Information Structure: Theoretical, Typological, and Experimental Perspectives* (pp. 206-233). Oxford: Oxford University Press.
- Levin, B., & Hovav, M. R. (1995). *Unaccusativity: At the syntax-lexical semantics interface* (Vol. 26). MIT press.
- Marten, L. (2006). Locative Inversion in Otjiherero: More on morpho- syntactic variation in Bantu. *ZAS Papers in Linguistics*, 43, 97-122.
- Marten, L., & van der Wal, J. (2014). A typology of Bantu subject inversion. *Linguistic Variation*, 14(2), 318-368.
- Mugari, V. (2013). Pronoun Dropping in ChiShona: A Comparative Approach. *Studies About Languages*, 23, 20-27.
- Ngonyani, D. (1998). V-to-I Movement in Kiswahili. In R.M. Beck, T. Geider, & W. Graebner (Eds), *Afrikanistische Arbeitspapiere 55, Swahili Forum V* (pp. 129-144). Institut für Afrikanistik, Universität zu Köln Köln, Germany.
- Rizzi, L., & Shlonsky, U. (2006). Satisfying the subject criterion by a non subject: English locative inversion and Heavy NP shift. In M. Frascarelli (Ed), *Phases of interpretation* (pp. 341-361). Berlin: Mouton de Gruyter.
- Van der Wal, J. (2008). Agreement inthetic VS sentences in Bantu and Romance. In C. De Cat & K. Demuth (Eds), *The Bantu-Romance connection. A comparative investigation of verbal agreement, DPs, and information structure* (pp. 323-350). Amsterdam: John Benjamins.
- Wu, H. H. I. (2008). Generalized inversion and theory of agree. Doctoral dissertation, Massachusetts Institute of Technology.
- Zeller, J. (2013). Locative inversion in Bantu and predication. *Linguistics*, 51(6), 1107-1146.
- Zentz, Jason. (2016). Forming Wh-Questions in Shona: A Comparative Bantu Perspective. Doctoral dissertation, Yale University.
- Zerbian, S. (2006). Inversion structures in Northern Sotho. *Southern African Linguistics and Applied Language Studies*, 24(3), 361-376.

Subjective Containment*

Harris Constantinou, Hans van de Koot

Abstract

This article suggests that subjective categories, including certain epistemic modals, evaluative adverbs, metalinguistic comparatives and outside negation, must take scope over non-subjective ones, such as quantified NP, tense, question and focus operators. This requirement captures previously noted containment effects (von Fintel & Iatridou (2003) and subsequent work), but also gives rise to additional scope freezing effects not previously discussed. The latter type of data also provide an argument for the view that subjective categories have a truth-conditional impact, contra previous work (e.g. Halliday, 1970; Palmer, 1986). We argue that the full range of effects necessitates an analysis in which scope shift is encoded through percolation of scope indices constrained by a minimality condition (as in Neeleman & Van de Koot (2012)).

Keywords: quantification, subjectivity, epistemic modality, focus, scope freezing

1 Introduction

Among the many restrictions on the scope of quantified NPs (QPs) discovered in recent years, von Fintel and Iatridou (2003) have contributed the observation in (1), the Epistemic Containment Principle, which is concerned with the relative scope of QPs and epistemic modals:

- (1) *Epistemic Containment Principle (ECP)*¹
A quantifier [(i.e. QP)] cannot have scope over an epistemic modal.
(von Fintel & Iatridou, 2003, p. 174).

The ECP accounts for sentences like (2) and (3), in which a QP is unable to scope over auxiliary verbs of epistemic possibility and necessity, respectively.^{2,3} In each example, the continuation forces a wide scope reading of the QP, which – as the ECP leads us to expect – results in a contradiction.

- (2) #Every student may have left, but not every one of them has.
(**every* > *may*)

* Parts of the material in this article have been presented at MIT (NELS,45), University of Göttingen (GLOW39), Humboldt-Universität zu Berlin (Workshop on Questions, Answers and Negation), University of Ulster and the Hungarian Academy of Sciences (the Budapest Conference on Projection and Representation in Syntactic Theory). We thank the audiences for their helpful comments and criticism. We would also like to thank Patrick Elliott, Ad Neeleman, Floris Roelofsen and Jacopo Romoli for their valuable input, as well as our group of informants.

¹ The ECP or an approximation of it is also observed in other works, including Leech (1971), McDowell (1987), Picallo (1990), Brennan (1997), Drubig (2001) and Butler (2003).

² We avoid the combination of quantificational items with the same quantificational force in order to facilitate the diagnosis of scope (un)ambiguities. For instance, we do not combine *every* with modals of epistemic necessity, such as *must*.

³ Sentences with QPs and other types of modals (e.g. deontic *must*) or temporal elements (e.g. *will*) behave as expected, that is, they present scope ambiguities. See von Fintel and Iatridou (2003) for discussion.

- (3) #Fewer than half the students must have passed, but perhaps all of them did.
 (*fewer than half > must)

The ECP effect is by no means restricted to the combination of a QP and an epistemic modal auxiliary. As shown in (4), the epistemic adverb *probably* also limits the scope of the QP *every student*. Example (5) shows that the effect is also manifested with an epistemic raising predicate like *certain*, whereas (6), a Dutch version of (4), illustrates that the phenomenon is not restricted to English.

- (4) #Every student has probably left, but not everyone has.
- (5) *Context: John is hiring employees and has just interviewed 20 applicants.*
 #John is certain to have hired fewer than half the applicants, but perhaps he hired all of them.
- (6) #Iedere student is waarschijnlijk vertrokken, maar niet iedereen is
Everyone student is probably left but not everyone is
 vertrokken.
left
 ‘Every student has probably left, but not everyone has left.’

Since the ECP was proposed by von Stechow and Iatridou (2003), several authors have argued that exceptions to it relate to whether epistemic modals are interpreted subjectively or objectively (Tancredi, 2007; Huitink, 2008; Anand & Hacquard, 2009). The subjective interpretation of epistemics corresponds to the invocation of less widely accepted evidence in the evaluation of the proposition in the scope of the epistemic. In particular, on the subjective use of an epistemic this proposition is evaluated with respect to the speaker’s personal belief state, and hence her full commitment to this evaluation is highlighted. By contrast, the objective interpretation of an epistemic category relies on evidence that is widely accepted by the relevant community (Lyons, 1977).⁴ “What differentiates a subjective modal statement from others is that it is one that cares only about the speaker’s subjective belief state. All other modals involve something additional. An objective, metaphysical, modal statement for example is based on speaker-external facts” (Tancredi, 2007, p. 2); that is, the relevant knowledge and beliefs of the discourse participants are taken into consideration when evaluating such a statement.

A statement whose evaluation is dependent on other speakers’ mental models, in particular their knowledge and beliefs, which in turn depend on the state of the world, could in principle be false. On the other hand, a subjective statement, whose interpretation solely depends on the rational speaker’s mental model, must always be true. This could imply that subjective categories make no truth-conditional contribution, as has previously been suggested for (subjective) epistemic modality (e.g. Halliday, 1970; Palmer, 1986; Drubig, 2001), whereas objective categories do. Since epistemic modals can be either subjective or objective, this position is at odds with standard semantic treatments of modality (see, for

⁴ Lyons (1977) discusses the following example to illustrate the different interpretations of *may*:

- (i) It may rain tomorrow.

On the subjective (layman’s) reading of the modal in (i), the speaker reasons on the basis of private concerns, whereas on its objective reading (for example when (i) is uttered by a meteorologist), the evaluation is understood to be based on reliable, complete and accessible scientific evidence.

instance, Kratzer (1991)). Here, we contribute to this discussion by demonstrating (in section 0) that subjective categories, including subjective epistemic modals, do have truth-conditional effects as they scopally interact in non-trivial ways with other intra-clausal scopal elements (see Papafragou (2006) for additional arguments).

The above observations impact on the evaluation of the ECP. In out of the blue contexts (see (7a)), epistemics tend to default to the subjective interpretation and exhibit ECP effects. Indeed, (7a), when rendered with neutral stress, only allows an interpretation where the modal out-scopes the universal QP (i.e. *It is possible that every party guest is the murderer*) and this results in infelicity (unless there was a collective murder). However, as shown in (7b), it is possible to provide a context that strongly favours the objective interpretation, and this in turn gives rise to obviation of the ECP effect.

- (7) a. *Every party guest might be the murderer. (**every*>*might*)
 b. Given the currently available evidence/Objectively speaking, every party guest might be the murderer. (✓*every*>*might*)

The literature identifies three further exceptions to the ECP, illustrated in (8) - (10).

(8) Each girl might be in love with John, but some of them aren't. (Hacquard, 2006)

(9) EVERY party guest might be the murderer. (Anand & Hacquard, 2009)

(10) *Context: Same as (5).*

Fewer than half the applicants are certain to have been hired, but perhaps all of them were (hired).

(Constantinou & Van de Koot, 2015)

In (8) a strongly distributive universal QP is able to out-scope the (subjective) epistemic modal. Similarly, in (9) a prosodically prominent universal QP can out-scope the same modal. (10) shows that a QP like *fewer than half the applicants* can out-scope the epistemic raising predicate *certain* when it raises across it (cf. (5)).

In what follows we argue that the ECP-related facts reviewed above are in fact unrelated to the epistemic nature of the categories involved and should be attributed exclusively to their subjective interpretation. In other words, we propose to replace the ECP with the more general Subjective Containment Principle:

(11) *Subjective Containment Principle (SCP)*

A category carrying a subjective interpretation must take scope over other operators in the proposition with which the subjective attitude is connected.

The main evidence supporting the SCP comes from the existence of containment effects in sentences involving speaker-oriented *non*-epistemic operators, such as evaluative adverbs (Ernst, 2009), metalinguistic comparatives (Giannakidou & Yoon, 2011) and outside negation (Ladd, 1981). These facts are presented in section 0. In section 0 we show how these data can be captured once the SCP is combined with the theory of scope put forward in Neeleman and Van de Koot (2012). In fact, this theory predicts that subjective categories should trigger scope freezing effects that go well beyond the containment effects that motivated the SCP, and we show that these further predictions are borne out as well. In section 0 we turn to information-structural factors that affect the distribution of subjective categories. We first discuss structures in which an A'-moved operator blocks obligatory

scope extension of a category in its c-command domain. We show that, as predicted by the SCP, subjective categories cannot occur in the relevant environment. We then turn to the exceptional data in (8) and (9) and argue that their status is due to a conspiracy of independent, information-structural factors and should not be taken as presenting an exception to the ECP, or our SCP. Section 0 lists our main conclusions.

2 Subjective containment

2.1 Speaker-oriented adverbs

The subjective-objective distinction has also been invoked to explain various apparently non-containment-related facts about the distribution of epistemic modals; Papafragou (2006) and Ernst (2009) show that the occasional acceptability of an epistemic modal in the antecedent of a conditional and in questions correlates with the way that modal is interpreted. An epistemic adverb like *probably* can be interpreted objectively, and can thus occur in such environments (see (12a-b)), whereas an epistemic auxiliary verb like *must* lexically resists such an interpretation; that is, *must* is strongly subjective, hence the ungrammaticality of (13a-b).

- (12) a. Where have they probably put the loot? (=Ernst's (56b))
 b. If he will probably get drunk, I am not coming to the party.
 (=adapted from Papafragou's (33))
- (13) a. *Where must they have put the loot?
 b. *If he must get drunk, I am not coming to the party.

In (12a) the speaker-oriented adverb (henceforth SpOA) *probably* turns into an addressee-oriented one where the speaker is asking about the likelihood of the loot being in some location, given the evidence accessible to the addressee (i.e. *Where is it probable that they put the loot?*).⁵ *Probably* receives a similar interpretation in (12b) (i.e. *If it is probable that he will get drunk, I am not coming to the party*); the sentence is felicitous in a context in which the high likelihood of him getting drunk is based on some common knowledge (e.g. it is well-known among the interlocutors that he usually gets drunk at parties). As noted by Ernst (2009), subjective epistemics express the full commitment of the speaker in regards to the evaluation of the proposition *p* the epistemic scopes over. Put differently, SUB(jective) epistemic (*p*) (=Q) is true in all worlds in the speaker's belief set; OBJ(jective) epistemic (*p*) (=Q) on the other hand could in principle be at odds with the speaker's belief set (despite the available evidence). Assuming that questions and conditionals allow Q to be true in some worlds and false in others (Karttunen, 1977; Groenendijk & Stokhof, 1984), the facts in (12) and (13) directly follow. Subjective epistemics can occur in assertions because these are semantic entities that can be true in all worlds in the speaker's belief set.

In line with much other work (e.g. Papafragou, 2006; Anand & Hacquard 2009; Ernst, 2009), we assume that subjective categories quantify over possible worlds and adopt Ernst's (2009) semantics for such categories, where subjective categories are taken to characterize a relation between a proposition *P* and the speaker's belief set $M_B(s)$ (see also Papafragou

⁵ As noted by Kratzer (1991) and Høye (1997), epistemic modal adverbs (e.g. *probably*) and auxiliary verbs (e.g. *must*) tend to be subjective, whereas the corresponding adjectives (e.g. *probable*) and nouns (e.g. *probability*) are objective (with *certain* being the only exception we are aware of). We thus frequently use paraphrases with the adjectival/nominal versions of the adverb in question to highlight an objective interpretation.

(2006)). The semantics are specific to adverbs, but, as Ernst (2009) points out, they are potentially applicable to other categories (by replacing ‘Adv’ with ‘Aux’, for instance).

(14) *Subjectivity (for Speaker-Orientation)*

Where a speaker asserts $Q = \text{ADV}(p)$ (thus Q is in $M_B(s)$),

- a) ADV is subjective iff all worlds by which Q is evaluated are consistent with respect to $M_B(s)$ at the time of utterance;
- b) otherwise ADV is objective (Ernst, 2009, p. 516)

(15) Consistency: a set of worlds (q -worlds) is consistent with a belief state M if the proposition q is true both in q -worlds and in all the worlds in M .

(Ernst, 2009, p. 516)

In effect, the above semantics suggests that subjective categories must be true for the speaker’s entire belief set – the speaker brooks no possibility of the proposition $\text{SUBJ}(p)$ being false.

Whilst some epistemic modals (e.g. *probably*) can be either subjective or objective, Ernst (2009) identifies a class of SpOAs that are consistently interpreted subjectively, as well as a class of SpOAs that are consistently interpreted objectively. Strong evaluative adverbs such as *bizarrely*, *unbelievably* and *(un)fortunately* have a strong emotive flavor in that they express the speaker’s “extreme judgment of good or bad, or of surprise, astonishment, disbelief, or the like. [...] Emotions are of course highly subjective, and a strong emotional attachment to a given proposition does not depend on objective information” (p. 514). On the opposite end of the subjective-objective scale, evidentials like *apparently*, *clearly*, *evidently* and *obviously* “are objective, since their use depends on evidence that is either physically perceptible, or a matter of very easy, transparent inference from publicly available evidence. [...] modals are in the middle: They do not have a strong emotive flavor, nor is their use based on overt evidence to the same extent as evidentials. Thus [most] epistemic modal[s] [...] can be used either subjectively or objectively, as has frequently been noted (Lyons, 1977; Nuyts, 2001a; Papafragou, 2000, 2006; Kratzer, 2002).” (p. 514).

The subjective character of strong evaluatives is confirmed by the fact that they are never acceptable in the antecedent of a conditional or in questions; evidentials, on the other hand, are usually acceptable in these environments, a strong indication of their objective interpretation (compare (16) and (17)).

(16) a. *Has she unfortunately found a scorpion in her boot? (=Ernst’s (67))

b. *If she has unfortunately found a scorpion in her boot, she’ll be sorry.
(=Ernst’s (69))

(17) a. Are they {obviously/clearly} going to be eligible for the competition?
(=Ernst’s (52b))

b. If Allison has {obviously/clearly} completed her analysis, there’s no need for you to wait around.
(=Ernst’s (52c))

The generalized containment principle in (11) predicts that QPs will exhibit containment effects in the presence of (non-epistemic) strong evaluative SpOAs, because they are subjective, but not with evidential SpOAs, because they are objective. Indeed, (18) shows that strong evaluative adverbs like *unfortunately* and *sadly* forcibly take wide scope over a clausemate QP, whereas the same QP is able to outscope evidentials like *apparently*,

obviously and *clearly* in (19). The contexts in (18) and (19) force wide scope of the QP *fewer than half of the students* in the associated example. This is because in both cases, the context requires the predicate containing the speaker-oriented adverb to be read as distributing over a subset of the students. Thus, the subject cannot be included in the scope of the adverb. The contextual inappropriateness of (18) therefore indicates that the relevant reading ('Fewer than half of the students are such that they have unfortunately/sadly passed the exam') is inaccessible. By contrast, (19) does permit an analogous reading ('Fewer than half of the students are such that they have apparently/obviously/clearly passed the exam').

- (18) *Context: I had 30 students in my final year syntax class. To obtain their degree, they had to pass my exam. For 10 of these students, I considered it better if they failed as they were just not ready for the big world. After the exam had taken place, I quickly marked all the scripts and discovered that everybody had passed. So as far as I'm concerned...*

#Fewer than half of the students have {unfortunately/sadly} passed the exam.

- (19) *Context: I had 30 students in my final year syntax class. After the final exam had taken place, I asked my TAs to quickly mark the scripts. They have now marked all of the scripts, but I ask specifically how the first 10 students on the class list did, as I was quite worried that they wouldn't pass due to their poor performance on another test. They are saying that those 10 students have done really well. So ...*

Fewer than half of the students have {obviously/apparently/clearly} passed the exam.

A containment effect can also be observed with other combinations of QPs and strong evaluative adverbs, even in cases where the narrow scope reading of the QP is also compatible with the context (but pragmatically less appropriate); (20) cannot be assigned the reading in (ii). This reading would be facilitated by placing contrastive stress on *unbelievably/surprisingly* (since for the remaining students it was believable/unsurprising that they passed). The same QP can nevertheless be interpreted in this way (i.e. outscope the evidentials) in (21), rendering the sentence grammatical in the given context.

- (20) *Context: I had 30 students in my final year syntax class. To obtain their degree, they had to pass my exam. For 20 of them, I was pretty certain they would fail it, as they are weak students. After the exam had taken place, I quickly marked all the scripts and discovered that everybody had passed. So as far as I'm concerned...*

Most students have {unbelievably/surprisingly} passed the exam.

- (i) ✓ It is unbelievable/surprising that most students have passed the exam.
 (ii) *Most students are such that it is unbelievable/surprising that they passed the exam.

- (21) *Context: I had 30 students in my final year syntax class. After the final exam had taken place, I asked my TAs to quickly mark the scripts. They marked all of the scripts, but I ask specifically how the first 20 students on the class list did, as I was quite worried that they wouldn't pass due to their poor performance on another test. They are saying that those 20 students have done really well. So ...*

Most of the students have {obviously/apparently/clearly} passed the exam.

- (i) ✓ It is obvious/apparent/clear that most students have passed the exam.
 (ii) ✓ Most students are such that it is obvious/apparent/clear that they passed the exam.

2.2 Metalinguistic Comparatives⁶

This section provides further evidence for the SCP by demonstrating that a containment effect also arises when combining QPs with so-called metalinguistic comparatives (MCs).

MCs are a type of clausal comparative construction introducing the speaker's point of view towards a sentence, in particular her preference of one sentence in a given context over another. They are thus similar to other metalinguistic constructions (e.g. metalinguistic negation) in expressing the speaker's subjective evaluation of the appropriateness of a sentence. This aspect of interpretation is what distinguishes them from regular clausal comparatives, which are restricted to mere comparison of predicates/clauses to express an opinion. In their discussion of MCs, Giannakidou and Yoon (2011) show that languages such as Modern Greek and Korean employ distinct linguistic devices in the formation of the two types of clausal comparatives; MCs in Modern Greek, for instance, are realized with the preposition *para* 'than', which is lexically distinct from the 'regular' clausal comparative *apoti* (Giannakidou & Yoon, 2011). Consider (22) and (23), which differ in terms of the possibility of using *para* vs *apoti*.

- (22) Ta provlimata su einai perissotero oikonomika para/apoti nomika.
The problems yours are more financial than legal.
 'Your problems are financial rather than legal.' (McCawley, 1968)
- (23) Kalitera na pethano para/#apoti na ton pantrefto.
Better to die than to him marry.
 'I would prefer to die than marry him.' (Giannakidou & Yoon, 2011)

As pointed out by Giannakidou and Yoon (2011), the difference between (22) and (23) lies in whether they express an assessment of accuracy or appropriateness. Whereas (22) is understood to be expressing the speaker's view in relation to the nature of her addressee's problems, (23) is "not really about what the speaker believes or deems appropriate, but rather, [it] seems to express the speaker's dispreference towards the *para* proposition and its content, i.e. the course of action that this proposition implies" (Giannakidou & Yoon, 2011, p. 623). In this case, "the dispreference is so extreme that something obviously dispreferred in objective terms ('to die') is presented as more preferred than the action of the *para* sentence ('to marry him'), which, in objective terms, should be more preferred" (Giannakidou & Yoon, 2011, p. 623). On the basis of examples like (22) and (23) these authors conclude that as the clausal comparative becomes more preferential of an action, the choice between *apoti* and *para* becomes more constrained, and only *para* is used felicitously. Their conclusion is important for our purposes in two respects; first, a clausal comparative with *para* expresses the speaker's subjective attitude, in particular a personal *preference*, towards one sentence over another; crucially, the speaker's assessment is not epistemic in nature (Giannakidou & Yoon, 2011). Second, *apoti* comparatives can only be used felicitously when the speaker's assessment is more objective, that is, they (are intended to) express mere statements of regular comparative assessment on the basis of some concrete evidence. For instance, a speaker can go on and explain why they hold the view expressed by (22), but they cannot do so when uttering (23), because, in objective terms, they do not hold the view it expresses to begin with.

⁶ For a detailed description of the meaning and syntax of MCs, see Giannakidou and Yoon (2011) and references therein.

With this background in mind, the SCP predicts a containment effect with *para* MCs, because they are consistently subjective (Giannakidou & Yoon, 2011), but not with *apoti* comparatives, because they are objective.

The example in (24) shows that indeed in the context of a *para* MC a QP is unable to take wide scope. As before, the context in (24) is only compatible with wide scope of the QP; this is because on the intended reading the comparative predicate distributes over the individuals picked out by “fewer than half of the students”. In other words, the context is compatible with a statement along the lines: fewer than half the students are such that the degree to which the speaker likes the sentence “he/she is a little child” as a description of him/her is greater than the degree to which she likes “he/she is an adult”. The same context is incompatible with a reading according to which the degree to which the speaker likes the sentence “fewer than half the students are little children” is greater than the degree to which she likes “fewer than half the students are adults”.⁷ We follow Giannakidou and Yoon (2011) in assuming that *para* MCs involve the subjective category *more_{ML}*, corresponding to *perissotero* in (24), and assume that this is the item blocking the QP from taking wide scope, in accordance with the SCP.

- (24) *Context: I had 30 students in my final year syntax class. To obtain their degree, they had to pass my exam. For 10 of these students, I considered it better if they failed as they were just not ready for the big world. After the exam had taken place, I quickly marked all the scripts and discovered that everybody had passed. In order to explain to my TAs why this is not necessarily good news for the 10 who are not prepared for post-university life, I say...*

#Ligoteroi apo tous misous fitites einai perissotero mikra paidia
Fewer than the half students are more_{ML} little children
 para/(*apoti) enilikes.
than adults
 ‘Fewer than half the students are little children rather than adults.’

The example in (25) demonstrates that an *apoti* comparative does not prevent the QP from taking wide scope. Again, given the explicit reference to a subset of students in the context, which is the same as in (24), the statement in (25) can only take a QP wide scope interpretation, along the lines: fewer than half the students are such that the degree to which they are unemployed citizens is greater than the degree to which they are students. The same context is incompatible with an interpretation of the sort: the degree to which fewer than half the students are unemployed citizens is greater than the degree to which they are students.

- (25) *Context: Same as in (24).*

Ligoteroi apo tous misous fitites twra einai perissotero anergoi
Fewer than the half students now are more unemployed
 polites apoti fitites.
citizens than students
 ‘Fewer than half the students are now unemployed citizens rather than students.’

⁷ The interpretations provided for the comparatives in (24) and (25) are based on Giannakidou & Yoon’s (2011) semantics of the relevant constructions.

2.3 Outside Negation Polar Questions

This section demonstrates that containment effects also arise when combining QPs and so-called ‘outside’ negation (ON) in the context of polar questions (PQs). Semantically, ONPQs express that the speaker seeks confirmation for a proposition held to be true on the basis of *private* concerns. In line with this, they are incompatible with explicit evidence supporting the truth of the relevant proposition (Büring & Gunlogson, 2000; see also Sudo, 2013). Another important property of ONPQs, which grammatically distinguishes them from PQs with regular truth-conditional negation (‘inside’ negation or IN), is that they license PPIs like *too*, but not NPIs like *either* (Ladd, 1981). (26a) shows the incompatibility of ONPQs with explicit evidence supporting the positive answer:

- (26) *Context: For a psychological experiment, we are looking for some left-handed subjects. We have asked some of our friends, but only Mary was left-handed so far. To my surprise, John is using a pencil with his left hand.*
- a. #Isn’t John left-handed(, too)?
- b. Isn’t John right-handed(, too)? (Sudo, 2013)

ONPQs further differ from INPQs in that the latter do not necessarily express any bias (even though they may do so, towards the negative answer; see Sudo (2013)), nor require fronting of negation along with the (auxiliary) verb.

- (27) a. Is John not right-handed, either? (meaning: Is it the case that John is not right handed, either?)
- b. Isn’t John right-handed, either? (meaning: Is it the case that John is not right handed, either?)

Another property of ONPQs, already noted by Ladd (1981), is that they are answered in exactly the same way as positive PQs (see Holmberg (2015) for extensive discussion of strategies for answering INPQs):

- (28) a. Does Mary like spinach?
Yes (she does).
No (she doesn’t).
- b. Doesn’t Mary like spinach(, too)?
Yes (she does).
No (she doesn’t).

In line with these observations, Ladd (1981) suggests that, in a very concrete sense, ON is not contained in the question (hence the name ‘outside’). Holmberg (2015) makes this intuition more precise by assuming that ON must out-scope the question operator; on this view, the properties that distinguish ON from IN, including its non-truth-conditional meaning contribution and inability to license NPIs, are derived from the fact that negation is interpreted extremely high in the clause, negating, in this way, quite literally, the question operator. This results in the question not being interpreted as a question, but instead as some sort of assertion regarding the proposition *p* the question operator scopes over, hence the bias towards the truth of *p*.

We agree with Holmberg that the simplest analysis would assume that ON is just another instance of a property-reversing operator that is nevertheless interpreted

idiosyncratically due to it scoping very high.⁸ This would explain the grammatical behavior of ONPQs, including the data we discuss shortly. However, in our opinion the assumption that ON negates the question operator faces an important issue; although ONPQs are understood to express a bias towards the truth of *p*, they are still questions, again in a literal sense. One way of overcoming this objection would be to assume that ON does not really negate the interrogative force of the clause, but instead, the *content* of the Q-operator (see Giannakidou and Yoon (2011) for how subjective evaluations may target either the *form* or the *content* of a sentence). In particular, one could assume that ON targets the likelihood of the alternatives made available by the Q-operator. Assuming that the Q-operator expresses that *p* and $\neg p$ are equally likely in root PQs, *not* has the effect that the alternatives are not equally likely. Note that this kind of modification is not truth-conditional, but still influences the way the question is to be interpreted because one of the alternatives is taken to be more likely to be true than the other. If the alternative that is taken to be more likely to be true is *p*, then we get a good approximation of the positive bias expressed by ONPQs.⁹

If ON is indeed a subjective category, its wide scope over very high operators like the Q-operator is forced by the SCP (on the null assumption that Q is a non-subjective operator). This naturally leads to the conclusion that the inability of other subjective categories to occur in questions (see section 0) need not be due to the fact that they occur in the scope of the Q-operator, as suggested by Ernst (2009), but may instead result from semantic incoherence associated with scope extension past the Q-operator.

We now demonstrate that ON exhibits similar scopal behavior to other subjective categories. The context in (29) forces a wide scope reading of the QP and, as shown by the deviance of (29a) in this context, the QP cannot be interpreted outside the scope of ON. Note that the example becomes fully interpretable if ON is moved to a following tag question, as in (29b).

- (29) *Context: I had 30 students in my final year syntax class. They all passed the coursework, but to obtain their degree, they had to pass my exam. I was so impressed with 10 of my students that I was certain they would pass the exam no matter what. But on past experience, I was confident that in fact the vast majority of them would. Now my TAs have marked all the scripts and I ask ...*
- a. #Haven't fewer than half of the students managed to pass the exam, too?
 - b. Fewer than half of the students managed to pass the exam, haven't they?

Since IN is not subjective, we expect a QP to be able to out-scope it. This prediction is borne out, as shown in (30), which is felicitous in a context forcing wide scope of the QP.

⁸ It should be noted that there are other analyses on the market that assume that ONPQs involve an epistemic operator; either negation itself is such an operator (e.g. Van de Koot et al., 2015) or some other covert element (Romero & Han, 2004). One major problem we see with regard to the former type of approaches is that cross-linguistically, ON looks like regular negation, morphologically, and, to our knowledge, never like an epistemic modal. With regard to Romero and Han (2004), they assume that ONPQs involve a covert epistemic modal that is responsible for the epistemic bias these constructions express. However, as pointed out to us by Floris Roelofsen (p.c.), one major issue for this account is that it is not expected to be possible for ONPQs to be answered in the same way as neutral PQs, contrary to fact (see (28)). Since the focus of the present paper is different, we do not wish to dwell on the pros and cons of each account separately. If it turns out that ONPQs involve an epistemic operator, then the observations in this section should be taken as evidence that this epistemic operator is a subjective one, behaving in accordance with the SCP.

⁹ A full analysis of ONPQs must have a way of linking this qualification of interrogative force to the perspective of the attitude holder for ON, but space does not permit us to dwell on this issue here.

(30) *Context: I had 30 students in my final year syntax class. 10 of them failed the coursework but could still pass the module if they passed my exam. I was so disappointed with the performance of those 10 students at the coursework that I suspected they would not pass the exam either. Now my TAs have marked all the scripts and I ask ...*

Haven't fewer than half of the students passed the exam, either?

(30) is important in another respect, as it demonstrates that movement of negation to C does not necessarily result in wide scope over the QP.

That containment effects are in any case not dependent on movement of negation is easily demonstrated in Dutch, where negation is located below a subject QP irrespective of whether it is a manifestation of ON or IN (compare (31) and (32)). Crucially, Dutch ON gives rise to containment effects even so, as shown in (31a).

(31) *Context: Same as in (29).*

a. *Zijn minder dan de helft van de studenten niet ook voor het
are fewer than the half of the students not also for the
 examen geslaagd? Of niet?
exam passed or not

'Haven't fewer than half of the students managed to pass the exam, too?'

b. Minder dan de helft van de studenten zijn ook voor het
fewer than the half of the students are also for the
 examen geslaagd. Of niet?
exam passed or not

'Fewer than half of the students have managed to pass the exam, too. Haven't they?'

(32) *Context: Same as in (30).*

Zijn minder dan de helft van de studenten ook niet voor het
are fewer than the half of the students also not for the
 examen geslaagd?
exam passed

'Haven't fewer than half of the students passed the exam, either?'

2.4 Subjective categories and tense

In the preceding sections we have reviewed evidence that subjective categories consistently outscope QPs in the proposition with which they are associated. However, unlike von Stechow and Iatridou's (2003) ECP, the SCP is stated in more general terms; putting aside the precise property giving rise to widest scope (i.e. subjectivity vs epistemicity), the SCP states that a subjective category must out-scope *any* operator (as opposed to just QPs) in the relevant proposition. In the last section, we saw that the assumption that ON scopes over the question operator may account for the fact that a PPI, such as *too*, may be licensed in the relevant proposition, but not an NPI, such as *either*, thereby providing some support for the generality of the SCP. In this section we build on observations of Papafragou (2006) that subjective epistemic modals, but not objective ones, consistently out-scope tense and show that this pattern extends to subjective categories more generally.

Papafragou (2006, p. 1695) suggests that "the main difference between subjective and objective epistemic modality is that the former, but not the latter, is indexical, in the sense that the possible worlds in the conversational background are restricted to what the current

speaker knows as of the time of utterance. By contrast, in the case of objective epistemic modality, possible worlds in the conversational background include what is generally known to some community, or, in other words, what the publicly available evidence is". Subjective epistemic modality is thus tied to the here-and-now of the conversation, whereas objective epistemic modality may be used for discussing past and future possibilities. Put differently, subjective epistemic modality consistently out-scopes the tense operator (34), whereas objective epistemic modality does not have to (33).

- (33) Yesterday it was possible that the stock market would go up today.
(Papafragou, 2006, p. 1695)
- (34) John was perhaps arrested by the police yesterday. (#But, at the moment, I don't think he was.)

Evaluative adverbs (36), ON (38) and MC comparatives (40) pattern with the strongly subjective epistemic modal *perhaps*, whereas evidential adverbs (35), IN (37) and RC comparatives (39) pattern with *possibly*, as shown below.

- (35) Until yesterday he was apparently arrested by the police (✓but now it's no longer that apparent.)
- (36) Yesterday, he was unfortunately arrested by the police (#but, I no longer think it's that unfortunate).
- (37) Was he not arrested by the police, yesterday? (means: Is it the case that he was not arrested yesterday?)
- (38) (#Even though I don't believe at the moment that he was,) wasn't he arrested yesterday?
- (39) Persi ta provlimata sou itan perissotero oikonomika apoti
Last year the problems you were more financial than
nomika, alla twra pistevw to antitheto.
legal but now believe the opposite
'Last year, your problems were more financial than legal, but now I believe the opposite.'
- (40) Persi ta provlimata sou itan perissotero oikonomika para
Last year the problems you were more financial than
nomika, (#alla twra pistevw to antitheto.)
legal but now believe the opposite
'Last year, your problems were rather financial than legal, but now I believe the opposite.'

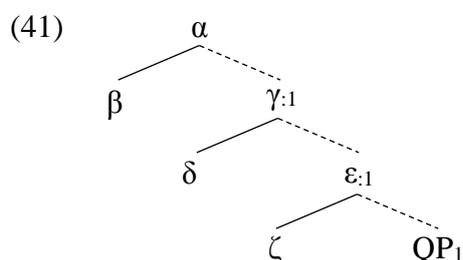
We conclude that the generality of the SCP is warranted by the fact that subjective categories consistently out-scope not only QPs, but also very high propositional operators, such as tense. Contrary to the ECP, which is peculiarly specific to epistemic modals out-scoping QPs, the SCP restricts the scope of any subjective category with respect to any non-subjective operator.

3 Accounting for Subjective Containment

3.1 The Condition on Scope Shift

Constantinou and Van de Koot (2015) provide an account of the containment and scope freezing effects induced by subjective epistemics based on the theory of scope shift developed in Neeleman and Van de Koot (2012). We first provide a brief outline of the latter work and then propose a variant of it that captures the containment effects with subjective categories observed in previous sections.

Williams (1994) assumes that an argument QP carries a scope index that may percolate to a dominating node to mark the QP's extended scope. Thus, the scope of QP_1 in (41) below corresponds to the largest category that carries its scope index (γ), minus the QP itself. Note that a scope index inherited by a node α is placed after a colon. This distinguishes it from an index introduced by α , which appears in front of any colon; we omit the colon when there is no inherited index.



In some languages, such as German and Japanese, a QP argument may also mark its scope through overt A'-movement, in which case scope is marked in the landing site. Scope-marking through movement is discussed in more detail in section 0, where it becomes directly relevant. Finally, an argument QP may also fail to percolate a scope index altogether, in which case it takes surface scope.

Neeleman and Van de Koot combine the index-based scope marking mechanism with the following minimality condition on inheritance of quantificational indices:

- (42) *Condition on Scope Shift (CSS)*
No node may inherit two scope indices.

It is easy to see that this constraint is incompatible with the widely held view, originating in the work of Chomsky (1976) and May (1977), that there is a syntactic level of Logical Form (LF) that provides a transparent and complete representation of scope relations. A translation of this view into the index-based representations of Williams (1994) yields the representation in (43a) for a structure containing two QPs interpreted as taking surface scope and that in (43b) for the same structure with inverse scope. As can be easily verified, both (43a) and (43b) violate the CSS.

- (43) a. $*[_{:1} \dots [_{:1,2} \dots QP_1 [_{:2} \dots QP_2 \dots]]]$ $QP_1 > QP_2$
b. $*[_{:2} \dots [_{:1,2} \dots QP_1 [_{:2} \dots QP_2 \dots]]]$ $QP_2 > QP_1$

There is an alternative view of scope according to which LF only represents deviations from surface scope (see Reinhart (1983, 1995, 2006); see also Lakoff (1972), Huang (1982) and Hoji (1985)). On this view, scope extension is limited to structures in which it generates an

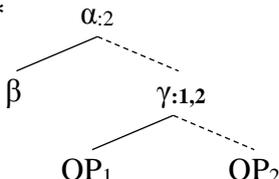
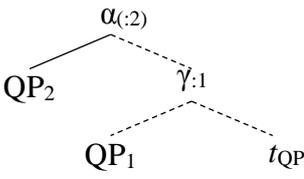
interpretation that is otherwise unavailable.¹⁰ Reinhart treats scope extension as QR; a translation of her proposal in terms of indices expresses the readings in (43) with the slightly simpler structures in (44), neither of which violates the CSS, as required.

- (44) a. [... [... QP₁ [... QP₂ ...]]] QP₁ > QP₂
 b. [:₂ ... [:₂ ... QP₁ [:₂ ... QP₂ ...]]] QP₂ > QP₁

We summarize the main tenets of the proposal in (45). The Economy principle in (45b) is intended to block scope extension where it does not give rise to inverse scope, while (45c) is a mapping principle that, in the absence of scope extension, regulates the association of syntactic structures at LF with their semantic representations.

- (45) a. *Scope Extension*
 If a QP percolates its index to a dominating node α , then its scope coincides with α minus the QP itself.
 b. *Economy*
 Scope extension must give rise to an otherwise unavailable interpretation.
 c. *Default Scope Rule*
 If a QP does not percolate its index, it takes scope over its c-command domain.

An important fact about the CSS is that it creates an asymmetry between covert scope shift, analyzed here as index percolation, and scope taking by overt movement. Covert scope shift involves index percolation. It is therefore incompatible with any other scope extensions in its percolation path, as these give rise to CSS violations. This is shown in (46a), where covert scope extension of QP₂ freezes the scope of QP₁. By contrast, overt movement of a QP cannot trigger CSS violations in the movement path, whether the moved QP marks scope in its landing site (because it has undergone overt QR; see section 0 for details) or not (because it has undergone A-movement, which does not mark scope). This is shown in (46b), where movement of QP₂ does not prevent covert scope extension by QP₁.

- (46) a. *  b. 

Thus, overt scope marking is freer than covert scope shift, a prediction that is corroborated by a range of scope interactions (see Neeleman and Van de Koot (2012) for extensive discussion). As will be clear, this predicted difference is hard to reconcile with theories that treat overt and covert scope shift as mediated by the same operation, for example movement.¹¹ With this background, we now return to subjective categories.

¹⁰ The view that scope extension is subject to Economy and relativized to an interpretation has been argued for in Fox (1999; 2000). However, unlike Reinhart, Fox assumes that every QP must move.

¹¹ The key point is that our scope theory predicts the existence of an asymmetry between overt and covert scope extension. To be sure, there may well be alternative theories of scope (extension) that make the same prediction and it is a reasonable expectation that the account of the containment and scope freezing effects to be presented in what follows would be easily adaptable to such alternatives.

3.2 Interactions between the CSS and the SCP

Subjective categories are quantificational and their distribution is, therefore, partly regulated by the system of scope outlined in the previous section. In particular, whenever a subjective category finds itself in the scope of another operator in the clause with which the subjective attitude is connected, it must extend its scope through index percolation, or a violation of the SCP will result.

In sentences with a single QP, this gives rise to the containment effects already discussed,¹² as summarized in (47) (the structure in (47c) in which the subjective category fails to extend its scope violates the SCP).

- (47) a. [... SUBJ₁ [... QP₂ ...]] (SUBJ₁ > QP₂)
 b. [:₁ ... QP₂ [:₁ ... SUBJ₁ ...]] (SUBJ₁ > QP₂)
 c. *[: ... QP₂ [... SUBJ₁ ...]] (QP₂ > SUBJ₁)

These facts follow in any proposal that adopts the SCP or a similar constraint and do not specifically rely on any particular theory of scope.

However, as we will now show, the combination of the CSS-based scope theory and the SCP predicts scope-freezing effects in sentences with multiple QPs that are unexpected on a QR-based approach. In all six structures below, the subjective category satisfies the SCP. In (48a,b), SUBJ outscopes the two QPs on the surface, and QP₃ may or may not extend its scope across QP₂. By contrast, in (48c-f), QP₂ outscopes SUBJ on the surface. Satisfaction of the SCP therefore requires SUBJ to percolate its scope index past this QP. This, in turn, has the consequence that QP₃ cannot outscope QP₂, because the relevant configurations constitute CSS violations.

- (48) a. [... SUBJ₁ [... QP₂ [... QP₃ ...]]] (SUBJ₁ > QP₂ > QP₃)
 b. [... SUBJ₁ [:₃ ... QP₂ [:₃ ... QP₃ ...]]] (SUBJ₁ > QP₃ > QP₂)
 c. [:₁ ... QP₂ [:₁ ... SUBJ₁ [... QP₃ ...]]] (SUBJ₁ > QP₂ > QP₃)
 d. *[:_{1,3} ... QP₂ [:_{1,3} ... SUBJ₁ [:₃ ... QP₃ ...]]] (SUBJ₁ > QP₃ > QP₂)
 e. [:₁ ... QP₂ [:₁ ... QP₃ [:₁ ... SUBJ₁ ...]]] (SUBJ₁ > QP₂ > QP₃)
 f. *[:_{1,3} ... QP₂ [:_{1,3} ... QP₃ [:₁ ... SUBJ₁ ...]]] (SUBJ₁ > QP₃ > QP₂)

A QR-based theory that assumes the SCP does not make these predictions. In particular, it is unclear why QR of SUBJ to some very high position should ever interfere with QR of QP₃ across QP₂ in the movement path of SUBJ.

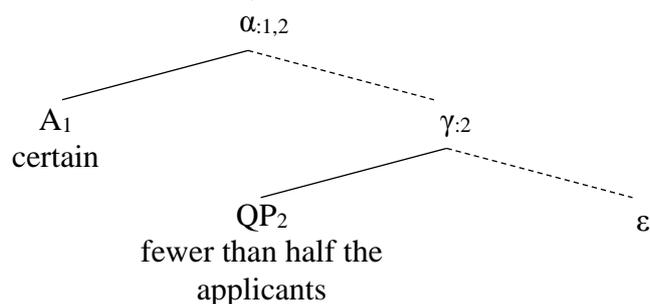
In section 0, we show in more detail how the CSS-based approach accounts for the containment effects discussed in the previous sections. Then, in section 0, we evaluate the predictions of the CSS-based theory regarding scope-freezing effects and show that they are corroborated by the data.

3.3 Subjective Containment as a CSS effect

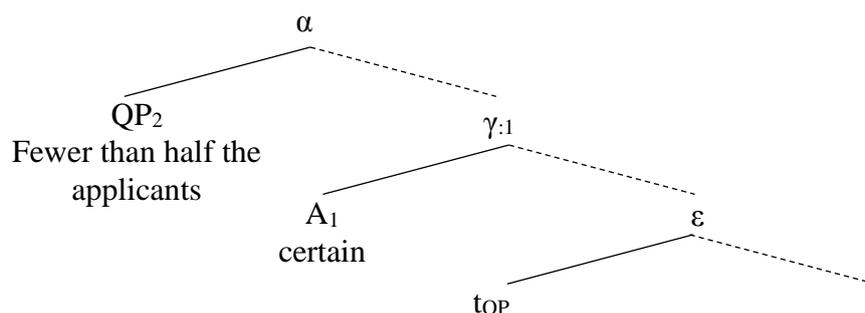
Consider first the examples in (2) and (4), which involve a subjective epistemic auxiliary verb and adverb, respectively. As shown in (49), these instantiate cases in which a subjective category merges in a position from which it has to percolate its index to the top of its own clause (presumably at least up to the TP level) in order to satisfy (11).

¹² To simplify the discussion, we disregard the conclusions of section 0, where we discuss scope interactions between subjective categories and tense. This does not affect the overall analysis.

(50) a. * = (5)

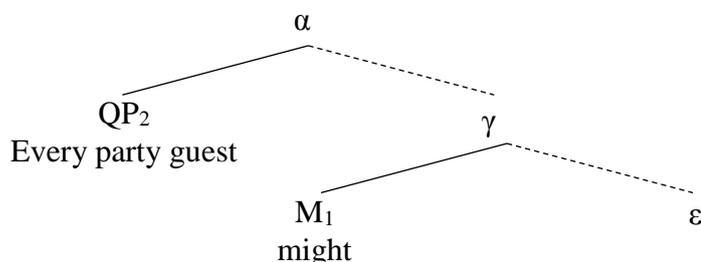


b. = (10)



Since the SCP only applies to subjective categories, an objective interpretation of an epistemic modal does not result in percolation of the scope index of that modal. Instead, the default scope rule in (45c) applies. Thus, in (51), the LF of (7b), the structurally higher QP is able to out-scope *might*.

(51) = (7b)



The facts relating to SpOAs and MCs receive a parallel treatment. As regards SpOAs, evaluative adverbs, being subjective, adhere to the SCP and thus prevent a QP from percolating an index. Evidentials, on the other hand, being objective, do not adhere to the SCP, hence the c-commanding QP can out-scope them (compare (52a-b)). (53a-b) show how the contrast between metalinguistic and regular clausal comparatives (24) and (25) is captured. As discussed in 0, we follow Giannakidou and Yoon (2011) in assuming that *para* MCs involve the subjective category more_{ML} , i.e. *perissotero* in (53a), the item preventing the QP from extending its own scope index. Similarly to evidentials, regular comparatives are

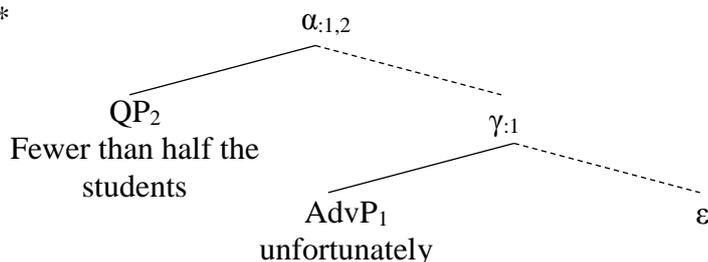
modal auxiliaries in examples like (2) and (3). The state of affairs described in (a) and (b) also obtains for the raising case in (10) as the object QP of the complement clause A-moves past the epistemic modal *certain*.

(ii) [Fewer than half the applicants]₁ are certain t_1 to have been hired t_1 , ...

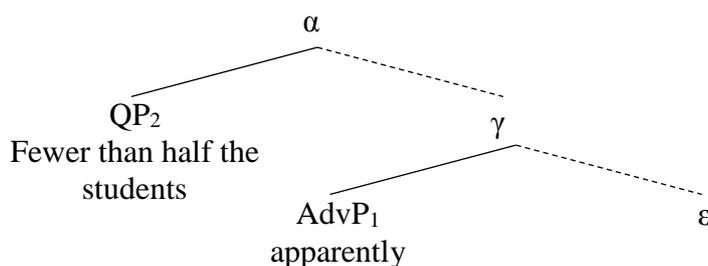
Reconstruction of the QP for scope explains the *certain* > *fewer than half* reading, however, (i) leads to the expectation that this should be the only available reading, contrary to fact.

objective; *perissotero* ‘more’ is thus not forced to mark its scope via index percolation, allowing in this way the higher QP to out-scope it in (53b).

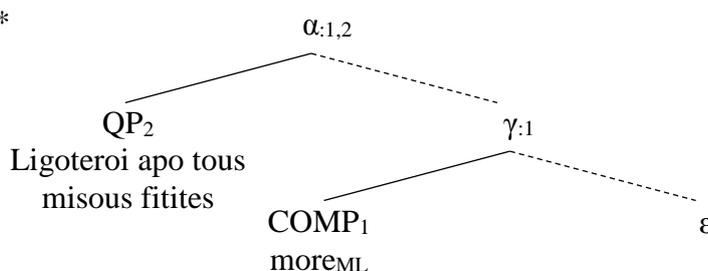
(52) a. * = (18)



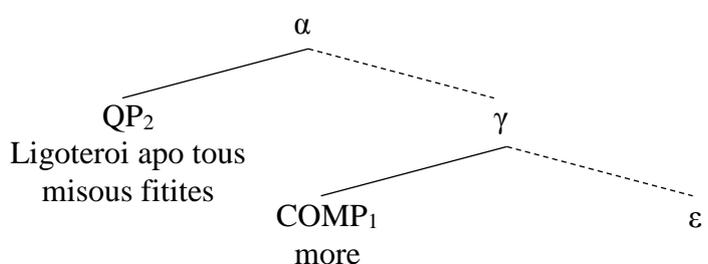
b. = (19)



(53) a. * = (24)



b. = (25)



Consider finally the contrast between ONPQs and INPQs in terms of containment effects (see (29a) and (30)). The contrast is accounted for in the same way if, in accordance with the SCP, ON must out-scope other operators by percolating an index. It then follows that the QP in (29a) cannot also extend an index to outscope negation without a CSS violation. IN behaves in the same way as objective epistemics, evidentials and regular comparatives in that index percolation is not necessary. The QP can thus take wide scope either by percolating its own index past the landing position of negation or by the default scope rule (we assume that negation takes/extends scope from its base position, as the movement of negation is not motivated by scope-related considerations).^{14,15}

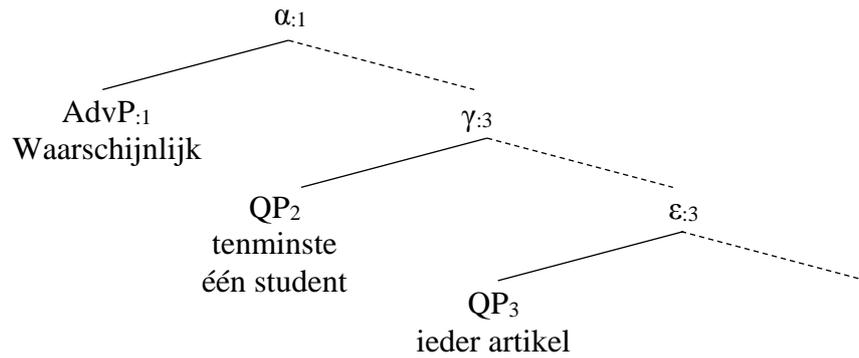
¹⁴ The structure in (54) assumes that the QP can extend its scope in the left periphery, that is, past the TP level. This is in accordance with evidence suggesting that QPs can indeed extend their scope quite freely, even beyond this level, e.g. specCP. For instance, universal QPs may take wide scope over a moved wh-expression (Krifka, 2001), as evidenced by the pair-list reading of (i).

- b. Tenminste één student heeft helaas ieder artikel zonder
at-least one student has unfortunately every article without
 morren gelezen.
complaining read
 $(\exists > \forall; * \forall > \exists)$
- c. Tenminste één student heeft ieder artikel helaas zonder
at-least one student has every article unfortunately without
 morren gelezen.
complaining read
 $(\exists > \forall; * \forall > \exists)$
 ‘At least one student has unfortunately read every article without complaining.’
- (57) a. Had niet tenminste één student ieder artikel zonder morren
had not at-least one student every article without complaining
 gelezen?
read
 ONPQ: $\exists > \forall; \forall > \exists$
- b. Had tenminste één student niet ieder artikel zonder morren
had at-least one student not every article without complaining
 gelezen?
read
 ONPQ: $\exists > \forall; * \forall > \exists$
- c. Had tenminste één student ieder artikel niet zonder morren
had at-least one student every article not without complaining
 gelezen?
read
 ONPQ: $\exists > \forall; * \forall > \exists$
 ‘Hasn’t at least one student read every article without complaining?’

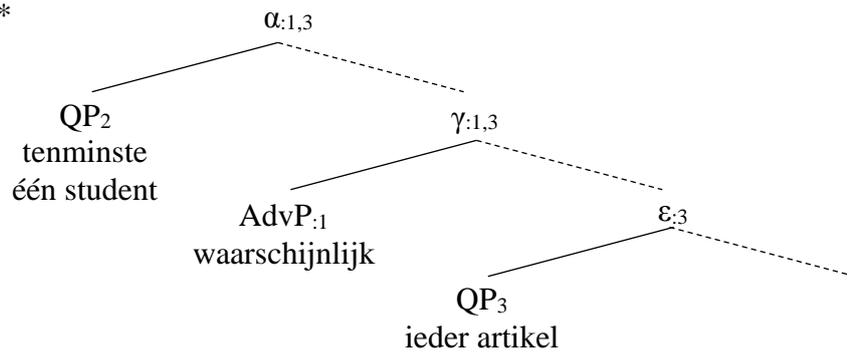
For a QR-based theory of scope, the effects in (55) – (57) are entirely unexpected; if the subjective category achieves widest scope over other operators by covert movement, there is simply no way of understanding why there are QP scope freezing effects in the (b) and (c) sentences, but not in the (a) ones.

As already explained in section 0, the pattern in (55) – (57) follows directly from the combined effect of the SCP and the CSS. In (55b), (56b) and (57b), the subjective category sits between the two QPs and must therefore percolate its scope index to outscope the highest one. Scope extension of the universal (to outscope the indefinite) is blocked, since its index would travel the same path as that of the subjective category. The scope freezing effect in the (c) examples, where the subjective category is c-commanded by both QPs, follows from our proposal in much the same way: scope extension by the subjective category freezes the scope of QPs in its percolation path. Finally, the well-formedness of (55a), (56a) and (57a) on the inverse scope reading falls out from the fact that the percolation path of the universal may terminate in a node that is below the percolation path of the subjective category. Hence, there is no overlap of the two scope extension paths. Below we only provide partial structures for the inverse scope reading of the data in (55), but the same explanation applies to the data in (56) and (57).

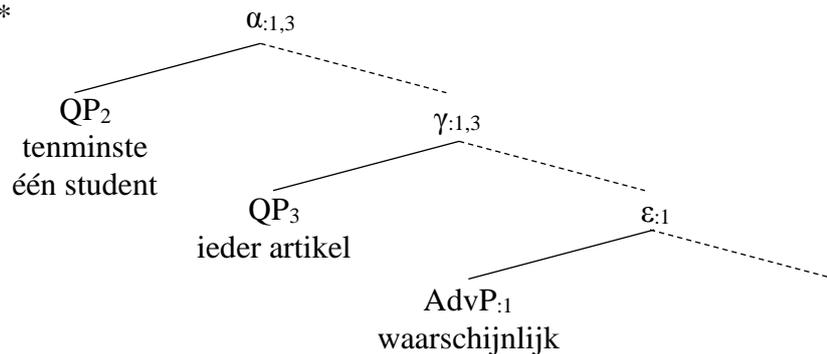
(58) a. = (55a)



b. * (=55b)



c. * (=55c)



Note that the scope freezing effects just reviewed are also found when outside negation or a subjective SpOA intervenes between the direct object and a to-dative, as shown in (59) and (60).

- (59) a. Heeft Jan niet tenminste één boek aan ieder meisje zonder
has John not at-least one book to every girl without
 morren gegeven?
complaining given
 ($\exists > \forall; \forall > \exists$)
- b. Heeft Jan tenminste één boek niet aan ieder meisje zonder
has John at-least one book not to every girl without
 morren gegeven?
complaining given
 ($\exists > \forall; * \forall > \exists$)

- c. Heeft Jan tenminste één boek aan ieder meisje niet zonder
has John at-least one book to every girl not without
 morren gegeven?
complaining given
 $(\exists > \forall; * \forall > \exists)$
 ‘Has John not given at least one book to every girl without complaining?’
- (60) a. Jan heeft waarschijnlijk/helaas tenminste één boek aan ieder
John has probably/unfortunately at-least one book to every
 meisje zonder morren gegeven.
girl without complaining given
 $(\exists > \forall; \forall > \exists)$
- b. Jan heeft tenminste één boek waarschijnlijk/helaas aan ieder
John has at-least one book probably/unfortunately to every
 meisje zonder morren gegeven.
girl without complaining given
 $(\exists > \forall; * \forall > \exists)$
- c. Jan heeft tenminste één boek aan ieder meisje
John has at-least one book to every girl
 waarschijnlijk/helaas zonder morren gegeven.
probably/unfortunately without complaining given
 $(\exists > \forall; * \forall > \exists)$
 ‘John has probably/unfortunately given at least one book to every girl without complaining.’

The problems for any QR-based approach to these data are compounded by the observation that (for the same group of speakers) an overt A'-movement operation across two QPs does not result in freezing effects. This is demonstrated by the data in (61). In (61b), movement of a contrastively focused direct object (shown with SMALL CAPS) across the subject does not prevent the indirect object from taking scope over the former.

- (61) a. Tenminste één student heeft iedere professor DAT verhaal verteld.
at-least one student has every professor that story told
 $(\exists > \forall; \forall > \exists)$
 ‘At least one student has told every professor THAT story.’
- b. DAT verhaal heeft tenminste één student iedere professor *t* verteld.
that story has at-least one student every professor told
 $(\exists > \forall; \forall > \exists)$
 THAT story at least one student has told every professor.

The CSS-based account successfully captures these facts, because overt A'-movement does not give rise to index percolation in the movement path (i.e. the path from the trace to the landing site). This allows the universal to extend its scope index past the indefinite without triggering a CSS violation.

4 Information Structure and the SCP

4.1 Interactions between IS categories and subjective categories

Neeleman and Van de Koot (2012) motivate their CSS-based theory of scope by showing that it correctly predicts scope restrictions in structures with three quantifiers, ordering restrictions on contrastive topics and foci, and scope interactions between quantifiers and contrastive categories. Of particular interest for the present proposal is their account of topic-focus order restrictions, because it relies on an information-structural condition, Topic Externality, that forces a topic to have wider scope than a focus at LF (see Krifka (2007) and Tomioka (2009) for discussion):

- (62) **Topic Externality**
 a. topic [...focus...]
 b. *focus [...topic...]

We then minimally expect the following predictions to be borne out: (i) any structural environment in which a topic is prevented from satisfying Topic Externality should also be an environment in which a subjective category is prevented from satisfying the SCP; (ii) in a structure in which a topic is forced to percolate a scope index to satisfy Topic Externality, the presence of a subjective category in the percolation path of the topic should give rise to a CSS violation.

We will review these predictions in turn, but before doing so we should clarify the notions of topic and focus we assume. The term focus is reserved for constituents that receive the main stress of the sentence (e.g. Selkirk (1984, 1996)). These constituents usually express new information. For instance, the constituent answering a wh-expression is taken to be in focus. With respect to topics, we follow Reinhart (1981) and subsequent work, in characterizing them in terms of ‘aboutness’. We distinguish between ‘discourse topics’ and ‘sentence topics’. Discourse topics refer to entities that a unit of discourse is about, whereas sentence topics are syntactic constituents that introduce a new discourse topic or narrow down the current discourse topic. On this view, constituents that are merely discourse-anaphoric do not qualify as sentence-topics (not even if they are anaphoric to the current discourse topic). For what follows it is essential that the notion ‘topic’, including its use in (62), is understood to mean sentence topic.

Let us now consider structures in which a topic is prevented from satisfying Topic externality. The basic pattern observed in Neeleman and Van de Koot (2008, 2012) can be summarized as follows: topics and foci may appear in either order, with the topic c-commanding the focus or vice versa, but as soon as one of these categories moves, a surface structure must be generated in which the topic c-commands the focus. In short, a topic may not be contained in the c-command domain of a moved focus. This pattern is illustrated below with data from Neeleman and Van de Koot (2008). (In all examples, topics are doubly underlined, foci appear in SMALL CAPS and contrastive categories are *italicised*.)¹⁶

- (63) A: Hoe zit het met *FRED*? Wat heeft *HIJ* gegeten?
 B: Nou, dat weet ik niet, maar ik geloof...
 A: ‘What about Fred? What did he eat?’
 B: ‘Well, I don’t know, but I believe...’

¹⁶ We deliberately use a negative quantifier for the contrastive focus. This makes it very difficult to perform a topic-focus swap, since negative quantifiers are very poor topics (except if the topic of discussion is a quantity).

- a. dat Wim *NERGENS* van meer gegeten heeft dan vorig jaar.
that Bill nothing from more eaten has than last year
- b. #dat [*NERGENS* van]₁ Wim *t₁* meer gegeten heeft dan vorig jaar.
that nothing of Bill more eaten has than last year
 ‘...that Bill has not eaten more of anything than last year.’

(64) A: Hoe zit het met de *SOEP*? Wie heeft *DIE* gegeten?

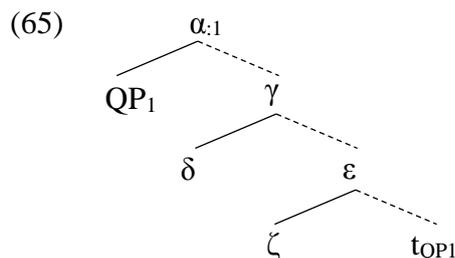
B: Nou, dat weet ik niet, maar ik geloof...

A: ‘What about the soup? Who ate that?’

B: ‘Well, I don’t know, but I believe...’

- a. dat *NOBODY* van de bonen meer gegeten heeft dan vorig jaar
that niemand from the beans more eaten has than last year
- b. dat [*van de bonen*]₁ *NOBODY* *t₁* meer gegeten heeft dan vorig jaar.
that from the beans NOBODY more eaten has than last year
 ‘...that nobody has eaten more from the beans than last year.’

Neeleman and Van de Koot (2012) account for these facts using the CSS. They propose that A²-movement of a contrastively marked element determines its scope, much like A²-movement of other quantificational elements.¹⁷ This scope marking operation is formalized as the mandatory percolation of the scope index from the moved category to the node that immediately dominates its landing site (see also Williams’s (1994) adjunct scope rule). Thus, in a structure like (65), the scope of QP is α minus the QP itself. This equals to γ .



Contrastive elements that remain in-situ are assigned scope via the default scope rule in (45c), which results in surface scope, or they may extend their scope through index percolation (the scope extension rule in (41)).

Now consider a structure containing a contrastive topic and a contrastive focus. If both are in situ, as in (66), the topic can freely percolate an index past the c-commanding focus, since the focus can be assigned default scope. If the topic has moved across the focus, as in (67a), it marks scope in its landing site and satisfies Topic externality. However, if the focus moves across the topic, as in (67b), it marks scope in its landing site. The CSS therefore prevents the topic from satisfying Topic Externality by percolating an index past the focus.

- (66) a. [... topic ... *FOCUS* ...]
 b. [:₂ ... *FOCUS* ... [:₂ ... topic₂ ...]]
- (67) a. [:₂ topic₂ ... [... *FOCUS* ... t_{topic} ...]]
 b. *[:_{1,2} *FOCUS* ₁ ... [:₂ ... topic₂ ... t_{focus} ...]]

¹⁷ In particular, it is assumed that contrastive elements involve a negative operator in their semantics, and it is the presence of this operator that makes contrast quantificational (see also Kiss (1998)). Thus, the movement of the contrastive constituent marks what material is included in the scope of the negative operator.

This brings us to prediction (i): if a moved focus prevents a topic in its c-command domain from satisfying Topic Externality, then it should also prevent a subjective category in its c-command domain from satisfying the SCP. Similarly, since an in situ focus does not prevent a topic in its c-command domain from percolating its index, it should also not prevent a subjective category from doing so. The relevant structures are shown schematically in (68).

- (68) a. [₂ ... *FOCUS* ... [₂ ... unfortunately/unbelievably₂ ...]]
 b. *[:_{1,2}... *FOCUS*₁ ... [₂ ... unfortunately/unbelievably₂ ... *t_{focus}* ...]]

The data in (69) and (70) show that this prediction is borne out. A strong evaluative adverb can occur below an in situ contrastive focus in English and Dutch, as in (69), but an occurrence of the same adverb below a *moved* contrastive focus gives a degraded result, as (70) demonstrates.

- (69) A: John has failed the exam.
 B: No, *BILL* has unfortunately failed the exam.
 B: Nee, *WILLEM* is helaas gezakt voor het examen.
no Bill is unfortunately failed for the exam
- (70) A: John has failed one exam.
 B: #No, *THREE* exams John has unfortunately failed.
 B: #Nee, voor *DRIE* examens is Jan helaas geslaagd.
no for three exams is John unfortunately failed

Let us now consider prediction (ii): in a structure in which a topic is forced to percolate a scope index to satisfy Topic Externality, the presence of a subjective category in the percolation path of the topic should give rise to a CSS violation. The relevant structure is a variant of (66b) above, in which a subjective category occurs between the focus and the topic, as in (71b). By contrast, a structure like (71a), in which a subjective category intervenes between a topic and a focus in a configuration that satisfies Topic Externality, is predicted to be fully acceptable.

- (71) a. [₁ *topic*... [SUBJ₁ ... [... *FOCUS* ...]]]
 b. *[:_{1,2} *FOCUS* ... [:_{1,2} SUBJ₁ ... [:₂ ... *topic*₂ ...]]]

Example (72) instantiates the structure in (71a) and is indeed completely unobjectionable. As predicted, example (73), which instantiates (71b), is indeed quite degraded.

- (72) A: Hoe zit het met *FRED*? Wat heeft *HIJ* gegeten?
 B: Nou, dat weet ik niet, maar ik geloof...
 A: ‘What about Fred? What did he eat?’
 B: ‘Well, I don’t know, but I believe...’
 dat *Wim* helaas *NERGENS* van meer gegeten heeft dan vorig jaar.
that Bill unfortunately nothing of more eaten has than last year

- (73) A: Hoe zit het met de *SOEP*? Wie heeft *DIE* gegeten?
 B: Nou, dat weet ik niet, maar ik geloof...
 A: ‘What about the soup? Who ate that?’
 B: ‘Well, I don’t know, but I believe...’
 ??dat *NIEMAND* helaas van *de bonen* meer gegeten heeft dan vorig jaar.
that nobody unfortunately from the beans more eaten has than last year

We may, therefore, conclude that our analysis of subjective categories receives further support from the fact that it makes correct predictions regarding interactions between such categories and topics and foci.

4.2 Each and EVERY

Consider finally the two remaining exceptions to von Stechow and Iatridou’s (2003) ECP:

- (74) Each girl might be in love with John, but some of them aren’t. (Hacquard, 2006)
 (75) EVERY party guest might be the murderer. (Anand & Hacquard, 2009)

The strongly distributive universal QP in (74) is able to out-scope the epistemic modal. In (75), a prosodically prominent universal QP can out-scope the same modal. Given the subjective–objective distinction discussed above, one possible explanation for these apparent counterexamples to the SCP would be that the epistemic modal is interpreted objectively and therefore may take narrow scope with respect to the QPs. After all, *might* is one of those modals that allow for an objective interpretation, as indicated by the fact that it can occur in the antecedent of a conditional and a question.

- (76) a. If it might rain tomorrow, people should take their umbrellas.
 b. Might it rain tomorrow? (Ernst, 2009)

We do not think, however, that this is the right explanation. This is because the scope judgements for the examples in (74) and (75) are unaffected by embedding under a doxastic predicate like *believe*, which is known to force a subjective reading of the modal (Anand & Hacquard, 2009).

- (77) I believe that each girl might be in love with John, but some of them aren’t.

- (78) I believe that EVERY party guest might be the murderer.

We speculate that the universal QPs out-scope the subjective epistemic categories in (77) and (78) because they undergo a type shift in virtue of being interpreted as a topic (see Krifka (2001) for a proposal). This operation removes them from the scope of the subjective category: in effect, they are interpreted as hanging topics. On this analysis, the predication that follows the QP is about the elements in the restrictor set rather than about the QP itself. A universal QP with a strongly distributive reading (e.g. (74)) is thus an ideal candidate for such interpretation; an example like (74) ends up being interpreted as in (79).

- (79) As for girl x, x might be in love with John & as for girl y, y might be in love with John, & as for girl z, z might be in love with John, etc.

Suppose now that the maximum domain with which a subjective attitude can be connected syntactically corresponds to a clause, and nothing more; the raising case in (10), where the raising adjective *certain* does not extend its scope past the raised QP, supports this view. Since hanging topics are extra-clausal, they must necessarily be outside the domain with which the subjective attitude is connected, and hence can out-scope the subjective category in the clause the two are associated with.

5 Concluding remarks

We have argued that the interpretation of subjective categories is constrained by the *Subjective Containment Principle*. The proposal is motivated by the fact that operators with a subjective interpretation, in particular, evaluative adverbs, MC comparatives, outside negation and subjective epistemic modals (there may be more) consistently out-scope non-subjective ones, including QP, tense, question and focus operators. Categories that adhere to the SCP further give rise to scope freezing effects. Thus, although subjective categories may not contribute a truth-conditional meaning component, they do affect truth conditions through the SCP. We have shown that the data can be captured if scope extension is modeled as index percolation constrained by a minimality condition (the CSS), as in Neeleman and Van de Koot (2012).

References

- Anand, P., & Hacquard, V. (2009). Epistemics with Attitude. In T. Friedman & I. Satoshi (Eds.), *Semantics and Linguistic Theory (SALT)* (Vol. 18, pp. 37-54). CLC Publications.
- Brennan, V. (1997). Quantificational Modals. *Linguistic Inquiry* 28(1), 165-69.
- Büring, D., & Gunlogson, C. (2000). Aren't Positive and Negative Polar Questions Really the Same?. Ms. USCS/UCLA.
- Butler, J. (2003). A Minimalist Treatment of Modality. *Lingua* 113, 967-96.
- Chomsky, N. (1976). Conditions on Rules of Grammar. *Linguistic Analysis* 2, 303-51.
- Constantinou, H., & Van de Koot, H. (2015). Epistemic Containment and the Encoding of Scope. In T. Bui, & D. Ozyildiz (Eds.), *NELS* (Vol. 45, pp. 151-64). Amherst, MA: GLSA.
- Drubig, B. H. (2001). On the Syntactic Form of Epistemic Modality. Ms. University of Tübingen. <http://www.lingexp.uni-tuebingen.de/sfb441/b2/papers/DrubigModality.pdf>.
- Ernst, T. (2002). *The Syntax of Adverbs*. Cambridge, UK: Cambridge University Press.
- . (2009). Speaker-Oriented Adverbs. *Natural Language and Linguistic Theory* 27(3), 497-544.
- Fox, D. (1999). Reconstruction, Binding Theory, and the Interpretation of Chains. *Linguistic Inquiry* 30(2), 157-96.
- . (2000). *Economy and Semantic Interpretation*. Cambridge, MA: MIT Press.
- Giannakidou, A., & Yoon, S. (2011). The subjective mode of comparison: Metalinguistic comparatives in Greek and Korean. *Natural Language & Linguistic Theory* 29, 621-655.
- Groenendijk, J., & Stokhof, M. (1984). Studies on the semantics of questions and the pragmatics of answers. Doctoral dissertation. Amsterdam: University of Amsterdam.
- Hacquard, V. (2006). Aspects of Modality. Doctoral dissertation, MIT.
- Halliday, M. (1970). Functional diversity in language as seen from a consideration of modality and mood in English. *Foundations of Language* 6, 322-361.
- Hoye, L. (1997). *Adverbs and modality in English*. Harlow: Addison-Wesley, Longman
- Hoji, H. (1985). Logical Form Constraints and Configurational Structures in Japanese. Doctoral dissertation, University of Washington.
- Holmberg, A. (2015). The Syntax of Yes and No. Ms. Newcastle University.
- Huang, J. C.-T. (1982). Logical Relations in Chinese and the Theory of Grammar. Doctoral dissertation, MIT.
- Huitink, J. (2008). Scoping over Epistemics in English and Dutch. *Current issues in unity and diversity of languages*, 2077-2089. Radboud University Nijmegen. <http://www.ryukyu.bun.kyoto-u.ac.jp/infling/events/cil18/huitink.pdf>.

- Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Karttunen, L. (1977). The syntax and semantics of questions. *Linguistics and Philosophy* 1, 3-44.
- Kiss, K. (1998). Identificational focus versus information focus. *Language* 74, 245-273.
- Kratzer, A. (1991). Modality. In A. von Stechow, & D. Wunderlich (Eds.), *Semantics: An International Handbook of Contemporary Research* (pp. 639-650). Berlin: De Gruyter.
- . (2002). The notional category of modality. In P. Portner, & B. Partee (Eds.) *Formal semantics: the essential readings*. Oxford: Blackwell.
- Krifka, M. (2001). Quantifying into question acts. *Natural Language Semantics* 9, 1-40.
- . (2007). Basic notions of information structure. In C. Féry, & M. Krifka (Eds.), *Interdisciplinary studies of information structure 6*. Potsdam.
- Ladd, D. R. (1981). A First Look at the Semantics and Pragmatics of Negative Questions and Tag Questions. *Papers from the ... Regional Meeting. Chicago Linguistics Society* (Vol. 17, pp. 164-71). Chicago, Ill.
- Lakoff, G. (1972). On Generative Semantics. In D. Steinberg, & L. Jakobovits (Eds.) *Semantics* (pp. 232-96). Cambridge: Cambridge University Press.
- Leech, G. (1971). *Meaning and the English Verb*. London: Longman.
- Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.
- Nuyts, J. (2001). *Epistemic modality, language, and conceptualization*. Amsterdam: John Benjamins.
- May, R. (1977). The Grammar of Quantification. Doctoral dissertation, MIT.
- McCawley, J. (1968). The role of semantics in a grammar. In E. Bach, & R. Harms (Eds.), *Universals in linguistic theory* (pp. 124-69). New York: Holt, Rinehart, and Winston.
- McDowell, J. (1987). Assertion and Modality. Doctoral dissertation, University of Southern California.
- Neeleman, A., & Van de Koot. (2008). Dutch Scrambling and the Nature of Discourse Templates. *The Journal of Comparative Germanic Linguistics* 11(2), 137-89.
- . (2012). Towards a Unified Theory of Contrast and Scope. In A. Neeleman, & R. Vermeulen (Eds.) *The Syntax of Topic, Focus, and Contrast: An Interface-Based Approach* (pp. 39-76). Berlin: Mouton de Gruyter.
- Palmer, F. (1986). *Mood and Modality*. Cambridge: Cambridge University Press.
- Papafragou, A. (2000). *Modality: issues in the semantic-pragmatics interface*. Amsterdam: Elsevier.
- . (2006). Epistemic Modality and Truth Conditions. *Lingua* 116(10), 1688-1702.
- Picallo, M. C. (1990). Modal Verbs in Catalan. *Natural Language and Linguistic Theory* 8(2), 285-312.
- Reinhart, T. (1981). Pragmatics and Linguistics: an analysis of sentence topics. *Philosophica* 27(1), 53-94.
- . (1983). *Anaphora and Semantic Interpretation*. London: Croom Helm.
- . (1995). *Interface Strategies-Reference-set computation*.
- . (2006). *Interface Strategies*. Cambridge, MA: MIT Press.
- Romero, M., & Han, C. (2004). On negative yes/no questions. *Linguistics and Philosophy* 27(5), 609-658.
- Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- . (1996). Sentence prosody: Intonation, stress and phrasing. In J. A. Goldsmith (Ed.), *The handbook of phonological theory*. Oxford: Blackwell.
- Sudo, Y. (2013). Biased Polar Questions in English and Japanese. In D. Gutzmann, & H.-M. Gaertner (Eds.), *Beyond Expressives: Explorations in Use-Conditional Meaning*. Current Research in the Semantics/Pragmatics Interface (CRiSPI) (Vol. 28). Leiden: Brill.
- Tancredi, C. (2007). A Multi-Modal Theory of I-Semantics. Ms. University of Tokyo.
- Tomioka, S. (2009). Contrastive topics operate on speech acts. In M. Zimmermann & C. Féry (Eds.), *Information structure from different perspectives* (pp. 115-138). Oxford: Oxford University Press.
- Von Stechow, K., & Iatridou, S. (2003). Epistemic Containment. *Linguistic Inquiry* 34(2), 173-98.
- Van de Koot, H., Constantinou, H. & Yan, Z. (2015). Epistemic Bias in Embedded Polar Questions. *UCL Working Papers in Linguistics* (Vol. 27, pp. 1-25).
- Williams, E. (1994). *Thematic Structure in Syntax*. Cambridge, MA: MIT Press.
- Wurmbrand, S. (1999). Modal Verbs Must be Raising Verbs. In S. Bird, A. Carnie, J. D. Haugen, P. Norquest (Eds.), *WCCFL 18 Proceedings* (pp. 599-612). Somerville, MA: Cascadilla Press.

Examining the nature of referential metonymy

Josephine Bowerman

Abstract

Working within the framework of Relevance Theory, I argue that referential metonymy can be analysed in properly inferential terms, and can be treated as a distinct type of lexical innovation, alongside lexical ‘modulation’ (narrowing and broadening) and motivated neologism. I address such issues as motivations for referring metonymically and factors constraining the choice of metonymic referring term. In addition, I consider how the interpretation of referential metonymy may routinize, and how frequency of use and routinization may affect the storage and processing of a metonymic interpretation, examining empirical data on metonymy processing from Frisson and Pickering (1999).

Keywords: Referential metonymy; Relevance Theory, lexical innovation; underspecification

1 Introduction

Metonymy is traditionally treated as merely a rhetorical device or figure of speech in which, for the name of a thing, we substitute the name of an attribute of that thing, or of something closely related to it (e.g. *hands* for *workers*, or *skirt* for *woman*). In order to further our knowledge of the phenomenon, I set out to elucidate the nature of referential metonymy, exemplified in (1)-(3), wherein the referring terms ‘*ham sandwich*’, ‘*Shakespeare*’ and ‘*muscle*’ are being used metonymically because they express attributes of/things relating to the actual target referent:

- (1) The *ham sandwich* has left without paying.
- (2) Put *Shakespeare* on the top shelf.
- (3) If they won’t pay, send in the *muscle*.

In (1) the referent is a restaurant customer but the referring term is the word for his food order; in (2) the referent is a book but the referring term is the name of the author of the text; and in (3) the referents are hired tough guys but the referring term is the word for their most notable property, i.e. their sizeable muscles.

I restrict my focus to referential metonymy because it is the prototypical example of metonymy and as such has been studied in most depth. Working within the framework of Relevance Theory (hereafter, RT) I aim to discover whether referential metonymy can be treated as a type of lexical innovation in its own right, and whether it is amenable to a fully inferential analysis.

I will show that referential metonymy is a motivated phenomenon in which the linguistic properties of the metonymically-used word provide evidence of the speaker’s intended referent, such that the target interpretation can be recovered inferentially. The comprehension of referential metonymy is therefore compatible with the RT view of utterance processing, which sees interpretation as fully inferential (Sperber & Wilson 1986/1995, 1998, 2002; Wilson & Sperber 2002). I will also argue, by comparison with other lexical pragmatic phenomena, namely ‘modulation’ (narrowing and broadening) and motivated neologism, that referential metonymy is indeed a distinct type of lexical innovation which involves ‘repurposing’ an existing expression to pick out a target entity. Finally, I will consider empirical evidence relating to my arguments.

2 Motivations for referring metonymically

Before I examine the interpretation of referential metonymy, and assess whether or not the phenomenon may be classed as an independent type of innovative word-use, I will begin by investigating why and how speakers refer metonymically. I take as my starting point RT accounts such as Papafragou (1996), Jiang (2013) and Rebollar (2015), in which referential metonymy is seen as a speaker's intentional use of an expression to refer to an entity that does not fall under the literal denotation of the expression. These accounts draw on Kaplan's (1989, p. 559-561) claim that a speaker's intention to use an expression to refer to a given entity irrespective of any prior meanings associated with, or pre-established uses of, the expression (the 'referential' intention) can dominate over the intention to use an expression with the meaning given it by the person from whom the expression was learnt (the 'attributive' intention). For example, although the established usage of the expression '*ham sandwich*' is to pick out the bread-and-meat snack, in (1) above, uttered in a restaurant context, the speaker's critical intention in using this expression is to refer to a particular customer who ordered a ham sandwich. Also pertinent is Donnellan's (1966) slogan: expressions denote, people refer. Consider (4):

(4) I spent the summer reading *Dickens*.

While the encoded meaning of the expression '*Dickens*' denotes the flesh-and-blood author Charles Dickens, the person who utters (4) uses this expression to refer to the works of Charles Dickens. This highlights a distinction between the linguistically-specified referent of an expression and the entity/entities which a speaker can use the expression to pick out. It also raises the question of why this distinction should arise: what reasons might a speaker have to fix a novel referent for an existing expression?

Speakers may refer metonymically in situations where referring literally would impose unnecessary extra processing costs on the audience. For example, suppose that a waitress in a crowded restaurant full of businessmen on their lunch break intends to refer to a specific diner. However, using an expression which literally refers to her target entity (e.g. '*businessman*') would, in the communicative context, lead to referential ambiguity because it would not identify a unique individual, and would thus require of the audience extra cognitive effort to resolve the ambiguity and work out the intended referent. Furthermore, in a hectic restaurant environment the audience, i.e. other waitresses, will have to deal with many other cognitive tasks in addition to utterance interpretation, and therefore may not have the time or processing resources to interpret a literal referring expression such as '*the man who ordered the ham sandwich*' which, although it provides disambiguating information, is long and syntactically complex.

Given the RT assumptions that all acts of ostensive communication, including utterances, communicate a presumption of their own optimal relevance (i.e. that the ostensive stimulus is both sufficiently relevant to be worth the audience's processing effort, and also the most relevant stimulus the communicator is willing and able to produce), and that the overall relevance of an utterance is a function of cognitive effects against processing effort (Sperber & Wilson, 1986/1995), the extra effort required for ambiguity resolution or processing a complex expression may result in the utterance failing to meet expectations of relevance, such that the audience abandon utterance processing without having recovered the speaker's intended referent. It is therefore in the best interests of the speaker to choose a referring expression that makes reference resolution easy— quick and unambiguous. Minimising the effort required for interpretation increases the likelihood that the audience will maintain

attention and fully process the speaker's utterance, thus increasing the likelihood that the intended referent will be recovered (Allott, 2013, p. 81).

Further, as Papafragou (1996, p. 186) notes, a speaker might refer metonymically to provide access to more contextual implications than would be available if reference were made literally, or to express attitudinal/affective information. By way of illustration, consider an alternative scenario. A guest at an office Christmas party is talking to her friend and, wishing to refer to their colleague Dave, who is wearing a distinctive outfit that night, utters (5):

(5) Look at the *green trousers* dance!

While it would be quicker and easier to refer to Dave by simply using his name, which is known to both speaker and addressee, picking him out by using the expression '*green trousers*' creates a humorous effect through reducing a human being to his garments. It also encourages the addressee to explore more carefully the implications of the referring expression '*green trousers*' and thereby derive extra conclusions, e.g. that Dave is, like his outfit, loud and tasteless. Here, the extra effects derived presumably outweigh any extra processing costs that their derivation incurs. Just as minimising the effort required for utterance processing is in the speaker's best interests, so too is maximising the benefits of the enterprise for the audience: providing extra, useful information will also increase the chances that the audience will pay attention, fully process the utterance, and remember the message that the speaker wants to convey (Allott, 2013, p. 81).

Finally, Jiang (2013, p. 11-13) and Rebollar (2015, p. 196-7) claim that speakers use metonymy to fill gaps in the lexicon if there is no existing word to refer to the target entity. For example, if the two friends at the office party do not know the name of the enthusiastically jitterbugging man in emerald flares, they must use an alternative referring expression which allows for economical-efficient referent identification, and possibly also implies additional conclusions. Summarising these points, we could thus say that the overarching motivation for referential metonymy is the speaker's realisation that, with respect to the entity she has in mind and intends that her audience should pick out, the available means for referring to that entity are not fully adequate.

3 Selecting a referring expression

When using the available linguistic resources as they are ordinarily used would prevent the speaker from fulfilling her communicative intention, as regards getting her audience to come to pick out a specific entity, the speaker's challenge is then to find a better way of referring. The most economical-efficient way to do so is to 'repurpose' an existing expression for the task. Yet this leads us to question how the existing term which is to be 'repurposed' in referential metonymy is chosen. For example, why does the restaurant waitress who wishes to refer to a specific customer use the expression '*ham sandwich*' to pick out her target entity?

If metonymy is indeed to optimise reference-making in the absence of an adequate literal way to refer, the speaker must ensure that although the referring term she uses does not literally identify her target entity, it is nevertheless the best 'evidence' she is able to provide for her referential intention (and, in some cases, for her intention to achieve certain effects), given the RT view that linguistic meaning, rather than fully determining communicated content, acts as a clue or pointer to the speaker's intended meaning. The 'evidence' must allow the audience to work out the intended referent in a way that is as low in processing costs and/or as rich in beneficial effects as possible, such that relevance is maximised. This can be achieved by choosing an existing expression which refers to a cognitively salient feature/property of the

target entity, i.e. a feature/property within the mental representation of the target entity that is more highly activated or accessible than other properties. Entities can be identified through their salient properties (Rebollar, 2015, p. 195), thus by getting the audience to focus on a particular, especially attention-commanding aspect of the target entity (e.g. for the restaurant customer to whom the waitress wishes to refer, the property of ‘being a ham-sandwich orderer’), it is made easier for the audience to pick out the entity itself.

When we speak of a ‘salient’ feature/property of an entity, we must bear in mind that cognitive salience is a relative notion that depends on the demands of the situation and thus cannot be independently defined, as Jiang (2013) emphasises. While there may be properties of an entity that are ‘inherently’ attention-commanding, e.g. concrete, functional or interactional properties (see Langacker (1987, p. 385-6) and Radden & Kovecses (1999) on the general cognitive determinants of salience), ultimately, the most salient property of an entity is the property that in the context at hand is the most relevant (yielding the most cognitive effects relative to processing effort) and therefore the most worthwhile for an agent to attend to. Importantly, the inherent salience of a property does not guarantee its usefulness in every situation. For example, if the situation demands economical-efficient identification of a target entity, the most relevant property of the entity will be one which, regardless of whether or not it is inherently salient, is individuating, i.e. in the context, uniquely picks out the entity without introducing any referential ambiguity which would require extra processing effort to resolve. Alternatively, if the situation requires us to derive certain implicit conclusions concerning a target entity, it may be the case that the assumptions associated with an inherently salient property of the entity in question would not allow us to draw the intended conclusions, and thus the inherently salient property would not be the most relevant property of the entity.

A speaker whose communicative goal is to achieve economical-efficient reference resolution and/or to convey contextual implications about a particular entity must therefore, in making reference to her target entity, get her audience to focus on the property of the entity that is most relevant with respect to her intentions. To do this, she must ‘repurpose’ as her referring expression the existing expression that, in the communicative context, introduces a property of the target entity which is uniquely identifying, thereby making the entity maximally accessible and minimising the effort required to pick it out, and/or which achieves additional effects.

We are now able to account for why the waitress chooses the referring expression ‘*ham sandwich*’ to identify her target customer. In the restaurant context, ‘food ordered’ is already a highly relevant property of customers, as attending to this property enables waitresses to plan and act successfully (taking the correct order to the correct customer). Moreover, for any given customer, this property is likely to be one of the only properties to which a waitress attends. This makes ‘food ordered’ an effective individuating property of customers. Therefore, by referring to a specific target customer using a term e.g. ‘*ham sandwich*’ that expresses the customer’s food order, the waitress is likely to have provided sufficient evidence of her referential intention to enable an addressee to identify the intended customer without incurring unnecessary processing costs. In addition, given sufficient time and processing resources, addressees can consider the assumptions associated with the referring term ‘*ham sandwich*’ and derive extra conclusions, e.g. that the intended referent—the customer who ordered the ham sandwich—is, like his choice of snack, traditional and boring. Although on this occasion drawing such conclusions is not central to the success of the metonymic use of ‘*ham sandwich*’ because the speaker’s primary goal is to ensure economical-efficient identification of her intended referent, any further implications that addressees can derive may increase the overall relevance of the utterance.

4 Interpreting referential metonymy

Having considered motivations for and factors constraining the production of referential metonymy, we can now explore how addressees interpret metonymically-used referring expressions. First, I will briefly present the RT approach to utterance interpretation. RT claims that in comprehension, our goal is to construct a hypothesis about the speaker's intended meaning that satisfies the expectations of relevance raised by the utterance. This task comprises three sub-tasks: (i) constructing hypotheses about the explicit content of the utterance, (ii) constructing hypotheses about intended contextual assumptions, and (iii) constructing hypotheses about intended contextual implications (implicated conclusions). These hypotheses are developed on-line and in parallel (Sperber & Wilson, 2002, p. 261). Considerations of relevance may lead us to expect a particular conclusion, or type of conclusion; these expectations may contribute via backwards inference to the identification of explicit content and contextual assumptions which act as explicit and implicit premises that warrant, via forwards inference, the anticipated conclusion (Carston, 2004). Utterance interpretation is therefore fully inferential: each sub-task in the overall process of constructing a plausible interpretation for the speaker's utterance involves a non-demonstrative inference process (Sperber & Wilson, 2002, p. 262). The mutual parallel adjustment of explicit content, context and cognitive effects is constrained by the RT comprehension strategy, which states that interpreters should follow a path of least effort and test interpretive hypotheses in order of accessibility, stopping when expectations of relevance are satisfied (or abandoned) (Sperber & Wilson, 2002, p. 259).

Let us apply the RT approach to the following examples of referential metonymy to see that, although these cases differ in terms of 'creativity', and in terms of how addressees make use of the metonymic referring expression in reference resolution, they are nevertheless both interpreted using the same mechanism of mutual parallel adjustment.

We will begin with (4), repeated here as (6), which is a less creative instance of referential metonymy:

(6) I spent the summer reading *Dickens*.

The speaker of (6) does not intend '*Dickens*' to be understood literally, as referring to Charles Dickens. This would yield an ungrammatical, nonsensical interpretation of (6) that would fail to satisfy our expectations of relevance. Rather, her intended referent is the works of Dickens. However, she has chosen '*Dickens*' as her referring expression because the property of her target entity that this word expresses (the name of the author of the works) is individuating, thereby enabling us to pick out that entity with minimum effort; and it allows us to derive her intended conclusions.

In order to grasp the intended referent of the metonymic use of '*Dickens*' in (6), we 'zoom in' on especially accessible encyclopaedic assumptions of the encoded meaning of the word, e.g. the information that Dickens was an author. Further, given that online processing proceeds incrementally (e.g. Just & Carpenter, 1980; Tyler & Marslen-Wilson, 1977), incoming words may influence predictions concerning the intended interpretation of the utterance. In (6), when the verb '*read*' is encountered, its selectional properties will make certain interpretive hypotheses about the explicit content of the utterance highly accessible, for instance that if the verb is followed by a noun-phrase complement, that noun phrase will refer to something readable. These assumptions and hypotheses combine in mutual parallel adjustment. Inference runs forwards from the hypothesis that the noun phrase will refer to a readable object, and from the assumption that Dickens was an author: authors write readable works, therefore the most plausible referent for '*Dickens*' is the works of Dickens. There may also be backwards

inference from expected cognitive effects (for example the anticipatory hypothesis that the speaker's utterance will achieve relevance by describing an activity that took up her whole summer), which contributes to the identification of explicit content and contextual assumptions that together will warrant the derivation of the contextual implications of the utterance. We therefore derive an interpretation for (6) where the term '*Dickens*' is understood as picking out the works of Dickens. This allows for sound inference to our expected conclusions: given that the works of Dickens are numerous and lengthy, we can infer that the speaker of (6) was occupied with her reading for the entirety of the summer. As this derivation shows, the comprehension of (6) is a fully inferential process, with both explicit and implicit content worked out by abductive reasoning.

Now consider a more novel example, (7), uttered in a crowded bar to refer to a hirsute man sitting nearby:

(7) The *beard* looks miserable.

Once again, the linguistic context will influence the accessibility of interpretive hypotheses about the explicit content of (7). For example, the verb phrase '*looks miserable*' may lead us to predict that its subject will be human. However in this instance, to identify the speaker's intended referent we do not need to focus on contextually relevant encyclopaedic assumptions about the referring expression '*beard*'. This is because a literal beard is physically present and directly perceivable, and the speaker's use of the word '*beard*' draws our attention to this entity. Crucially, when we attend to the beard, we simultaneously attend to an individual who could plausibly be the speaker's intended referent, by virtue of the fact that the beard is on his face. Thereby, through her choice of referring expression, the speaker leads us to focus on her target entity. Moreover, the speaker has presumably chosen the word '*beard*' as her referring expression because no other men present in the bar but her intended referent have facial hair, and the word '*beard*' therefore expresses an individuating property of her target individual. Thus, when we attend to the literal beard we can see nearby, we simultaneously attend to the only person who could be the speaker's intended referent.

Given that interpretation follows a path of least effort, the first interpretive hypotheses to be tested and developed by the mutual parallel adjustment process will be the highly accessible hypotheses that the subject of (7) is human and that the speaker's intended referent is most likely to be the only bearded man visible in the immediate environment. Inference runs forward from these hypotheses to the conclusion that the intended referent of '*beard*' is indeed the bearded man nearby. In addition, backwards inference from specific expectations about the cognitive effects that the utterance of (7) will achieve, for example the expectation that the speaker will provide us with useful information about other people in the bar, may also help us to recover an interpretation in which the referent of '*beard*' is taken to be a fellow drinker. Having derived the explicit content of the utterance, we can go on to infer further contextual implications which will contribute to the overall relevance of the utterance, e.g. that if the '*beard*', i.e. the bearded man nearby, is looking miserable, we should go over and attempt to cheer him up. Here, too, we can see that each of the sub-tasks contributing to the recovery of a plausible reading of (7) that satisfies our expectations of relevance involves non-demonstrative inference, such that the overall comprehension process is fully inferential.

5 Comparison with lexical modulation

As my analyses of the above examples of referential metonymy make clear, it is possible to provide an adequate explanation of the phenomenon in RT terms without claiming that in the

course of interpretation, an occasion-specific sense, or ad-hoc concept, is constructed from the linguistically-specified meaning of the metonymically-used referring term as the result of interaction between encoded information, contextual information and relevance-based expectations (see e.g. Wilson and Carston (2007) on ad-hoc concepts). That is to say, in order to deal with the phenomenon on a properly inferential approach, there is no need to treat it as a variety of lexical ‘modulation’ in which the linguistically-specified meaning of the metonymically-used referring term is narrowed, such that the communicated sense is more specific than the encoded sense; broadened, such that the communicated sense is more general than the encoded sense; or a combination of the two (Carston, 1997). Indeed, a modulation account appears to be wrong for referential metonymy.

Consider first narrowing, which involves using a word to pick out a proper subset of the items that fall under its linguistically-specified denotation, as in (8) and (9):

- (8) I’ve got a *temperature*.
 (9) Josie is avoiding *bachelors*.

(8) communicates the message that the speaker has a specific kind of temperature, i.e. a higher-than-normal one. (9), given a particular set of background assumptions, could be understood as conveying that Josie is steering clear of a certain subtype of unmarried men, for example the kind that still live with their mothers. The words ‘*temperature*’ and ‘*bachelors*’ come to express the ad-hoc concepts TEMPERATURE* and BACHELORS*, which have a restricted denotation compared to that of the encoded meaning from which they are derived. In order to construct such ad-hoc concepts, contextually relevant encyclopaedic properties are added to the encoded meaning, e.g. for ‘*bachelors*’, the property of living with their parents.

In referential metonymy, however, the overall effect is not to highlight a proper subpart of the linguistically-specified denotation of the metonymically-used referring term. It is plainly not the case that the intended referent in the ‘*ham sandwich*’ example—the customer who ordered the ham sandwich—could be thought of as a subtype of ham sandwiches. Likewise, in (10), there appears to be no way in which the entity that the speaker intends to refer to by her use of the word ‘*Mozart*’, i.e. a piece of music by that composer, is a specific subtype of the entity which is picked out by the encoded meaning of the word, i.e. the man Wolfgang Amadeus Mozart:

- (10) I’m playing *Mozart*.

Rather, the metonymic use serves to pick out an entity which lies outside of the linguistically-specified denotation of the referring term.

A further important difference between referential metonymy and narrowing is that in referential metonymy, the referring term itself continues to mean what it means on its normal usage. It does not come to express, through the addition of encyclopaedic properties, an ad-hoc concept which is a more specific version of the sense it is ordinarily used to convey. All that changes is the entity that it is used to pick out. It can thus be concluded that referential metonymy cannot be explained as a variety of narrowing.

We turn next to broadening. Broadening results in a widening of the linguistically-specified denotation of the broadened term, such that the term can apply to entities that lie outside of its linguistically-specified denotation. The denotation of the ad-hoc concept which results from broadening contains all the entities contained in the denotation of the original literal concept, plus more. In (11) for example, ‘*boiling*’ expresses the ad-hoc concept BOILING*, which includes in its denotation not only things that are boiling in the strict sense (having a

temperature of 100 C) but also further cases of things that are hot but not boiling, such as very warm bathwater.

- (11) The water is *boiling*.

Yet compare the metonymic use of the expression ‘*ham sandwich*’, which denotes, or refers to, a person, not a person in addition to all the bread-and-ham snacks. The referring expression comes to pick out an entity that does not fall within its linguistically-specified denotation, which may appear similar to the way in which a broadened expression can apply to entities that lie outside of its linguistically-specified denotation. Importantly, however, the metonymic denotation does not stand in a superset relation with the denotation of the encoded literal concept, unlike the denotation of a broadened concept.

Broadening crucially involves dropping logical/definitional properties from the encoded meaning of the broadened expression, e.g. for ‘*boiling*’, the property of being at 100 C. However, this again seems very different to what occurs in the comprehension of referential metonymy. Consider some different examples: (12), a familiar case, where the word ‘*Vietnam*’ is used to refer to the Vietnam War; and (13), a more novel case, where the word ‘*sheep*’ is used to refer to someone born in the Year of the Sheep (originally from a television program; cited in Gerrig (1989) and Papafragou (1996)).

- (12) *Vietnam* was a military disaster.
 (13) You should avoid marrying a *sheep* at all costs.

In order to derive an interpretation of (12) and (13) that satisfies our expectations of relevance, we cannot take the literal referent of ‘*Vietnam*’ and ‘*sheep*’ to be the speaker’s target entity, and must find a novel referent for these expressions. However, in (12), we do not come to pick out an entity of which the property of being a military disaster can appropriately be predicated and which is therefore most plausibly the speaker’s intended entity by dropping logical properties of the encoded meaning of ‘*Vietnam*’ (i.e. the property of being an South-East Asian country), and possibly also adding encyclopaedic properties in the process, to end up with an ad-hoc concept, VIETNAM*, that includes the target entity in its denotation. Likewise, although (13) is a more creative instance of referential metonymy, in order to identify the speaker’s intended referent we do not need to drop from the encoded meaning of ‘*sheep*’ the logical property of literal sheephood, such that ‘*sheep*’ comes to convey an ad-hoc concept, SHEEP*, which can apply to human beings.

Indeed, we do not modulate the linguistically-specified meaning of the metonymically-used referring expression in any way. The referring expression conveys the same sense when used metonymically as it does when used literally. Although the referring expression comes to pick out a novel referent, this is not achieved by dropping logical properties from its encoded meaning to derive an ad-hoc concept whose denotation is sufficiently expanded compared to that of the input meaning as to include the target entity. Rather, the communicative context makes certain encyclopaedic assumptions of the encoded meaning of the referring expression especially relevant and accessible: in (12), that a disastrous war took place in Vietnam, and in (13), that the sheep is a Chinese zodiac symbol. These assumptions provide a uniquely identifying property of the speaker’s intended referent—the location of the target war in (12), and the Chinese zodiac sign of the target individual in (13)—thereby enabling the audience to economically-efficiently pick out that referent. Thus it is by virtue of the unmodulated linguistically-specified meaning of the metonymically-used referring expression that recovery of a novel referent is accomplished.

This shows that referential metonymy cannot be treated as a variety of broadening. It therefore appears that referential metonymy is not amenable to an analysis in terms of either type of lexical ‘modulation’, narrowing or broadening, as also noted by e.g. Recanati (2004, p. 26) and Wilson and Carston (2007, p. 254). Whereas in cases of modulation the linguistically-specified meaning of an expression is pragmatically adjusted to derive the speaker’s intended interpretation, in cases of referential metonymy, no such ad-hoc concept construction is involved in the recovery of the speaker’s target referent. Note, too, that not only is no new occasion-meaning expressed by the metonymically-used referring term, but my derivations in §4 also assume no change to the standing meaning of the metonymically-used referring expression.

6 Comparison with neologism

Given that it seems referential metonymy is not reducible to lexical ‘modulation’, we could argue that the phenomenon is a distinct type of innovative word-use, involving the fixing of a novel referent for the metonymically-used referring expression. However, it may be possible to treat referential metonymy as a variety of neologism, or word coinage, which would undermine the hypothesis that referential metonymy is an independent phenomenon of innovation. For example, Papafragou (1996, p. 182) claims that metonymically-used referring terms function as ‘newly coined names’ for the speaker’s intended referent, and Rebollar (2015, p. 196) speaks of ‘ad-hoc name creation’ for identification purposes. In addition, the need to fill gaps in the lexicon is identified as a motivation for metonymy by Jiang (2013) and Rebollar (2015), which implies that metonymic usages can be seen as supplying new words.

Wilson (in Wilson & Falkum, 2015) goes further, explicitly arguing that metonymic usages arise through spontaneous processes of word coinage and should be analysed as ‘motivated’ neologisms, i.e. neologisms in which the linguistic properties of the new word (here, the metonymically-used referring term) provide a clue to the speaker’s intended meaning. This argument seems to be driven by two main goals. Firstly, Wilson aims to provide an account of metonymy that is not formulated in terms of lexical modulation, but also does not resort to citing code-like ‘transfer of meaning’ rules (‘author for work’, ‘foodstuff for consumer’, etc.). While these mappings are central to the cognitive linguistics approach to metonymy (see e.g. Lakoff, 1987; Kövecses & Radden, 1998; Radden & Kövecses, 1999; Panther & Thornburg, 2003), wherein metonymy is seen as the natural language reflex of what is primarily a cognitive phenomenon involving a relationship of contiguity between two items in a single cognitive domain, such a code-based account seems only able to describe the input-output relation for certain highly frequent, familiar metonymic usages, and cannot account for how these relations arise nor how they are comprehended when they are first used, unless we adopt the unsatisfactory position that these rules are innate. Secondly, Wilson—while acknowledging the important role for associative relations, i.e. spreading activation patterns, in metonymy comprehension—seeks to accommodate these within a fully inferential approach. Wilson thus claims that metonymy interpretation involves the same mechanism as is used in the interpretation of other motivated neologisms such as denominal verbs or noun-noun compounds. That is to say, a hearer using the RT comprehension heuristic will treat the metonymically-used referring term as an ostensive stimulus and infer the intended meaning by mutual adjustment of explicit content, context and cognitive effects, following a path of least effort.

Wilson may also be influenced by evidence from acquisition which shows that even prelinguistic infants spontaneously produce motivated ostensive stimuli, including stimuli that seem to be based on the ‘metonymic’ principle of picking out a salient property/feature of the

target referent as a way to identify that referent, e.g. smacking lips for 'food' or making a 'vroom vroom' noise for 'car' (Falkum, Recasens & Clark, 2013). It is likely that infants 'innovate' in this way to compensate for a lack of established stimuli for conveying the intended message. The same applies to older infants' uses of novel denominal verbs, e.g. '*to gun*' for 'shoot', and noun-noun compounds, e.g. '*clown boy*' for 'boy who is a clown' (Falkum, Recasens & Clark, 2013): young children use language creatively to fill vocabulary gaps, when they want to pick out a specific referent but lack the words to refer to their target entity. This suggests that metonymic usages and phenomena which are traditionally taken as clear cases of neologism have a common motivation. Adults may also be driven to innovate for the same reason, but it is likely that their motives are more varied and less 'egocentric': adults may take into account the needs of their audience, and innovate in order to reduce processing costs and/or increase positive cognitive effects. Given that children's creative uses of language, including metonymic uses, fill vocabulary gaps, their innovations are clearly creating 'new words'. Adults typically have an established word in cases where a child might employ an innovative usage, therefore the novelty of the child's innovation is more apparent. This highlights the sense in which metonymic uses can be seen as new coinages.

Yet, although we can identify phenomena such as denominal verbs and noun-noun compounds as examples of neologism/new coinage, there nevertheless appears to be a lack of clarity as to how we should understand the notion of 'new coinage'. For instance, on the basis of the observation that children use language creatively when they are lacking a 'label' for, or way to pick out, some entity/concept they want to talk about, we could view a 'new coinage' as a 'new name/label'. This interpretation would also encompass adult's lexical innovations. For example, a metonymic usage motivated by the need to refer more efficiently or with more effects could be seen as the creation of a new name/label for the target referent, simply with the caveat that a new name/label is needed not because of a vocabulary gap, but because the existing way to pick out the target referent is inadequate in the communicative context. Accepting this understanding of 'new coinage' would mean that the notion applies to four distinct phenomena: metonymy, modulation, classical neologisms like denominal verbs, and out-and-out inventions e.g. *wilgrid* 'thorny, knotty problem'. These phenomena could be seen as falling on a continuum of creativity, with respect to how much is changed in order to coin a new name/label. Metonymy, at the lower end, only changes the referent of the metonymically-used word. At the upper end, invention is the most creative way to coin a new name/label: rather than repurposing an existing word, we produce an entirely novel form.

Invention aside, metonymy, modulation and classical neologisms all involve reusing an existing expression in some way; in Wilson's terms, these are the 'motivated' phenomena, in that the linguistic properties of the repurposed word provide evidence of the speaker's intended meaning (presumably, this is what influences the speaker's choice of word to be repurposed). These phenomena would therefore be amenable to a fully inferential analysis, and Wilson's goal of getting away from an unexplanatory analysis of metonymy which appeals to 'transfer of meaning' rules would thus be met. This seems to be a significant advantage of understanding 'new coinage' as 'new name/label'. However, classing both metonymy and modulation as new coinages may be an unwanted outcome for Wilson, who seems to want to keep metonymy firmly apart from modulation. Given that metonymy, modulation and classical neologisms would all receive an inferential analysis on this approach, the challenge would then be to find a way of clearly distinguishing the three phenomena. This would involve showing adequately that although the interpretive mechanism, i.e. the mutual parallel adjustment of explicit and implicit content which driven by the search for relevance and constrained by the RT comprehension procedure, is the same in all three cases, the outcomes of the operation of that mechanism are different (thereby meeting Wilson's goal of not treating metonymy as a case of

narrowing or broadening). It would also involve considering how the demands of communication and interpretation might lead to one phenomenon rather than another.

A further problem is that there is a strong sense in which the cases with no form-changes, i.e. metonymy and modulation, are not overtly novel, especially in contrast to e.g. a new compound or an outright invention, and even when compared to denominal verbs (e.g. ‘party_V’ = to attend lots of parties_N) and deverbal nouns (e.g. ‘scratch_N’ = slight mark/injury caused by scratching_V) which, although identical to the ‘parent’ noun/verb in bare form, nevertheless appear with verbal/nominal inflections (*he partied all night; three scratches*), such that the surface form of the new word is different to that of the existing word from which it is derived. Although in metonymy and modulation the reference (metonymy) or the meaning (modulation) of the innovatively-used word changes, on the surface it retains its old form and for this reason it could be argued that there is no true ‘coinage’ taking place.

Alternatively, we could view a ‘new coinage’ as a ‘new form/meaning pairing’. That is to say, a new coinage should involve, minimally, a change to the meaning of an existing expression. Note, though, that the term ‘meaning’ can be ambiguous between ‘encoded’ (or ‘standing’) meaning and ‘expressed’ (or ‘occasion’) meaning. We must therefore clarify whether a new coinage requires a change to the standing or the occasion meaning of an expression. Armstrong (2016, p. 101-2) draws a distinction between ‘lexical shifts’, in which the semantic properties of an existing word are revised, and ‘lexical expansions’, in which a new ‘character’ is introduced into the language, claiming that only with the latter type of lexical innovation is the set of linguistic expressions available for use between a speaker and her audience enriched by the addition of a new linguistic expression not already in circulation among the interlocutors. Presumably, ‘character’ is meant in the Kaplanian sense of ‘standing meaning’, a context-invariant constraint on the range of semantic contents that can be assigned to utterances of an expression. If so, this suggests that the coinage of a new word requires a new encoded meaning.

However, given that on my account a metonymically-used referring term comes to pick out a new referent but keeps its old standing meaning, it appears that referential metonymy does not involve a new form/meaning pairing and should therefore be excluded from the class of new coinages. This would be an undesirable result for Wilson. It also seems that ‘modulation’, in which a new ‘occasion’ meaning is derived for the modulated word on the basis of its constant standing meaning, cannot be seen as a variety of new coinage. This result is less problematic for Wilson, who holds cases of modulation—old words getting new ‘meanings’—apart from ‘true’ motivated neologisms.

Yet what of the ‘form’ part of the ‘form/meaning pairing’? We could focus on the factor of overt novelty, discussed above, and argue that for a lexical innovation to count as a ‘new coinage’, not only must we end up with a new standing meaning, we must also create a new surface form. On this approach, both classical neologisms (new meaning, new form created from an old word/old words by morphosyntactic operations such as category change or compounding) and inventions (new meaning, entirely novel form) would count as new coinages. However, a serious issue immediately becomes apparent. If we take true motivated neologisms to crucially involve the creation of a new linguistic expression, there again seems to be no way in which referential metonymy could be included in this class, given the claim that in referential metonymy, only the referent of the metonymically-used word changes. Wilson describes metonymies as ‘denominal nouns’, yet the creation of such expressions does not seem to correspond to any kind of morphosyntactic operation of the sort undergone by classical motivated neologisms: there is necessarily no category change, unlike with e.g. denominal verbs, and there is no compounding, affixing or other surface evidence of a form-change. It appears that if we accept the ‘new coinage as new form’ treatment, referential

metonymy is in fact excluded from the very class of phenomena to which Wilson argues it belongs.

While this outcome would clearly be problematic for Wilson, it does not rule out a pragmatic account of referential metonymy in which the phenomenon is not reduced to modulation. In §4 I have provided fully inferential analyses of some examples of referential metonymy which, if plausible, suggest there is no need to group metonymy with other phenomena of lexical innovation such as modulation or motivated neologism in order to deal with it on a properly inferential approach. Indeed, we have seen that metonymy is not modulation, and nor does it seem to be new coinage. Rather, metonymy is plausibly a distinct motivated phenomenon alongside modulation and neologism, which is interpreted by the same single inferential mechanism of mutual parallel adjustment. While it would be undesirable to posit multiple interpretive mechanisms (for example, to claim that certain linguistic phenomena are interpreted by an associative, code-based mechanism while others are interpreted inferentially), I can see no objection to proposing that there are multiple different kinds of lexical innovation, especially given that we already take modulation alone as subdividing into narrowing and broadening, and that RT sees broadening as covering a range of cases including hyperbole and metaphor (e.g. Wilson & Carston, 2007).

Due to the demands of communication (the message a speaker wants to convey, the ways in which she can use existing linguistic resources to satisfy her communicative intention) and of interpretation (the way an audience uses a speaker's linguistic 'evidence' to recover the intended meaning given the time available, their processing resources and their own informational needs/preferences), a metonymic usage of an expression does something specific: it fixes a new referent for the expression and provides us with a novel way of picking out the target entity (this is what makes it a case of 'innovation'). Yet no new sense comes to be conveyed by the expression, and its linguistically-specified meaning does not change. Further, the creation of a new form is not part of the innovation. Referential metonymy, as a phenomenon in its own right, can therefore be clearly distinguished from modulation (old word, new occasion meaning) and neologism (new word, new standing meaning). However, this distinction can be made without also needing to claim that metonymy requires a specific interpretive mechanism involving 'transfer of meaning' rules.

7 Metonymic patterns

Using an expression metonymically to pick out a novel referent may be a one-off event, or a metonymic usage may be limited to a specific context (e.g. the '*ham sandwich*' case, which is localised to its restaurant context) and/or a specific group of people: the context creates the conditions which motivate speakers to use words metonymically, as well as providing information and assumptions which are used to interpret the metonymy. Metonymic usages may also be more general. If a particular metonymy is highly relevant and useful, it will come to be used increasingly frequently as a way to refer to the target entity, and the interpretation process will routinize. This reduces processing costs, as it means that effortful, (more) conscious derivation of the intended interpretation is no longer required on every occasion the metonymy is encountered. In addition, a consequence of routinization is that interpreters may need to rely less on contextual support, and the metonymy can be used outside of its original context.

Interestingly, exploring how metonymy interpretation routinizes could shed new light on the 'conceptual metonymies' or associative mappings (e.g. 'producer for product', 'place for inhabitants', 'clothing for wearer', etc.) that are so central to the cognitive linguistics account of metonymy. As noted, metonymy interpretation is likely to routinize when the metonymy in

question is used very frequently—whether in a specific context or generally. A frequent metonymy becomes so on account of being an especially useful—that is to say, economical-efficient and/or effect-rich—way of picking out a certain referent. It picks out its target referent so successfully because the metonymically-used expression refers to a highly relevant, uniquely identifying property of that target referent. It is therefore plausible that, across the type or class that a given referent belongs to, the same properties will be useful identifying factors for all tokens of the type or all members of the class. For example, the referent in the ‘*ham sandwich*’ case belongs to the class of restaurant customers; for all members of that class, their food order is (in the restaurant context) a useful identifying property.

This suggests a generalisation: a ‘class-level’ identifying factor can be used to metonymically refer to any member of the class in question (e.g. ‘food ordered’ can be used to refer to any restaurant customer). Extracting, and storing in memory, patterns of this kind is useful to us in both metonymy production and comprehension. In production, it provides us with a set of ‘templates’ we can make use of on any occasion when we need to find a more efficient way of referring to a particular entity: if we can identify that entity as belonging to a specific class to which one of our stored patterns applies, all we need to do to extend the pattern to the case at hand is to find the contextually appropriate token of the class-level identifying factor and use as our referring expression the expression for which this token is the linguistically-specified referent. We are thereby likely to have met our goal of referring more efficiently, because the pattern we have applied has been extracted and stored due to its usefulness in reference-making. This could account for the productivity of metonymic patterns (see e.g. Pustejovsky (1995) on ‘metonymic polysemy’). In comprehension, metonymic patterns reduce processing effort by constraining our interpretive hypotheses. If we are able to identify the linguistically-specified referent of the metonymically-used referring term as belonging to a particular category of class-level identifying factors, we can access the relevant pattern and use it to identify the class to which the target referent belongs; we need then only home in on a specific class-member in order to recover the speaker’s intended referent. For instance, when a waitress utters (14), her colleague must identify the referring expression ‘*falafel wrap*’ as an instance of the identifying-factor type ‘food ordered’, but she can then retrieve from memory the pattern ‘food ordered for food orderer’, and this will constrain her search for the speaker’s intended referent to just the class of food orderers.

(14) The *falafel wrap* wants to know if the coffee is organic.

This is an example of a very localised pattern which is only useful, and therefore worthwhile extracting and storing, for a restricted group of language users, i.e. restaurant staff. However, there are also generalisations which are less-context specific and common to a greater number of language users, and indeed appear to be so broadly useful as to recur cross-linguistically, such as ‘container for contents’ as in (15), ‘producer for product’ as in (16) and ‘place for people’ as in (17) (Pustejovsky, 1995; Srinivasan & Rabagliati, 2015).

(15) The *kettle* is boiling. (= the water in the kettle)

(16) The guests admired his new *Picasso*. (= painting by Picasso)

(17) *France* has a healthier relationship with food than *America*. (= the French, the Americans)

Yet it is plausible that we can treat all these generalisations as arising from the routinization of metonymy interpretation, driven by the pressure to reduce processing costs in online interpretation which ultimately stems from the human tendency to maximise relevance when processing stimuli. Unlike on the cognitive linguistics approach, which sees such patterns as

somehow ‘inbuilt’, we are able to see them as emergent and, moreover, motivated. We can thus account for the crucial role of associative relations in metonymy interpretation in an explanatorily adequate manner, rather than by arbitrarily listing contiguity-based metonymic ‘mappings’, and can integrate our account into a properly inferential approach to referential metonymy.

8 Effects from frequency and routinization

Given that the frequency of a metonymic usage can lead to the routinization of its interpretation, we might also ask whether high frequency and routinization affect the storage and processing of a metonymic interpretation. This is an important question, as the answer could threaten the conclusions I have drawn concerning the nature of metonymy.

One type of answer may be that frequency and routinization create pressure to encode the metonymic interpretation of an expression as a separate lexical entry, meaning that at least certain metonymic interpretations are stored like the distinct meanings of a homonym (a case of two (or more) unrelated lexical items which happen to share a phonological form, e.g. ‘bank₁’ = side of a river, ‘bank₂’ = financial institution). For example, for the word ‘*Dickens*’, we may end up with two lexical entries, ‘*Dickens*₁’ = the author and ‘*Dickens*₂’ = the author’s works. This would threaten my claim that metonymy is not a variety of neologism, as it would suggest that in using an expression metonymically, a new standing meaning comes to be introduced into the language, albeit one which shares a linguistic form with the old (literal) standing meaning.

The hypothesis that frequent, routinized metonymic interpretations end up stored as distinct lexical entries makes several predictions, firstly that metonymic interpretations of this type will be processed differently to other metonymic interpretations. Instead of the metonymic interpretation being derived online, when the metonymic expression is encountered, representations of both its literal and its figurative referent would be initially activated before contextual information rapidly resolves the ambiguity, with the inappropriate interpretation being quickly suppressed, just as in homonymy processing (e.g. Swinney, 1979). Furthermore, we would expect to find relative frequency acting as a key determinant of the order in which the potential interpretations of the metonymic expression are accessed (Frisson & Pickering, 2007, p. 152).

However, empirical evidence suggests that these predictions do not hold. Frisson and Pickering (1999) conducted an eye-tracking experiment with a self-paced reading task, which focussed on the processing of sentences containing (i) expressions with a familiar ‘place for institution’ metonymic interpretation, e.g. ‘*convent*’ in (18); and (ii) expressions with a familiar ‘place for event’ metonymic interpretation, e.g. ‘*Vietnam*’ in (19):

- (18) That blasphemous woman had to answer to the *convent*.
- (19) A lot of Americans protested during *Vietnam*.

Their results revealed that people experience no difficulty with reading sentences like (18), where the context supports an interpretation of the sentence in which ‘*convent*’ receives its metonymic ‘institution’ reading, compared to sentences like (20), where the context supports an interpretation of the sentence in which ‘*convent*’ receives its literal ‘place’ reading:

- (20) These two businessmen tried to purchase the *convent*.

Processing times in the two conditions, literal and figurative, were equally fast, both during initial reading and later processing (as indicated by measures such as total fixation time for a word). Crucially, Frisson and Pickering (1999, p. 1376) discovered that the relative frequencies of the interpretations of a metonymic word did not affect processing times: participants were no quicker when reading a more frequent interpretation (whether literal or figurative) than when reading a less frequent interpretation. This suggests that the interpretations of a familiar metonymic item are not activated in order of frequency. By contrast, in homonymy processing relative frequency has a major effect on the order in which meanings are accessed. In addition, the ‘basic’ interpretation of a metonymic word, operationally defined by Frisson and Pickering (1999, p. 1373) as the interpretation listed first in dictionary definitions (for ‘*convent*’, the ‘place’ interpretation), was not found to be read faster than the derived interpretation (for ‘*convent*’, the ‘institution’ interpretation). This suggests that in processing frequent, routinized metonymies, language users do not initially access the ‘basic’ interpretation of the metonymic expression before subsequently accessing the derived interpretation. Frisson and Pickering (1999) thus concluded that there was no evidence for any ordering of the interpretations of a familiar metonymic item, whether in terms of frequency or of ‘basicness’. The data therefore indicate that metonymic expressions do not have the same processing profile as homonyms, which suggests that familiar, routinized metonymic interpretations are not stored in the same way as the distinct meanings of a homonym.

Frisson and Pickering (1999; see also Pickering & Frisson, 2001) argue that their results support an ‘underspecification’ model of processing for polysemous words, i.e. words that, like ‘*convent*’ and ‘*Vietnam*’, which have both a literal and a derived metonymic reading, are ambiguous between multiple different but related interpretations. The key claims of the model are that a polysemous word has only one encoded meaning which is an abstraction over the features of specific familiar interpretations, and that this single meaning is initially activated when the word is encountered. Note, however, that as Frisson (2009) highlights, there is a lack of clarity in the literature on ‘underspecified’ meanings as to whether the lexical semantic representation of a polysemous word should be seen as rich (as in Pustejovsky’s (1995) Generative Lexicon framework) or impoverished, possibly providing little more than pointers to conceptual space (e.g. Winograd, 1983; Carston, 2002; but see Carston, 2016, p. 157-8, and forthcoming, for a critique of the ‘impoverished’ position).

The underspecified meaning is taken to be compatible with all established interpretations of the polysemous word, making no distinction between primary and derived interpretations. For example, the meaning of ‘*Vietnam*’ which is initially accessed in (19) above would be underspecified with respect to whether it refers to the country or, metonymically, to the war that took place there. There is no immediate activation of any one fully specified referent. Contrast the processing of homonymous words, where all the different interpretations are initially activated. In addition, due to the fact that an underspecified meaning is the same for all related interpretations of a word, the model predicts that there should be no competition for activation between alternative interpretations, which again differs from many accounts of homonym disambiguation, where competition between meanings is taken to be the cause of processing difficulty (Rayner & Duffy, 1986). It thus appears from Frisson and Pickering’s (1999) evidence that, although a familiar, routinized metonymic interpretation of an expression is not encoded as a distinct entry in the mental lexicon, the standing meaning of the expression may nevertheless change in order to cover the established metonymic reading—yet at this stage, we cannot say more about the nature of such a change, e.g. whether the standing meaning becomes richer or leaner.

Problematically for my claim that referential metonymy is a distinct kind of lexical innovation, this change could be taken to indicate that at least certain, highly frequent and routinized metonymies are in fact examples of neologism as on Wilson and Falkum’s (2015)

account, involving the introduction of a new standing meaning, i.e. the underspecified meaning, into the language. However, the putative change can plausibly be seen as an 'adaptation' of the existing standing meaning of a metonymically-used word to cover its established interpretations—which seems fundamentally different in nature to the creation of an entirely novel standing meaning as in cases of neologism like compounds and denominal verbs, where the lexical entry for the new linguistic expression is presumably stored separately from the lexical entry for the word(s) from which it is derived. If there is indeed a change to the encoded meaning of a metonymically-used word, as Frisson and Pickering's (1999) evidence suggests, we could argue that while the resulting underspecified meaning would be undeniably novel, the change would not cause the lexical inventory to expand through the inclusion of a new standing meaning, and we could therefore conclude that no 'new coinage' takes place. The empirical data thus would not threaten my argument that referential metonymy is not reducible to neologism. Nevertheless, there is a certain tension between the notion of a change to the encoded meaning of a metonymically-used referring expression and my claim that a metonymically-used word has its old standing meaning and only its referent changes. In addition, given that the underspecified meaning for expressions with a familiar metonymic interpretation would be compatible with both the literal and metonymic readings of the expression, there is a sense in which such expressions could no longer be seen as 'metonymic', because they would no longer serve to pick out a new referent that is not covered by the linguistically-specified meaning of the expression.

However, let us consider why the postulated change to the encoded meaning of a referring expression with a highly frequent, routinized metonymic interpretation might occur. It may plausibly arise as a way to reduce processing costs, in response to usage of the expression. The resulting underspecified meaning would comprise features of, and thus be compatible with, the metonymic interpretation, which would facilitate the inferential move from encoded meaning to metonymic interpretation because on encounter with the expression, the metonymic interpretation would immediately be made highly accessible as a plausible hypothesis about the speaker's intended meaning and, according to the RT comprehension procedure which claims that interpretive hypotheses are tested in order of accessibility, may be the first hypothesis tested, should additional contextual information make this hypothesis more accessible than any other anticipated interpretations. The RT comprehension procedure further claims that our search for the speaker's intended meaning stops when we arrive at an interpretation which satisfies our expectations of relevance. Given that a frequent, routinized metonymic interpretation comes to be frequent and routinized on account of its relevance and usefulness, it is likely that on the occasion in question, the metonymic interpretation will indeed meet our expectations of relevance and will therefore be the interpretation we accept as the one most plausibly intended by the speaker. In this way, the underspecified nature of the encoded meaning of the referring expression would mean that when the metonymic interpretation is the interpretation required by considerations of relevance, we would be able to arrive at this reading quickly, with a minimum of processing effort. It is therefore possible not only to account for why the interpretation of referential metonymy may undergo routinization, as in §7, but also to propose a satisfactory explanation, which is compatible with the RT cognitive principle of relevance, of why the encoded meaning of a referring expression with a highly frequent, routinized metonymic interpretation may become underspecified.

Indeed, we could think of a diachronic progression for a metonymic reading: from an increase in frequency, depending on how relevant it is, to a routinization of its interpretation, and finally to a change to the encoded meaning of the metonymically-used expression. This may help to resolve the tension noted above, as it suggests that if the standing meaning of a metonymically-used expression changes, the change would be gradual, a function of the frequency with which the expression is used metonymically rather than an immediate and

necessary consequence of metonymic use. It also implies that such a change may not affect all metonymically-used words, only those whose metonymic usage attains an especially high frequency. We can thus acknowledge that repeated metonymic use of a referring expression may lead to diachronic changes to the standing meaning of that expression. However, we can also maintain that, for any given occasion when we use an expression to refer metonymically, although the expression comes to pick out a novel referent, its standing meaning does not change in the process of reference resolution.

9 Conclusion

Through close examination of examples of referential metonymy, comparison with other phenomena of lexical ‘flexibility’, namely modulation and neologism, and evaluation of empirical data, I have been able to elucidate the nature of referential metonymy and to show that referential metonymy can be accounted for in RT-compatible terms:

- Referential metonymy is a distinct phenomenon of lexical innovation involving the use of an existing expression to pick out a new referent.
- It is motivated by the communicative need to find a way of referring to a target entity which will enable the audience to recover the intended entity with minimum effort and/or maximum cognitive effects.
- The metonymically-used referring term is chosen because, in the communicative context, it introduces a property/feature of the target referent which is uniquely identifying and/or which conveys additional intended implications. It is therefore the best evidence of the speaker’s intended interpretation.
- The interpretation of referential metonymy is fully inferential.
- In interpretation, the metonymically-used referring expression does not come to express an ad-hoc concept. Referential metonymy therefore cannot be seen as a variety of lexical ‘modulation’.
- Further, referential metonymy does not involve ‘coining’ a new word, as in neologism.
- The interpretation of highly frequent metonymic usages may become routinized through the extraction of general metonymic patterns.
- Frequency of use and routinization of interpretation may lead to the encoded meaning of a metonymically-used referring expression changing to become underspecified as to whether the expression gets its literal or its metonymic interpretation.

Previous RT accounts of referential metonymy (e.g. Papafragou, 1996; Jiang, 2013; Rebollar, 2015) have argued for a fully inferential treatment of the phenomenon. The novel contributions here have been to closely analyse metonymy interpretation, thereby robustly demonstrating that referential metonymy is not reducible to ‘modulation’; to give an explanatorily adequate account of why we find metonymic patterns; and to offer a careful critique of the claim that metonymy can be treated as neologism, uniting theoretical arguments with experimental data from psycholinguistics in support of my position that referential metonymy is a distinct type of lexical innovation.

References

- Allott, N. (2013) Relevance Theory. In Capone, A., Lo Piparo, F., & Carapezza, M. (Eds.) *Perspectives on Linguistic Pragmatics*. New York: Springer. 57-98.
- Armstrong, J. (2016) The problem of lexical innovation. *Linguistics and Philosophy* 39 (2): 87-118.
- Carston, R. (1997) Enrichment and loosening: complementary processes in deriving the proposition expressed? *Linguistische Berichte* 8, Special Issue on Pragmatics: 103-127.
- Carston, R. (2002) *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.
- Carston, R. (2004) Relevance Theory and the saying/implicating distinction. In Horn, L., & Ward, G. (Eds.) *The Handbook of Pragmatics*. Oxford: Blackwell. 633-656.
- Carston, R. (2016) The heterogeneity of procedural meaning. *Lingua* 175-176: 154-166
- Carston, R. (forthcoming) Polysemy, pragmatics and lexicon(s).
- Donnellan, K. S. (1966) Reference and definite descriptions. *Philosophical Review* 77: 281-304.
- Falkum, I., Recasens, M., & Clark, E. V. (2016) 'The moustache sits down first': On the acquisition of metonymy. *Journal of Child Language* 43. Doi: 10.1017/S0305000915000720.
- Frisson, S. (2009) Semantic underspecification in language processing. *Language and Linguistics Compass* 3 (1): 111-127.
- Frisson, S., & Pickering, M. J. (1999) The processing of metonymy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory and Cognition* 25: 1366-1383.
- Frisson, S., & Pickering, M. J. (2007) The processing of familiar and novel senses of a word: Why reading Dickens is easy but reading Needham can be hard. *Language and Cognitive Processes* 22(4): 595-613.
- Gerrig, R. J. (1989) The time course of sense creation. *Memory and Cognition* 17(2): 194-207.
- Jiang, X. (2013) Referential metonymy: Reference transfer and pragmatic motivations. *UCL Working Papers in Linguistics* 25: 1-19.
- Just, M. A., & Carpenter, P. A. (1980) A theory of reading: From eye fixations to comprehension. *Psychological Review* 87: 329-354.
- Kaplan, D. (1989) Demonstratives. In Almog, J., Perry, J., & Wettstein, H. K. (Eds.) *Themes from Kaplan*. New York: Oxford University Press. 481-563.
- Kövecses, Z., & Radden, G. (1998) Metonymy: developing a cognitive linguistic view. *Cognitive Linguistics* 9(1): 37-77.
- Lakoff, G. (1987) *Women, fire and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.
- Langacker, R. (1987) *Foundations of cognitive grammar*. Stanford: Stanford University Press.
- Panther, K-U., & Thornburg, L. (2003) *Metonymy and Pragmatic Inferencing*. Amsterdam/Philadelphia: John Benjamins.
- Papafragou, A. (1996) On metonymy. *Lingua* 99: 169-195.
- Pickering, M. J., & Frisson, S. (2001) The semantic processing of verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory and Cognition* 27: 556-573.
- Pustejovsky, J. (1995) *The generative lexicon*. Cambridge, MA.: MIT Press.
- Radden, G., & Kövecses, Z. (1999) Toward a theory of metonymy. In Panther, K-U., & Radden, G. (Eds.) *Metonymy in Language and Thought*. Amsterdam: John Benjamins. 17-60.
- Rayner, K., & Duffy, S.A. (1986) Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition* 14: 191-201.
- Rebollar, B. E. (2015) A relevance-theoretic perspective on metonymy. *Social and Behavioral Sciences* 173: 191-198.
- Recanati, F. (2004) *Literal Meaning*. Cambridge: Cambridge University Press.
- Sperber, D., & Wilson, D. (1986/1995) *Relevance: Communication and cognition*. Oxford: Blackwell.
- Sperber, D., & Wilson, D. (1998) The mapping between the mental and the public lexicon. In Carruthers, P., & Boucher, J. (Eds.) *Language and Thought: Interdisciplinary Themes*. Cambridge: Cambridge University Press. 184-200.
- Sperber, D., & Wilson, D. (2002) Pragmatics, modularity and mind-reading. *Mind & Language* 17: 3-23.
- Srinivasan, M., & Rabagliati, H. (2015) How concepts and conventions structure the lexicon: Cross-linguistic evidence from polysemy. *Lingua* 157: 124-152.
- Swinney, D. A. (1979) Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior* 18: 645-659.
- Tyler, L. K., & Marslen-Wilson, W. D. (1977) The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior* 16: 683-692.
- Wilson, D., & Carston, R. (2007) A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts." In Burton-Roberts, N. (Ed.) *Pragmatics*. London: Palgrave. 230-259.

-
- Wilson, D., & Falkum, I. (2015) *Explaining metonymy*. Talk presented at the conference on New Developments in Linguistic Pragmatics, Lodz, Poland, April 2015.
- Wilson, D., & Sperber, D. (2002) Truthfulness and relevance. *Mind* 111: 583-632.
- Winograd, T. (1983) *Language as a Cognitive Process. Volume i: Syntax*. Reading, MA.: Addison-Wesley

Embedded pragmatic effects – what’s the problem?*

Robyn Carston

Abstract

It has been claimed that pragmatic effects that arise in embedded clauses pose a problem for the Gricean reasoning procedure. I maintain, however, that the real issue these phenomena raise for Grice, as he himself acknowledged, is their violation of his saying/implicating distinction. While these effects can be accounted for by Gricean reasoning, as is clearly demonstrated by Simons (in press), there is no way round this latter problem other than a major revision of Grice’s notion of ‘saying’ and hence of the saying/implicating distinction.

Keywords: local pragmatics, pragmatic enrichment, scalar implicature, Gricean reasoning, saying vs implicating, relevance theory

1 Introduction: local pragmatics and Grice

It is widely assumed that what are known as ‘embedded pragmatic effects’ present a problem for Grice’s programme of conversational logic. In a recent paper, Simons (in press) addresses this issue and shows that these local pragmatic effects are, in fact, explainable using standard Gricean reasoning. More specifically, she provides rational reconstructions of the inferences required to account for elements of pragmatic meaning such as those given in italics in the following examples:

- (1) A: What will you do for your mother’s birthday?
B: Either I’ll buy flowers or I’ll cook a nice dinner.
>> ‘buy flowers’ understood to mean: *buy flowers and give them to my mother*;
‘cook a nice dinner’ understood to mean: *cook a nice dinner to share with my mother*
- (2) A: What’s making noise up in the attic?
B: I’m not sure, but if there’s a nest up there, we’re going to have a heck of a mess to clean up.
>> ‘a nest’ understood to mean: *a nest occupied by birds*

The key point about these components of pragmatic meaning is that they arise in embedded clauses, that is clauses which fall in the scope of logical operators, disjunction in (1) and the conditional in (2); similar examples could involve negation or a propositional attitude verb (e.g. ‘John thinks there’s a nest up there’ in response to A’s question in (2)).

Had the embedded clause been uttered alone, the pragmatically inferred meaning would be described as a Gricean conversational implicature:

- (3) B: I’ll cook a nice dinner.
Implicature: The nice dinner I’ll cook will be to share with my mother.

* Many thanks to Emma Borg and Deirdre Wilson for helpful comments on an earlier draft of this paper. A slightly revised version is to appear in *Inquiry* as one of three commentaries on the paper ‘Local pragmatics in a Gricean framework’, by Mandy Simons, cited here as Simons (in press).

- (4) B: There's probably a nest up there.
Implicature: The nest is occupied by birds.

Refreshingly, all of Simons' key examples are cases of what Grice would have explained as relevance-based implicatures rather than the quantity or scalar implicatures, which have come to dominate the literature; indeed, these examples are very similar to Grice's example 'There's a garage round the corner' (Grice 1975/89: 32)¹, where the speaker, specifically observing the maxim 'Be relevant', conversationally implicates that the garage is open (and selling petrol). The sameness of the pragmatic contents derived for the embedded and unembedded cases has led to quite widespread talk of 'embedded implicatures' for cases like (1) and (2). This locution is, strictly speaking, a contradiction in terms, given Grice's clear characterisation of implicatures as non-truth-conditional components of utterance meaning, hence as not embeddable within the truth-conditional content ('what is said') of an utterance. Simons is faithful to Grice in this regard and does not employ this phrase but talks throughout of 'embedded pragmatic effects' or 'local pragmatic effects'.

In my view, Simons realises her goal of providing Gricean rational reconstructions of how these components of meaning could be recovered by hearers/interpreters. Following her catch phrase 'think globally, act locally', she provides convincing accounts of (1) and (2) above which employ global Gricean reasoning, based on an assessment of the compositional semantic content of the sentences uttered, in deriving the local (embedded) pragmatic effects. Thus, I am not taking issue here with the central point of her paper, but instead set out to assess the whole endeavour, suggesting it be extended and developed in certain ways and thereby made a bit more ambitious and (in my view) more interesting.

First, I'll raise some questions about the way in which the problem of 'embedded pragmatic effects' is construed within Grice's programme, based on his own expressed worries about this very phenomenon. I am not convinced that the real issue here is whether or not these pragmatic effects can be explained by the Gricean 'working out schema', as it seems to me that any and all pragmatic phenomena must be so explainable; rather, the issue lies with Grice's saying/implicating distinction, in particular his notion of 'what is said' and its essentially semantic nature. Second, I consider Simons' points that, given her successful rational reconstructions of how the components of pragmatic meaning at issue can be derived, there is both a local and a global view of the resulting contents, that is, they may either be taken to contribute to 'what is said' by the speaker (the proposition she has expressed, the truth-conditional content of her utterance) or be construed as conversational implicatures, and that it doesn't make much difference one way or the other. I think it does make a difference and the difference matters, and will suggest that the embedded (and other local) effects are components of a properly pragmatic notion of 'what is said', not of implicated propositions. Third, I maintain that it is time to move beyond normative rational reconstructions of the logic of speaker-hearer interactions, towards a more empirically cognitively-based account of communication and the processes of pragmatic interpretation. There is a lot more to 'local pragmatics' than 'embedded pragmatic effects' and there are several post-Gricean theories on the market nowadays which are pursuing accounts of these pragmatic phenomena which, arguably, move us closer to the cognitive mechanisms actually employed in utterance comprehension.

¹ Throughout this paper, all pages references to work by Grice refer to the versions (re)printed in Grice (1989).

2 The issue for the Gricean programme

To put my view simply and baldly: the issue that embedded pragmatic effects raise for the Gricean model is not to do with their reconstructibility by a reasoned argument of the sort Grice provides in his ‘working out schema’, but concerns how they can be accommodated by his particular view of the saying/implicating distinction, in particular, his concern to keep ‘what is said’ as close as possible to the conventional linguistic meaning of the sentence uttered. That is what I will try to argue for in what follows, first by looking at Grice’s own statements about these embedded cases, second by briefly examining the way in which the Gricean project was temporarily derailed by formal treatments of the phenomenon of scalar implicatures, and third by unpacking his problematic concept of ‘what is said’, which, as it stands, is not easily reconciled with embedded pragmatic effects, as he acknowledged.

2.1 Grice’s worry: the scope of conversational implicata

The problem that the original Cohen (1971) examples made for Grice concerned his truth-functional semantics for natural language connectives such as ‘and’ and ‘if’. Cohen maintained that Grice’s ‘conversationalist hypothesis’ regarding the frequent non-truth-functional meaning of ‘and’ had to give way to Cohen’s ‘semantical hypothesis’, because the sequential and consequential relations between the first and second conjuncts in examples like (5a) and (5b) clearly affect the truth-conditional content of the conditionals and thus the truth value of the statements made by asserting them:

- (5) a. If the old king died of a heart attack and a republic was declared, Tom will be happy.
b. If a republic was declared and the old king died of a heart attack, Tom will be happy.

Cohen makes no mention of there being any problem concerning the possibility of giving a Gricean rational reconstruction of the enriched content. His argument against the conversationalist hypothesis turns entirely on the impossibility of maintaining a truth-functional semantics for both ‘if’ and ‘and’ if implicated meaning is allowed to enter into the truth evaluation of (5a) and (5b), in that one or other (or both) of them would no longer be operating solely on the truth values of their clausal subparts.

As far as I am aware, Grice never attempted any head-on rebuttal of Cohen on this point and had, in fact, already expressed his own disquiet about a parallel problem arising for denials of some ordinary language conditionals, such as the example in (6), where it seems that, given the standard truth-functional account of ‘if’ (as equivalent to the material conditional), what the speaker denies is an implicature of the utterance of the unnegated conditional:

- (6) It is not the case that if X is given penicillin, he will get better.

Commenting on this sort of example, Grice said: ‘... if the affirmation of “if p, q” carries an implicature, its denial has to be interpreted as the denial of the implicature. This principle does not appear to be acceptable.’ (Grice 1967/89: 83). Later in his ‘Retrospective Epilogue’ (1987/89), when he looks back on his account of the ‘vulgar logical connectives’, he does not seem to have got much further with finding an answer to the vexing question ‘whether it is or is not required that a nonconventional implicature should always possess maximal scope’. He allows that there are some cases where negation has to be taken as a denial of the implicature of the embedded sentence/utterance, specifically, cases of metaphor where the literal content

is absurd (e.g. ‘She is not the cream in my coffee’), but ends by saying: ‘But where the limits of a license may lie which allows us to relate embedding operators to the standard implicata rather than to conventional meanings, I have to admit that I do not know.’ (Grice 1987/89, p. 375).

Both Grice and Cohen envisaged only two possibilities for explaining the enriched (non-truth-functional) meaning of the connection between conjuncts or between antecedent and consequent of a conditional: either it was a matter of conversational implicature or a matter of the (lexical) semantics of the connectives ‘and’ and ‘if’ (whether an ambiguity analysis, with the truth-functional meaning being just one of the conventional senses, or a single rich meaning, with contextual cancellation of certain of its components, as Cohen advocated). Neither Grice nor Cohen considered the possibility of a third kind of account, viz. a *local pragmatic enrichment* of the connection between propositional subparts of the utterance, while maintaining the truth-functional semantics of the connectives themselves. This was simply inconceivable, given the way in which the semantics/pragmatics (= saying/ implicating) distinction was construed at the time. Even now, after various reconstructions of the distinction, including differentiating ‘what is said’ from sentence semantics, the phenomenon at issue is still widely talked of negatively, as ‘pragmatic intrusion’ into what is said (the truth-conditional content of the utterance).

What I want to emphasise at this point is that there is no evidence that Grice’s worry about cases of implicature apparently falling in the scope of logical operators was a worry about their derivability, their susceptibility to being explained by a line of rational reconstructive reasoning involving his conversational maxims. In fact, if we take seriously Grice’s emphasis on the distinction between conventional and non-conventional components of utterance meaning, there is no obvious reason not to extend his general pattern of reasoning to other nonconventional (pragmatic) components of utterance meaning, including cases that clearly contribute to ‘what is said’, such as ascertaining the referents of demonstratives and indexicals, selecting the sense of ambiguous words, and (perhaps) completing apparent cases of nonpropositional sentence content (e.g. for utterances of ‘He is ready’, ‘She’s too tall’, etc.). It was pointed out long ago that Grice’s maxims apply to these aspects of speaker meaning (e.g. by Katz (1972, p. 449) and Walker (1975, pp. 156-7)), and, more recently, by Stalnaker (1989, p. 9): ‘the Gricean principles and maxims clearly play a role in resolving ambiguity and fixing contextual parameters as well as in generating conversational implicatures’. Some current Gricean pragmaticists provide explicit rational reconstructions of such phenomena; for instance, Geurts and Rubio-Fernandez (2015, p. 2) set out the following line of Gricean reconstructive reasoning for the disambiguation of the word ‘chestnut’ in the comprehension of an utterance of the sentence ‘The chestnuts are shedding their leaves’: ‘(a) She has used the word “chestnut”; (b) there is no reason to suppose that she is not observing the Cooperative Principle; (c) she could not be doing this unless she intended to refer to trees of the genus *Castanea*, for this is one of the standard meanings of the word, and it fits our discourse purpose better than any of the others; (d) she knows (and knows that I know that she knows) that I can see that the supposition that she intends the word in this meaning is required; (e) she has done nothing to stop me thinking that this is so; (f) she intends me to think, or is at least willing to allow me to think, that she intended to refer to this type of tree; (f) and so this is what she has implied.’ Now, clearly, the nature of the input to the reasoning (the premise in (a)) and the nature of the output (a word meaning) differ from the input (what is said) and the output (a conversational implicature) of Grice’s schema, but the actual pattern of reasoning is virtually identical (see Grice 1975/89: 31).

Some neo-Griceans (e.g. Simons (in press)) might object to this sort of account of disambiguation as not properly Gricean because it doesn’t start from the global consideration of the cooperativity of a complete speech act (what is said). Along with Geurts and Rubio-

Fernandez, I don't see that this makes the actual line of reasoning non-Gricean and, in any case, something of the required sort can be pretty easily supplied as the initial input to the rational reconstruction, given that this need not bear any close relation to the actual psychological processes of addressees: '(a) The speaker has said one of the following three things: (i) that the chestnut trees are shedding their leaves, (ii) that the nuts from the chestnut trees are shedding their leaves, (iii) that the chestnut horses are shedding their leaves; (b) there is no reason to suppose that she is not observing the Cooperative Principle; and so on (as above).' Furthermore, although Simons (in press) emphatically maintains that her Gricean accounts of the embedded pragmatic effects in (1) and (2) above are achieved without what she calls 'embedded pragmatic computation', it is not so clear that this really is the case. According to her definition, an embedded pragmatic computation involves general pragmatic principles operating on 'the content of an unasserted embedded clause; or the observation that the speaker has expressed this content ...'. In her account of the pragmatic enrichment of 'buy flowers' to *buy flowers to give to mother for her birthday* in (1), the disjunct 'I'll buy flowers' is found to be insufficiently relevant (not adequate as a potential answer to the question asked) and the pragmatic enrichment is performed so as to make it sufficiently relevant.² Her idea is that what kicks off the Gricean reasoning is the recognition that the compositionally derived content of the disjunction as a whole is not sufficiently relevant and (then) the hearer recognises that the problem lies with the first disjunct (and, presumably, also the second disjunct). However, nothing precludes it being the case that it is the recognition of the 'failure of relevance of the first disjunct', as she puts it, that triggers the Gricean reasoning in the first place.

Be that as it may, I see no problem at all with the application of pragmatic principles, Gricean or otherwise, to unasserted utterance parts, in the interests of recovering the intended interpretation of the utterance as a whole (as suggested way back by Walker 1975). The overriding point here is that it is inherent to pragmatic phenomena (as opposed to those that are a matter of linguistic convention or encoding) that they are derived inferentially and that, 'even if they can be intuitively grasped, unless the intuition is replaceable by an argument, the meaning will not count as pragmatic; it will be conventional/encoded' – here, I am echoing Grice's own statement about 'conversational' implicature versus 'conventional' implicature.³ So why has the idea caught on that there is a major problem with providing Gricean reconstructions of embedded or local pragmatic content? The answer, I think, lies, in part at least, with the way in which accounts of a particular kind of pragmatic inference, known as 'scalar implicature', have evolved.

² In Simons (2010), she says quite explicitly that what she is doing is departing from Grice in that, rather than seeing implicatures as arising from observations about what the speaker has said, 'we should instead see conversational implicatures as inferences generated by observations about what a speaker has (merely) expressed: this includes observations about sentence parts whose content is not part of what is asserted ...' (ibid: 163). The sentence parts at issue in that paper are exactly the same as those at issue in Simons (in press), i.e. disjuncts and antecedents of conditionals. Perhaps what Simons takes herself to be doing in this later paper is going back on her departure from Grice in the 2010 paper, having found a way to be more faithfully Gricean, but she doesn't say this explicitly (in fact, she doesn't refer at all to the 2010 paper, although it seems to be very much a precursor to the current one) and, as far as I can see, pragmatic principles (specifically, the maxim of relevance) are being applied to unasserted sentence parts here too.

³ 'The presence of a conversational implicature must be capable of being worked out; for even if it can in fact be intuitively grasped, unless the intuition is replaceable by an argument, the implicature (if present at all) will not count as a conversational implicature; it will be a conventional implicature' (Grice 1975/89: 31).

2.2 Scalar implicatures and formal rules of derivation

Much has been made in the vast literature on ‘scalar implicature’ of the difficulties of giving a Gricean account of the derivation of this kind of generalised quantity-based pragmatic inference when it occurs in embedded clauses. Simons mentions the example in (7) and another much discussed sort of example is that given in (8); in both cases, the desired pragmatic interpretation is the one given after >>, and the claim has been that Gricean reasoning does not deliver it, but rather only the too weak inference in (7) and the too strong inference in (8), as indicated after ##:

- (7) Prof Smith believes that some of her students are in her office.
 >> Prof Smith believes that some but not all of her students are in her office.
 ## It’s not the case that Prof Smith believes that all of her students are in her office.
- (8) Either Kai ate the broccoli or he ate some of the peas.
 >> Either Kai ate the broccoli or he ate some but not all of the peas.
 ## It’s not the case that either Kai ate the broccoli or he ate all of the peas.
 Hence: Kai didn’t eat the broccoli and he didn’t eat all of the peas.

What seems to have happened is that the phenomenon of quantity/scalar implicatures got reduced to a few cases of scalar terms (‘some’, ‘often’, ‘or’, and the number terms) which occupy a position on an entailment scale (e.g. <some, most, all>) and for which a simple automatic derivational procedure was proposed. Grounded in Grice’s first maxim of Quantity, but making no explicit reference to it, the procedure is as follows: ‘For every α higher on the entailment scale in which the uttered scalar term β occurs, infer, as an implicature, that the speaker believes/knows that the α counterpart of the uttered sentence is false.’ So, for the sentence/utterance ‘Some of the kids are sick’, the procedure generates the implicature ‘The speaker knows/believes that it is not the case that all of the kids are sick’ and then, assuming speaker reliability, the standard scalar implicature follows: ‘Not all of the kids are sick’. The instigator of this approach was Gazdar (1979: 57-61), but it has been taken up in variant forms by many others (e.g. Gamut 1991; Levinson 1983, 2000).

It turns out that this fixed procedure often doesn’t work when the scalar term (e.g. ‘some’, ‘often’, ‘four’, ‘or’) occurs in an embedded clause, as in (7) and (8) above.⁴ But this, I would maintain, is a problem with the rigid scalar inference procedure and *not* with the Gricean reasoning schema, which provides a very general and flexible method for calculating conversational implicatures of all sorts (and, as suggested above, for calculating other kinds of pragmatic phenomena too). This has now been demonstrated by, among others, Geurts (2010), who provides Gricean style inferential derivations for the seemingly problematic cases, in which the first maxim of Quantity plays a central role.

I don’t mean to underestimate the complexity of figuring out the key premises needed in providing a rational reconstruction of the pragmatic inferences involved in arriving at the intuitively correct interpretation of cases like (7) and (8). This is indeed tricky, as the reconstruction requires reasoning about the relevant stronger statements the speaker *could* have

⁴ It should be noted that Gazdar (1979) explicitly excluded occurrences of scalar terms in the scope of logical operators from his procedure, but, as others subsequently pointed out, such cases do give rise to implicatures of a scalar sort, just not the ones that the procedure would generate (see Hirschberg 1985, Recanati 2003, Sauerland 2004 on problems with extending Gazdar’s procedure to cover these cases). There are other issues with Gazdar’s procedure and it has been much tinkered with over the years, but the details don’t matter here.

made and assessments of her epistemic competence with regard to the facts of the situation under discussion. For instance, the reasoning that leads to the interpretation of (8) as Kai either ate the broccoli or some but not all of the peas involves consideration of four relevant alternative stronger utterances and of what the speaker can be reasonably supposed to believe about their truth/falsity given the utterance she chose to produce (see Geurts 2010: 166-68).⁵ However, the initial recognition that some further component of speaker meaning is required to preserve the presumption that the speaker is being cooperative (which Simons calls the step of *Gricean reasoning*) is entirely standard, and it is at the second stage (which she calls *the interpretive step*) that the theorist (the rational reconstructor) needs to exercise some ingenuity in providing the right premises.

My main point here is that it is the lengthy debate occasioned by these cases and their assumed automatized derivational procedure⁶ which has promoted the perception that there is something amiss with the Gricean mechanism when it comes to providing a reasoned argument for the presence of what seems to be embedded pragmatic material. But this is wrong: the apparent problem arose because the *formal procedure*, which generated the right implicature for utterances of stand-alone clauses containing scalar terms, generated an array of erroneous results when that clause occurs embedded in the scope of an operator. This is simply not a problem if we drop the procedure and go back to Grice's informal and very general reasoning schema. Furthermore, problems posed by a formal procedure of the sort set up for quantity-based scalar implicatures do not arise for the cases discussed in detail by Simons: she focuses on relevance-based pragmatic inferences and they are 'particularized' (context-specific) cases, so there is no temptation to see them as default inferences or routinized in any way. In his book in which he makes the case for default-rule accounts of a number of implicatures, in particular the scalar cases, Levinson (2000) does not try to extend this to such relevance-based cases, but instead invokes a very general 'Enrichment Rule' for hearers to follow: 'Amplify the informational content of the speaker's utterance by finding the most specific interpretation, up to what you judge to be the speaker's m-intended point, ...' (ibid: 114). Although this seems right, it is no more than a general guideline and Simons does us a service by providing explicit step-by-step demonstrations of how these context-specific amplifications can be derived by Gricean style reasoning employing the CP and maxim of relation, but, as far as I can see, there was never any reason to doubt the feasibility of this.⁷

Finally, returning to the passages where Grice discusses his worry about whether an embedding operator may sometimes have to be taken as governing a nonconventional (conversational) implicature (1967/89, p. 83, and 1987/89, p. 375), it's worth noting again that

⁵ Note that the result of Geurts' account is the *implicature* 'Kai didn't eat all of the peas'. Putting that together with 'what is said', the overall interpretation can be paraphrased as 'Kai either ate the broccoli or he ate some but not all of the peas'. His account, although very Gricean as compared with the grammatical/lexical accounts he is rejecting, still seems somewhat in the grip of the Gazdarian formal procedure approach. In this regard, I think Simons' (in press) account of disjunctions is better, in that it enables quantity-based reasoning on the second disjunct resulting in the (intuitively correct) *embedded enrichment* 'some but not all of the peas'.

⁶ There was a downward spiral, starting with the semi-formal procedure introduced by Gazdar (1979), to the default implicature account of Levinson (2000), where the key inferences were seen as generated by defeasible rules attached to specific lexical items, including 'some', 'often', 'or' (so inevitably derived and accepted unless contextually cancelled), through to the full-blown grammatical accounts of Chierchia (2004) and others.

⁷ Simons claims that relevance theorists 'are of the opinion that embedded pragmatic effects pose a challenge to the Gricean model of pragmatics. They take the Gricean model to be unable to account for the effects seen in Cohen's examples and others [with embedded pragmatic effects],' (p.2). As a relevance theorist, my issue is not with the ability of the Gricean reasoning schema to provide an account of how these effects might be derived, but rather with how they can be accommodated in Grice's 'model' given his particular construal of the saying/implicating distinction (Carston 2002, 2004; Carston & Hall 2012).

he maintained that for denials of metaphorical cases it is the implicature that is denied, not the absurd literal content:

- (9) Rex: She is not the cream in my coffee.
[Rex is denying that the female referred to is his pride and joy.]
- (10) Alf: I spent the summer cleaning the Augean stables.
Babs: No, you didn't. (You spent two months lazing at the seaside.)
[Babs is denying that Alf spent his summer working on a difficult and long overdue task.]

The embedding operator here is negation but the point would presumably carry over to the conditional too:

- (11) If she's the cream in his coffee, he's bound to give in to her demands.
If Mary can locate those old photos for him, she'll be the cream in his coffee.

At no point is it an issue for Grice whether or not the 'embedded implicature' is calculable by use of his standard reasoning schema; he simply assumes it is. What concerns him is the question of how to delimit the range of cases for which we must allow embedding operators to apply to 'the standard implicata rather than to conventional meaning'. I doubt that the Gricean framework can provide an answer to this question, given its very general (relatively unconstrained) reconstructive reasoning schema, on the one hand, and its overly constrained notion of 'what is said', on the other hand. An answer may, however, be found once we move to a more cognitively oriented account of pragmatic interpretation, as I suggest in section 4 below.

2.3 'Globalism', what is said, and the move to 'making as if to say'

Having shown that the pragmatic enrichments of 'buy flowers', 'cook a nice dinner' and 'nest', which occur within sub-clauses of the sentences uttered in (1) and (2) above, can be accounted for by pretty standard Gricean normative reasoning, Simons (in press) points out that the resulting propositional content can be construed either as what the speaker said (so the enrichment is truly local) or as a conversational implicature independent of what is said (hence as a wholly global inference). She doubts that there is any empirical evidence that could decide between the two construals and thinks that the preference is theory-driven, essentially dependent on one's construal of 'what is said'. Even if this is right, still one theory might be better than another and, on that basis, provide a reason for adopting one construal rather than the other. As discussed below in section 4, I believe that we can and should be moving beyond philosophers' rational reconstructions to accounts of pragmatics that come closer to the cognitive goals/constraints that direct actual processes of interpretation. This is what the relevance-theoretic account sets out to achieve and its cognitively constrained pragmatic processes deliver a localist implementation of the pragmatic effects under discussion.

However, staying for the moment within the Gricean framework, let's consider a line of defence that Simons adopts on behalf of the 'staunch globalist', who eschews 'pragmatic intrusion' into truth-conditional content. She suggests the availability of a move that I think really ought to be avoided. The globalist position would have it that B's utterance in (2), repeated here as (12), results in the contents as given in (13):

- (12) A: What's making noise up in the attic?
 B: I'm not sure, but if there's a nest up there, we're going to have a big mess to clean up.
- (13) What is said: If there's a nest up there, we're going to have a big mess to clean up.
 What is implicated: If there's a nest occupied by birds up there we're going to have a big mess to clean up.⁸

The semantically generated 'what is said' seems not to fall within the speaker's meaning (it's not the case that just any old nest, e.g. a disused one, will produce a big mess), so Simons' suggested move for the globalist here is to 'avail herself of Grice's notion of *making as if to say*' and to proceed from there to use standard Gricean reasoning to derive the conversational implicature, which constitutes the speaker's meaning. This does not seem to me to be a legitimate Gricean move. As noted earlier, this example, like all of Simons' cases, is an instance of particularised relevance-based Gricean reasoning and, along with Grice's example 'There's a garage round the corner', it falls into his Group A, 'Examples in which no maxim is violated, or at least it is not clear that any maxim is violated'. These are distinguished from his Group C, 'Examples that involve exploitation, that is, a procedure by which a maxim is flouted for the purpose of getting in a conversational implicature by means of something of the nature of a figure of speech.' (Grice 1975/89, pp. 32-33) It is examples in this latter group, which includes cases of metaphor, hyperbole, meiosis, and irony, that prompt the move to 'making as if to say' and all involve a blatant overt violation of a maxim. This is not what is going on in Simons' example; although the literal linguistic meaning here happens to be false, this is just a function of the particular embedding operator and not a matter of a speaker blatantly flouting truthfulness or any other maxim.⁹ Had B in (12) responded to A with a different embedding operator, e.g. 'Either there's a nest up there or some squirrels have moved in', or with no embedding, e.g. 'Oh, there's a nest up there', the literal linguistic meaning would be entailed by the (alleged) implicature, hence part of what is speaker meant, and there would be no talk of 'making as if to say' or maxim flouting. This would be unsatisfactory in a different way, in that the 'said and meant' content is essentially an idle wheel as it is subsumed and expanded upon in what is (allegedly) implicated. Even if this latter concern may be deemed an innocuous redundancy, the bigger point here can't be so easily dismissed. There are no principled (or empirical) grounds for splitting this set of cases, all of which, as Simons herself assumes, involve relevance-based enrichment, into two quite different kinds: the 'making as if to say'/maxim-flouting cases and the 'genuine saying'/maxim-observing cases. If this is the only way to salvage globalism, then it looks as if localism is the preferable option.

To recap, the problem Simons' examples present for the Gricean framework is not, as she convincingly shows, their calculability, but how they are to be accommodated by his saying/implicating distinction, specifically, his conception of what is said. I've argued at considerable length elsewhere that Grice packs two incompatible constraints into this single concept: (a) the (pragmatic) requirement to be speaker meant, and (b) the (semantic) requirement to be as close as possible to the conventional compositional meaning of the sentence (Carston 2002, 2004; Carston & Hall 2012). Different theorists deal with this problem

⁸ The referents of 'we' and 'there' would have been provided, but this doesn't affect the point I want to make here about the relationship between the said and the implicated.

⁹ The one example Grice gives in Group C which involves a 'real, as distinct from apparent, violation of the maxim of Relation', something he considers quite rare, is the bland statement by B about the weather which follows A's statement 'Mrs X is an old bag' (Grice 1975/89: 35). This is a blatant refusal by B to make what he says relevant to A's preceding remark. Nothing like this is going on in Simons' examples.

in different ways, but what they tend to have in common is the separating out of these two requirements into different levels of analysis, that is, the semantic content of a sentence (even relative to a context) is one kind of thing, while the propositional content that is meant (communicatively intended) by the speaker is another. It is clearly the latter that is Simons' focus here and the issues now are: (a) whether or not there is a distinction to be made, within speaker meaning, between globally derived conversational implicatures and a propositional content which is, in effect, the outcome of local pragmatic enrichment of the conventional/encoded meaning of the sentence uttered; and, if there is, (b) whether there is any reason to prefer one or the other of these as the locus of her cases of relevance-based pragmatic enrichment.

In the next section, I argue that there is a great deal more to *local* pragmatic effects than *embedded* pragmatic effects, in that the concepts (or senses) that speakers communicate with the words they utter are usually at least partially pragmatically derived. I believe that this goes quite a long way toward providing Simons with the kind of evidence she feels is lacking for choosing between the localist (what is said) construal of her pragmatic enrichment cases and the globalist (implicature) construal.

3 Local pragmatics, the dynamic lexicon and 'what is said'

Simons focuses on 'embedded' pragmatic effects, that is, components of pragmatically derived meaning which arise within clauses/propositions that fall in the scope of logical operators. In this section, I start by looking at the local pragmatics of utterances of simple sentences (no embedded clauses or logical operators). This goes well beyond the standard cases of ascertaining the intended referent of indexicals and the intended sense of ambiguous lexical/phrasal forms (e.g. 'bank', 'in the grip of a vice') and is essential to many (perhaps most) instances of grasping the meaning a speaker intends in uttering a word or phrase. The strategy is to first demonstrate the prevalence of local pragmatic adjustments in figuring out what the speaker said and, given this, to argue that there is no reason not to allow the exact same kinds of local adjustments in figuring out what the speaker said in uttering a more complex sentence containing an embedded clause.

Let's consider some cases of utterances of simple sentences. First, some examples adapted from Nunberg (1979):

- (14) a. He can hit the ball two *football fields*.
 b. He hit a home run two *games* ago.
 c. *IBM* went up three points last week.

What is expressed by this use of 'football field' is 'length of a football field'; 'game' is used to express a measure of time; 'IBM' is used to refer to 'IBM stocks'. Given their linguistic contexts here, there is no difficulty in seeing how these meanings can be pragmatically derived from the more well-established meaning of the word or phrase in each case (see Geurts 2010: 184, for a step-by-step Gricean derivation of the occasion-specific sense of 'football field' in (14a)). Next, consider some cases of novel 'denominal verbs', as discussed by Clark and Clark (1979):

- (15) a. The boy *porched* the newspaper.
 b. She *wristed* the ball over the net.
 c. He *houdinied* his way out of the closet.
 d. The police car *sirened* up to the accident site.

- e. Mum will *jam-spoon* you when she gets home.

Having run a series of experiments on people's ability to understand these, Clark and Clark (1979, p. 767) concluded: 'people readily create and understand denominal verbs they have never heard before'¹⁰ (and no doubt the same goes for deverbal nouns, e.g. 'buy some nice *eats*', 'an *embed* in Iraq', 'a good *get*', etc.). The examples in (14) and (15) are all cases of verbal innovation (novel uses of existing words in (14) and new, but motivated, word coinages in (15)), whose meaning is inferred pragmatically, using as key evidence the existing known senses of the words on which they are based. The new communicated senses cannot but contribute to what the speaker said (the truth-conditional content of her utterance) or else she would have to be deemed, highly implausibly, to have not said anything at all, nor even made as if to say something, given that there is nothing propositional without this contribution. Whatever implicatures these utterances may have depend on there being some fully propositional component (what was said) on the basis of which, together with other premises, including the presumption that Gricean maxims are being observed, they can be inferred.

According to the relevance-theoretic approach to 'lexical pragmatics', interpreting a word (or phrase) frequently involves the construction of an ad hoc concept or occasion-specific sense, based on an interaction between linguistically encoded concepts, contextual information and pragmatic principles (see Wilson and Carston (2007)). The constructed sense may be a narrowing or a broadening of the encoded sense, or a combination of both. Here is a possible example of each:

- (16) a. I'm not *drinking* tonight.
 b. If the kids go on this way, we'll be *bankrupted*.
 c. Either she's a *saint* or he's a lot nicer to her at home.

A likely interpretation of 'drinking' in (16a) is a narrowing to 'drinking alcohol' and this may be further modulated on the basis of contextual knowledge to 'not drinking much alcohol' or 'not drinking spirits', etc. Taking (16b) as uttered by an exasperated parent whose children are constantly demanding new clothes and gadgets, she most likely doesn't intend 'bankrupted' strictly literally but is expressing a broader concept along the lines of 'made considerably poorer'. And (16c), a kind of metaphorical use, probably communicates a concept which is both broader and narrower than the literal meaning.¹¹ The general point here is that our lexical resources do not consist of a static store of words and their meanings: we are constructing new senses (even new words) much of the time, some of which are one-off, others of which become routinised to varying degrees, and a few of which take on the status of established word senses.

Polysemy is the phenomenon of a word having a number of related senses, but unlike cases of arbitrary ambiguity, known as 'homonymy' (e.g. 'bank', 'coach'), polysemy is highly variable across individuals apparently speaking the 'same language'; what is a case of semantic polysemy for you may be monosemy for me, requiring me to pragmatically infer the other sense(s), and vice versa for other cases. For instance, as a result of frequent use of the

¹⁰ Of course, some need more context than others; (15e), which is my example, was used to express the proposition that Mum will beat the addressee with the (big wooden) spoon she uses for making jam when she gets home.

¹¹ Some of the earliest work on 'local' or 'lexical/phrasal' pragmatics was by Jerry Hobbs and colleagues (e.g. Hobbs & Martin 1987; for an overview, see Hobbs 2004). They focused, in particular, on cases of metonymy and noun-noun compounds, which are pretty much universally accepted as cases of meaning that contributes to the truth-conditional content of an utterance (what is said), thus as pragmatic contributions to (or 'intrusions' on) that content.

pragmatically derived sense, some English speakers may represent the words ‘drink’ or ‘saint’ as having two senses, while, for others who have encountered the derived use less often or not at all, the word may have only a single sense and the interpretation of the derived use would involve constructing (or reconstructing) an appropriate ad hoc concept. What this indicates, again, is that the lexical pragmatic processes responsible for concept construction are contributing to the truth-conditional content (what is said/asserted) of the utterance rather than merely affecting implicatures. It would be highly implausible to suppose that (16a) or (16c) would be understood as expressing or asserting different propositions depending on whether the intended sense of ‘drinking’ or ‘saint’ is recovered by disambiguation or by ad hoc concept construction. As we put it in Wilson and Carston (2007, p. 241): ‘One of the most important functions of pragmatic inference is to compensate for grammatical and lexical differences among members of a speech community, so that addressees with different encoded senses can end up with the same interpretations, albeit via different routes.’¹²

Note that each of the examples in (16) is an instance of an ‘embedded pragmatic effect’, in that each occurs within a clause that falls in the scope of a logical operator. Simons’ staunch globalist would presumably require that the intended sense in such cases occurs as an implicature and that each of these utterances is a case of ‘making as if to say’. But, following Grice, the globalist would take disambiguation (sense selection) as essential in reaching a full identification of what the speaker has said (‘drink-1’ or ‘drink-2’, ‘saint-1’ or ‘saint-2’). So sticking staunchly to globalism entails that for some addressees of (16a) the speaker has said that she is not drinking alcohol tonight, while for others she has made as if to say that she is not drinking anything at all tonight and implicated that she is not drinking alcohol. This is an absurd outcome and not one, I believe, that Simons would want to ascribe to.

Geurts (2010) makes similar points to mine about the pragmatic determining of speakers’ word meanings: ‘The dynamics of this process is Gricean: the logic that underwrites the selection and construction of word meanings is similar to that of conversational implicatures’ (ibid, p. 184). Cases of scalar inferences that he accommodates in this way include the following:

- (17) a. Around here, we don’t like coffee, we love it.
 b. I’d rather have a warm bath than a hot one.

Taken strictly literally, (17a) would be a contradiction but the lexical meaning of ‘like’ is pragmatically narrowed to ‘like but not love’ and the proposition expressed (what is said) is thus a contingent truth; essentially the same goes for ‘warm’ in (17b). Here, as he clearly argues, Gricean reasoning is employed in establishing the truth-conditional meaning of the utterance, or ‘what is said’ (acknowledging that Grice’s construal of ‘what is said’ has to be amended). He concludes: ‘interpretation processes are guided by the Cooperative Principle from the lexical level up’ (ibid: 85).¹³

¹² I am less comfortable now with this talk of words having ‘encoded senses’ and suspect that grasping the senses expressed by speakers is essentially always a matter of occasion-specific concept formation, but with the accessibility of those senses/concepts varying across individuals, a matter which is dependent on their particular communicative experiences. The nature of standing word meaning is currently a much debated issue; for an overview of the possibilities, see Recanati (2004).

¹³ In fact, pragmatic processes may well kick in at an even more local level if, as suggested by Lieber (2004, p. 179), some derivational morphemes are semantically underspecified and polysemous. They may even work with linguistic (or perceptual) properties of uttered words other than (and logically prior) to their meaning properties, e.g. phonological or phonetic properties; see Predelli (2010) on the comprehension of malapropisms (e.g. Mrs Malaprop understood as *saying* that X is a nice arrangement of epithets when she utters ‘That’s a nice derangement of epitaphs’).

Towards the end of her paper, Simons discusses the local pragmatic process of assigning a referent to an indexical, using the example of an utterance of ‘He is late’, for which ‘all that the semantics gives to the interpreter is an underspecified representation’. In such cases, ‘The interpreter need not wait to identify what is said before she can pose ... the Gricean question’. Her assumption seems to be that it is only when the meaning delivered linguistically falls short of determining a proposition that pragmatic inference (Gricean reasoning) comes into play: ‘In cases where the words uttered by the speaker do not determine a proposition, the reasoning begins from whatever semantic content or structure those words make available.’ On this basis, her examples (1) and (2) would not involve local pragmatic enrichments of ‘nest’, ‘buy flowers’, etc., because, arguably, a proposition is determined without the need for these pragmatic enrichments. The criterion seems to be: ‘if and only if nothing propositional is delivered by linguistic semantics, do whatever pragmatics is needed to make it propositional’. If this were applied to the lexical/phrasal cases discussed in this section, some of them would require local pragmatic work (e.g. the word coinages in (15)) and some would not (e.g. the meaning shifts in (14) and (16), although the proposition (allegedly) expressed/said in each of these cases is nonsensical), and this will vary from hearer to hearer depending on their lexicons, which, in turn, depend on their individual communicative histories. As far as I can see, there is nothing principled or warranted in this criterion and all of these should be treated in the same way, as instances in which pragmatic inference is essential in establishing the proposition the speaker said and meant (explicitly communicated).¹⁴

Simons’ examples in (1) and (2) are cases of lexical/phrasal narrowing: from ‘nest’ to ‘occupied nest’; from ‘buy flowers’ to ‘buy flowers to give to mother for her birthday’. Having given a Gricean rational reconstruction of how the enrichment can arise within a disjunct or the antecedent of a conditional, she leaves it open whether the resulting communicated proposition is to be construed as ‘what is said’ or ‘what is implicated’. As reviewed in this section, pragmatic inference plays an essential role in establishing the intended sense of words and phrases, hence in determining what the speaker has said. Putting this together with the absence of any motivation for an implicature construal of her cases, as discussed in the previous section, it looks as if the local ‘what is said’ construal of these cases is correct. In the next section, I outline a particular post-Gricean account of how such pragmatic enrichments can arise as components of the proposition said and meant by the speaker.

4 Going beyond Gricean rational reconstruction

The essence of the relevance-theoretic account of utterance interpretation is that hearers are entitled to expect a satisfactory array of contextual implications (and other cognitive effects) in return for the effort of attending to and processing an utterance addressed to them. The pragmatic theory is grounded in a general view of human cognition according to which our cognitive systems are geared toward maximising the cognitive effects they derive from

¹⁴ Simons’ position here is at least partly motivated as a response to the apparent problem of ‘Grice’s circle’ raised by Levinson (2000, p. 186): ‘what is said seems both to determine and to be determined by implicature’. However, this problem has been resolved in several post-Gricean frameworks without resorting to an arbitrary nonpropositionality constraint on the involvement of pragmatic processes in recovering the proposition said and meant. In the next section, I indicate how it is met in the relevance-theoretic framework (see also discussion in Carston (2002)). For other solutions to the circularity issue, see Recanati’s (2004) distinction between primary and secondary pragmatic processes, and the ‘critical pragmatics’ approach of Korta & Perry (2008), who distinguish ‘near-side pragmatics’ (i.e. contributions to ‘official referential content’ or what is said) from ‘far-side pragmatics’ (i.e. implicated contents).

incoming information while being as economical as possible in their expenditure of effort. Given the communicative intention with which utterances are produced, they are special among sources of information in that they convey a presumption of ‘optimal relevance’, that is, a presumption that the effort of processing them will be adequately rewarded by the cognitive effects they will yield.

This is not the place for a full account of the theory and the foregoing is just a brief and rudimentary characterisation (or reminder) of the framework, but I hope it suffices as an indication of its strong cognitive orientation.¹⁵ Commensurate with this orientation, the accounts of utterance interpretation it gives are not intended as idealised rational reconstructions of the reasoning by which utterance content could be derived, but as more closely meshed to the actual inferential processes employed by addressees and the ways in which these are constrained by cognitive factors such as different degrees of activation, hence of accessibility, of information in memory.¹⁶ With this in mind, consider how a relevance-based pragmatic account might explain A’s interpretation of B’s response to her question in (18):

- (18) A: What’s making noise up in the attic?
B: Oh, there’s a nest up there.

Given the expectation of (optimal) relevance that addressees of utterances are entitled to have, A can expect B to provide an answer to her question. But B’s answer here is indirect; she *implicates* that it is birds that are making the noise in the attic. What then is the proposition explicitly communicated (said and meant, in Grice’s terms)? Linguistic decoding of the utterance provides the atomic concept nest which has associated with it a cluster of general world knowledge about nests: that they are a construct, made by birds, of such and such size/shape/composition, that birds lay eggs in them and live in them, etc. Some aspects of this general knowledge about nests will be more activated than others; given the immediately preceding question about the source of noise, information about birds residing in nests and making certain kinds of noise is likely to be highly activated. Such readily accessible assumptions are employed as contextual premises, together with the proposition explicitly expressed, from which contextual implications are inferred, specifically, that it is birds that are making the noise in the attic. But this is not a warranted inference unless the general concept nest is enriched to something along the lines of nest occupied by birds. Thus, there is a process of ‘backwards inference’ from the relevance-based implication to the proposition expressed (the truth-conditional content of the utterance) which effects this enrichment. This is an instance of a general pragmatic mechanism of mutual parallel adjustment of (hypothesised) explicature and (hypothesised) implicatures of an utterance, a process which stops when a sound inference which meets the addressee’s expectation of relevance is reached. The interpretation of B’s utterance derived in this way is roughly the following:

¹⁵ For more detailed article-length accounts of the relevance-theoretic framework and its cognitive underpinnings, see Wilson & Sperber (2004) or Wearing (2014).

¹⁶ More generally, the approach heeds the findings of evolutionary psychology on the domain-specificity of our cognitive capacities and the bounded nature of our rational functioning (e.g. the use of ‘fast and frugal heuristics’, as discussed by Gigerenzer & Selten (2002)). Nonetheless, it probably qualifies as a ‘broadly Gricean’ account as characterized by Simons (in press): ‘any account of pragmatic inference which posits that an interpreter reasons about what the speaker meant and that this reasoning is guided by presumptions of rationality of the speaker is a Gricean account’.

- (19) Explicature: There is a NEST* up in the attic
 (where the ad hoc concept NEST* can be paraphrased as ‘nest occupied by birds’)
 Implicature: The noise in the attic is made by birds

The embedded cases that are Simons’ primary concern would work in much the same way, albeit with additional complications concerning the correct treatment of the logical operators involved. Take one of her disjunction cases:

- (20) A: What’s making noise up in the attic?
 B: Either there’s a nest up there or some squirrels have moved in.

The implicated answer to A’s question is that (B thinks that) the noise in the attic is being made by birds or by squirrels. As in Simons’ Gricean derivation, there is some pragmatic inferential work to be done on each of the disjuncts individually, given the relevance-based assumption that each is a potential answer to the question asked, although the speaker does not know which is the correct answer (as the use of the disjunction operator indicates). So the processing of the first disjunct proceeds as just discussed for (18), its relevance lying with the contextual implications it yields, in particular, the implication that nesting birds are a possible source of the noise in the attic. The processing of the second disjunct is similar, its relevance lying with the contextual implication that squirrels are a possible source of the noise in the attic. While the inferences hinging on the first disjunct require pragmatic enrichment of the nest concept in order to be properly warranted, there is not any obvious comparable concept enrichment in the second disjunct (although, arguably, ‘moved in’, is completed to ‘moved in to the attic’). On this account, the pragmatic enrichment of ‘nest’ is a component of the proposition explicitly communicated (what Simons nicely calls a ‘robust notion of what is said’), whether or not it occurs in a simple sentence or in a more complex structure in which that sentence is embedded. The same would go for all the various cases of lexical/phrasal pragmatics discussed in the previous section. As with Simons’ Gricean analyses, the enriching/adjusting of sentence subparts is ‘always in the service of rendering the content of the utterance as a whole pragmatically appropriate’ (Simons, in press), that is, the inferential processes involved are all in aid of recovering an optimally relevant interpretation of the utterance as a whole.

Recall again Grice’s worry: for certain uses of linguistic expressions in simple sentences, an implicature has to be recovered in order to grasp the speaker’s meaning (e.g. the metaphorical use of ‘cream in X’s coffee’; many uses of ‘and’-conjunctions), but when they occur in the same sentence embedded under a logical operator what was implicated content in the unembedded case is now a component of what is said (truth-conditional content). This strange inconsistency, which perplexed Grice, disappears on the account proposed here, as the pragmatic enrichment contributes to what is said and meant (truth-conditional content) whether the expression at issue occurs embedded or unembedded.

To conclude: Mandy Simons has set out impeccable Gricean accounts of how embedded pragmatic effects arise, so anyone who doubted the ability of the Gricean reasoning schema to do this should now be convinced that it can. In this response to her paper, I’ve suggested that, although this is indeed a job well done, the doubts about its feasibility were misplaced, largely arising from failed attempts to formalise the derivation of scalar implicatures. While Grice himself worried about cases where it seemed that ‘implicatures’ fall in the scope of logical operators, the basis of his worry, I believe, was the problem they posed for his saying/implicating (semantics/pragmatics) distinction. Finally, I have suggested that it is time to move on from purely rational reconstructions of pragmatic phenomena, whose aim is to

reveal the ‘logic’ of acts of linguistic communication, and to develop accounts that come closer to the cognitive reality of communication and interpretation.

References

- Carston, R. (2002). *Thoughts and utterances: the pragmatics of explicit communication*. Oxford: Blackwell.
- Carston, R. (2004). Relevance theory and the saying/implicating distinction. In L. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 633-656). Oxford: Blackwell.
- Carston, R. & Hall, A. (2012). Implicature and explicature. In H-J. Schmid (Ed.), *Cognitive pragmatics, Handbook of pragmatics* (Vol. 4, pp. 47-84). Berlin: de Gruyter Mouton.
- Chierchia, G. (2004). Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond: the cartography of syntactic structures* (Vol. 3, pp. 39-103). Oxford: Oxford University Press.
- Clark, E. & Clark, H. (1979). When nouns surface as verbs. *Language* 55, 767-811.
- Cohen, L.J. (1971). Some remarks on Grice’s views about the logical particles of natural language. In Y. Bar-Hillel (Ed.), *Pragmatics of natural language* (pp. 50-68). Dordrecht: Reidel.
- Gamut, L. (1991). *Logic, language, and meaning* (Vol. 1). Chicago, IL: University of Chicago Press.
- Gazdar, G. (1979). *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic Press.
- Geurts, B. (2010). *Quantity implicatures*. Cambridge: Cambridge University Press.
- Geurts, R. & Rubio-Fernandez, P. (2015). Pragmatics and processing. *Ratio* 28(4), 446-469.
- Gigerenzer, G. & Selten, R. (2002). *Bounded rationality*. Cambridge, MA: MIT Press.
- Grice, H. P. (1967/89). Indicative conditionals. Lecture 4 of ‘Logic and conversation’, the William James Lectures, Harvard University 1967. In H. P. Grice (1989), pp. 58-85.
- Grice, H. P. (1975/89). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts*, pp. 41-58. New York: Academic Press. Reprinted in H. P. Grice (1989), pp. 22-40.
- Grice, H. P. (1987/89). Retrospective epilogue. In H. P. Grice (1989), pp. 339-385.
- Grice, H.P. (1989). *Studies in the way of words*. Cambridge MA: Harvard University Press.
- Hirschberg, J. (1985). *A theory of scalar implicature*. PhD thesis, University of Pennsylvania, PA. Published 1991 by Garland New York.
- Hobbs, J. (2004). Abduction in natural language understanding. In L. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 724-741). Oxford: Blackwell.
- Hobbs, J. & Martin, P. (1987). Local pragmatics. *Proceedings: International joint conference on artificial intelligence* (pp. 520-523). Milano, Italy, August 1987.
- Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, S. (2000). *Presumptive meanings: The theory of generalized conversational implicatures*. Cambridge MA.: MIT Press.
- Katz, J.J. (1972). *Semantic theory*. New York: Harper and Row.
- Korta, K. & Perry, J. (2008). The pragmatic circle. *Synthese* 165(3), 347-357.
- Lieber, R. (2004). *Morphology and lexical semantics*. Cambridge: Cambridge University Press.
- Nunberg, G. (1979). The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy* 3, 143-184.
- Predelli, S. (2010). Malapropisms and the simple picture of communication. *Mind & Language* 25(3), 329-345.
- Recanati, F. (2003). Embedded implicatures. *Philosophical Perspectives* 17(1), 299-332.
- Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.
- Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27(3), 367-391.
- Simons, M. (2010). A Gricean view on intrusive implicatures. In K. Petrus (Ed.), *Meaning and analysis: New essays on H. Paul Grice* (pp. 138-169). London: Palgrave Macmillan.
- Simons, M. (in press). Local pragmatics in a Gricean framework. *Inquiry*
- Stalnaker, R. (1989). On Grandy on Grice on language. Unpublished paper, MIT. Abstract in *Journal of Philosophy* 86, 526-527.
- Walker, R. (1975). Conversational implicatures. In S. Blackburn (Ed.), *Meaning, reference and necessity* (pp. 133-181). Cambridge: Cambridge University Press.
- Wearing, C. (2014). Relevance theory: Pragmatics and cognition. *WIREs Cognitive Science*. doi: 10.1002/wcs.1331
- Wilson, D & Carston, R. (2007). A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.), *Pragmatics* (pp. 230-259). London: Palgrave Macmillan.
- Wilson, D. & Sperber, D. (2004). Relevance theory. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 607-32). Oxford: Blackwell.

The roles of lexical information and pragmatic processing in the interpretation of complement coercion structures*

Lewis Pollock

Abstract

Two theories of *complement coercion* (as in ‘Mary began the book’), termed the *lexico-semantic theory* and the *syntactico-pragmatic theory*, are examined. A new experiment motivated by two factors is reported. Firstly, it has been recently shown that previous studies of complement coercion overlook a confound in their stimuli that renders their findings open to question. Secondly, proponents of the lexico-semantic account assume that complement coercion structures have default interpretations, and that their theory explains how these default interpretations are generated. However, studies that support this assumption contain normed experimental materials that bias specific interpretations. The reaction time data of the new experiment returned null results, but the error data show that the assumption that complement coercion structures have default interpretations is unfounded. The syntactico-pragmatic theory is shown to be flawed on theoretical grounds. However, on the basis of the error data, the lexico-semantic theory is argued to be incomplete without acknowledging that pragmatic processing helps resolve the underspecification inherent in complement coercion.

Keywords: Complement coercion, lexical representations, pragmatic processing

1 Introduction

This paper evaluates two competing theories about how the process of complement coercion (e.g. ‘Mary began the book’) should be analysed and reports a psycholinguistic test of them using a new experimental paradigm based on comprehension-to-comprehension priming. The first of these theories will henceforth be referred to as the *lexico-semantic* theory. The second will henceforth be referred as the *syntactico-pragmatic* theory. Note that the labels of these two theories are intended to reflect the intentions of their proponents, but interpreting the claims made by either of these theoretical perspectives is not clear cut with respect to which systems govern the process under investigation. The structure of this paper is as follows. First, an overview of the lexico-semantic account is given. For this, a summary of Pustejovsky’s (1995) original account of enriched composition is required. Next, psycholinguistic research undertaken from the lexico-semantic perspective is reviewed. It is argued that the lexico-semantic claim that complement coercion structures generally have default interpretations is debateable because of how experimental materials in their studies have been normed. Then, the competing syntactico-pragmatic account described in de Almeida (2004) and Dwivedi and de Almeida (2008) is outlined. It is argued that the syntactico-pragmatic account is not clear with respect to how its components interact and so cannot function as a proper psycholinguistic theory. A recent experiment by Katsika, Braze, Aswini and Piñango (2012) identifies a serious confound in the experimental materials featured in previous complement coercion experiments. On the basis of Katsika et al.’s (2012) finding and the norming problem mentioned above, a new experiment was designed which investigated the extent to which different pairwise

* This work was supported by the Economic and Social Research Council [grant number ES/J500185/1]. I thank Robyn Carston and Andrea Santi for their help, expertise and encouragement.

combinations of structures exhibit priming in comprehension. This experiment is reported and its implications and limitations are discussed with respect to both theories. Finally, it is argued that complement coercion structures do not necessarily have default interpretations, and that the lexico-semantic account must ultimately rely on pragmatic theoretical constructs in order to account for how complement coercion structures are interpreted. Therefore, elements from both theories are required in order to account for the resolution of ambiguity in complement coercion structures.

2 The lexico-semantic account

2.1 Pustejovsky's (1995) account of complement coercion

There exist linguistic structures that are interpretable and yet seem to require processes beyond traditional compositional processes. Pustejovsky (1995) called this phenomenon *enriched composition*. The sentences in (1) below demonstrate *complement coercion*, which is a type of enriched composition that has received much attention in the psycholinguistic literature:

- (1) a. The girl began reading the book.
b. The girl began the book.

The meaning of Sentence (1a) can be derived through compositional language processing. However, Sentence (1b) appears to be different. In Sentence (1b), what it is that the girl is doing to the book is not explicitly expressed. An interpretation of (1b) is incomplete until this event is specified: one cannot simply begin a book; one must begin doing something *to* a book. Other things being equal, the interpretation of Sentence (1b) is likely to be the same as the interpretation of sentence (1a). As McElree, Traxler, Pickering, Seely and Jackendoff (2001, p. 18) point out, various features of the sentence and the discourse context in which the sentence appears can motivate different interpretations of these structures. Consider the sentences in (2):

- (2) a. The author began the book.
b. The author began writing the book.
c. The author began reading the book.

The interpretation of (2a) is more likely to be (2b) than (2c) unless the context provides enough information to make it clear that the author began reading the book as opposed to writing the book. Pustejovsky (1995) accounts for the potentially varying interpretations of complement coercion structures by positing a set of compositional processes that operate on *types* of semantic entities. Objects are one such type. The lexical representation of an object has a *qualia structure*. A qualia structure is part of the lexical representation of a word that contains information such as an object's physical components and the telic and agentive properties associated with the object that the word denotes. Sentence (2a) can be interpreted as (2b) in some contexts but as (2c) in other contexts because the telic representation of *book* contains two arguments bound to the verb *read*, whereas the agentive representation of *book* contains two arguments bound to the verb *write*. The subject of the verb also has a qualia structure. The information contained in the semantic representation of a verb and the qualia structures of a verb's arguments is combined via a process called *type-shifting* in order to generate an interpretation of sentences such as (1b) and (2a). Jackendoff (1997) notes that certain verbs select for a complement that denotes an event. The verb *begin* is one such example. When a verb such as *begin* appears with a complement that denotes an entity, such as *book*, a process

of type-shifting alters the semantic representation of the complement such that it conforms to the verb's selectional restrictions.

2.2 Psycholinguistic research on complement coercion from the lexico-semantic perspective

In the first psycholinguistic investigation of complement coercion, McElree et al. (2001) designed a self-paced reading time experiment in order to test whether there was behavioural evidence for Pustejovsky's enriched composition theory. McElree et al. (2001) report that reading times for the complement of the verb were longer for complement coercion structures than for traditional compositional structures. This finding is taken to be 'on-line evidence for a type of enriched *lexical* processing' (McElree et al. 2001, p. b17, emphasis added). However, at this stage McElree et al.'s (2001) evidence only shows that complement coercion structures take longer to process than control structures, and this evidence alone does not make it safe to assume anything about *why* complement coercion structures take longer to process than control structures, or which linguistic systems govern these processes.

Traxler, McElree, Williams and Pickering (2005) report four eye-tracking experiments designed to investigate the effects of context on processing complement coercion structures, and to test hypotheses as to which of a set of potential processes is responsible for their increased reading time. These processes are:

(A) When encountering [the complement noun], comprehenders access the word's lexical entry and attempt to integrate various stored senses of this word into the evolving semantic representation of the sentence.

(B) The mismatch between the verb's selectional restrictions and the stored senses of the noun triggers a coercion process.

(C) Comprehenders use salient properties associated with the complement noun and other relevant discourse elements (including but not necessarily limited to the agent phrase) to infer a plausible action that could be performed on the noun.

(D) Comprehenders incorporate the event sense into their semantic interpretation of the VP by reconfiguring the semantic representation of the complement, converting [began [the book]] into [began [reading the book]].

(Traxler et al., 2005, p. 4)

Traxler et al. discount process (A) on the grounds that it is obviously not unique to complement coercion and discount Process (B) in light of Pykkänen, Liinas, and McElree (2004), who show that complement coercion does not prompt the same MEG activity as semantic relation mismatches. This leaves processes (C) and (D), which are the main focus of Traxler et al.'s (2005) experiments. Experiment 1 aimed to test whether the increased processing cost of coercion should be attributed to process (C): is complement coercion costly because upon reading *The girl began the book*, a reader has to access world knowledge about books and representations of a speaker's communicative intentions in order to infer what the girl is doing to the book? This would not be a lexical process in the sense employed in McElree et al. (2001) because it relies on extra-lexical information that is not necessarily contained in the lexical entries *girl* and *book*. Rather, it is a pragmatic process that makes use of the human ability to operate on contextual information along with representations of a speaker's likely intentions in

order to derive an interpretation of an underspecified utterance. In experiment 1, participants read context sentences followed by target sentences while their eye movements were recorded. Traxler et al. (2005) hypothesise that if the cost of coercion structures lies in inferring implicit content, then a context sentence that makes this content explicit should reduce or eliminate the cost of a coercion structure in the target sentence. An example stimulus is reproduced below, where the underlined word in the first sentence constitutes the explicit introduction of the action that is left implicit in the complement coercion structure featured in the second sentence:

The contractor had been building in the suburbs. That spring, he began a condominium next to the shopping center.

Traxler et al. (2005, p. 7) find that explicit contextual information does not reduce the cost of complement coercion, and conclude that 'these results are difficult to reconcile with accounts that would attribute the cost to the time or effort needed to retrieve a possible activity associated with the NP complement'. Instead, they propose that the cost should be attributed to process (D): '[incorporation of] the event sense into the semantic interpretation of the VP by reconfiguring the semantic representation of the complement' (Traxler et al. 2005, p. 4), which is the core element of the lexico-semantic position. In addition to this empirical evidence in favour of their theory, Traxler et al. (2005, p. 18) state that 'a primary reason for assuming lexical representations of the sort proposed by Pustejovsky (1995) is that they account for the default interpretations of otherwise underspecified expressions'. There are two issues with this claim. Firstly, the sentences that appear in Traxler et al. (2005) have first been subjected to a fill-in-the-blank norming process. Participants were asked to fill in the blank for sentences such as 'The author began ____ the book', presented in isolation. The most frequent response to this task was taken and introduced in the context sentences in Traxler et al.'s (2005) experiments. Other complement coercion structures that are more ambiguous, or harder to interpret without a supporting context, were not featured in their study. It is possible that if structures with a greater degree of ambiguity were featured, there would be a greater effect of context on reading times, and Traxler et al. (2005, p. 18) themselves note that 'there may be expressions requiring coercion in which the recovery of a suitable action does have measurable consequences, perhaps because it may require a relatively slow inferential process'.

Secondly, even after this norming process, Traxler et al. (2005, p. 6) report that in their acceptability norming pre-test, participants only identified the 'correct' interpretation of complement coercion structures 67% of the time on average. Far from supporting the notion of default interpretations of complement coercion structures, this agreement rate suggests that in many cases, participants generate different interpretations of structures even when they have been designed to make one specific interpretation likely. Moreover, empirical studies to date have only investigated how long it takes participants to finish reading complement coercion structures, but have not examined whether participants tend to find them acceptable or not. If isolated complement coercion structures show a high rejection rate relative to straightforwardly compositional control sentences, then this is evidence against the notion that complement coercion structures have default interpretations and suggests a flaw in the lexico-semantic account. This issue is investigated in the experiment reported below.

3 The syntactico-pragmatic account

Other researchers have claimed that pragmatic processing must take place in order for complement coercion structures to be interpreted. De Almeida (2004) and de Almeida and Dwivedi (2008) argue against the lexico-semantic account of complement coercion discussed

above in three ways. First, de Almeida (2004) reports an experiment that does not replicate McElree et al.'s (2001) finding. Second, de Almeida (2004) and de Almeida and Dwivedi (2008) disagree with the lexico-semantic theorists over how some of their psycholinguistic data should be interpreted. Third, de Almeida and Dwivedi (2008) advance a theoretical argument against the lexico-semantic theory that denies that anything like complement coercion actually takes place when sentences such as (1b) or (2a) are interpreted. These arguments will now be examined in turn.

3.1 Psycholinguistic research on complement coercion from the syntactico-pragmatic perspective

De Almeida (2004) reports a self-paced reading experiment in two parts that does not replicate McElree et al.'s (2001) finding that complement coercion structures take longer to process than control structures. As before, the experimental stimuli consisted of sentence triads that contained either a coercing structure, a preferred structure, or a non-preferred structure. De Almeida (2004, p. 255) finds no significant effect of sentence type on reading times in the critical positions in the stimuli sentences. De Almeida (2004, p. 256) argues that his data and 'some of Traxler et al.'s [2002] results... suggest that type-shifting verbs are not more complex – or do not engender more complex semantic processing – than other verb types'. However, de Almeida's interpretation of Traxler, Pickering and McElree's (2002) experiment 1 results is dubious. Traxler et al. (2002, p. 536) report an eye tracking experiment that recorded a participant's gaze as they read sentences that contained a complement coercion structure, a preferred verb or a non-preferred verb, and report the following measures: first-pass time, first-pass regressions, second-pass time, total time and regression-path time. De Almeida (2004, p. 256) states: 'In the eye-tracking study reported by Traxler et al.... type-shifting effects were only marginally significant and even non-significant, in some cases, such as in first-pass reading times' and takes this as strong evidence against the psycholinguistic reality of complement coercion. However, he ignores the other measures on which there was a significant difference between complement coercion sentences and preferred/non-preferred sentences such as second-pass time, which is a measure of how much time a reader spent re-reading a word after an initial fixation. Indeed, the pattern of results reported in Traxler et al. (2002) seems to indicate that although readers do not immediately encounter increased difficulty with complement coercion structures (based on no significant difference on the first-pass time measure), subsequently readers do have to revert back to critical areas (based on a significant difference on second-pass time in all areas and significantly more first-pass regressions from the complement to the verb in complement coercion structures). Contrary to de Almeida's (2004) interpretation, these findings are consistent with the claim that complement coercion structures are more costly to process than control structures. Furthermore, in a direct response to de Almeida (2004), Pickering, McElree, and Traxler (2005) report an additional eye-tracking experiment that featured de Almeida's (2004) stimuli and do find evidence of increased processing difficulty for complement coercion structures. They argue that, given that a coercion effect has been demonstrated repeatedly across multiple experiments in McElree et al. (2001), Traxler et al. (2002) and Pickering et al. (2005), and that it is possible to interpret some of de Almeida's own data as supporting a coercion effect, the null results that de Almeida (2004) reports are probably due to Type II Error caused by a relatively small number of experimental items.

3.2 A theoretical critique of the syntactico-pragmatic account

The theoretical element of the syntactico-pragmatic account will now be examined. De Almeida and Dwivedi (2008, p. 302) make a central claim that they argue is incompatible with the lexico-semantic account: 'the basis for the interpretation of indeterminate linguistic structures is inferential or pragmatic in nature', and they include complement coercion in the category of 'indeterminate structures'. This claim is incompatible with the lexico-semantic account because the lexico-semantic account holds that the basis for interpreting these structures is not pragmatic, but lexical: interpretation is based on lexical information at the word level, and the effect of context is relatively limited. Traxler et al. (2005, p. 20) state that 'context can ameliorate the cost associated with coercion', but only if 'it contains a relevant event sense. In this way, a discourse representation appears to serve as an extended lexicon'. The 'discourse representations' in their experiments are lexically realised verb phrases. Traxler et al. claim that contextual information does not help participants infer the meaning of complement coercion structures. Instead, the full event sense, containing the relevant verb and its complement must be lexically realised in discourse prior to the ambiguous structure in order for processing to be facilitated.

De Almeida and Dwivedi (2008) propose a new analysis of complement coercion structures based on the existence of covert syntactic structure. Sentence (3a) reproduces their formalisation of this analysis. Sentence (3b) contains a construction without the proposed syntactic modifications for comparison purposes.

- (3) a. The secretary began [VP [V0 e] [NP the memo]]
 b. The secretary [VP wrote [NP the memo]]

De Almeida and Dwivedi (2008, p. 313) argue that instead of arising from the application of a complex operation on highly structured lexical representations, it is 'the extra VP structure in [3a which] accounts for the extra time taken in some experiments to process such sentences'. The difference between sentence (3a) and sentence (3b) is that sentence (3a) contains an extra VP node that itself contains an empty verb head 'V0 e'. The content of this empty verb head is the specific action undertaken by the agent of the coercing verb (*began*), and is pragmatically inferred from the discourse context and other elements of the sentence. De Almeida and Dwivedi (2008) seem to claim that the process of pragmatic inference that supplies the content of the empty verb head does not prompt a measurable increase in processing time, and that it is the processing of additional syntactic structure alone that explains any such increase. On this account the process of pragmatic inference is 'free' in the sense that it leaves no psycholinguistic footprint. Putting aside initial doubts as to how it would be possible for such a process to happen instantaneously and at no cost, at a later stage in their argument this position is confused by another element of their account. De Almeida and Dwivedi (2008, p. 321) argue that verbs that can appear in complement coercion structures come with certain entailments that may give rise to indeterminacy. For example, usage of an aspectual verb such as *finish* presupposes that there was a beginning to the event that *finish* is to be applied to. Now consider sentence (4), which contains a complement coercion structure:

- (4) John enjoyed the book.

De Almeida and Dwivedi (2008, p. 321) propose that 'the indeterminacy of [4] can be characterised as presupposing a relation between an activity and two NP referents... for someone to utter [4], a relation z (the verb-referring event) between x and y needs to be determined such that x enjoyed z -ing y '. They call the content of this relationship z -information

and equate it with the content of the empty verb head that occupies the extra syntactic structure postulated in their account. They claim that McElree et al. (2001) detected increased reading times for complement coercion structures because in their experimental stimuli:

“a filler for the empty V is not supplied overtly and because the presuppositional context does not supply a potential filler either. Our suggestion is, then, that increased reading times at post-verbal positions obtained in some of the experiments discussed above could be due to the lack of specification of the z relation.”

(De Almeida & Dwivedi, 2008, p. 322)

In the McElree et al. (2001) experiment, participants encountered complement coercion structures with no supporting context. De Almeida and Dwivedi (2008) argue that participants may have experienced difficulty processing such sentences simply because not enough information was present in either the sentence itself or the discourse context in order for them to derive a full interpretation. However if it is possible for z-information to reduce the processing time of complement coercion structures, where does this leave de Almeida and Dwivedi's (2008, p. 131) earlier claim that it is the extra VP structure in complement coercion that accounts for their extra processing time? It might be argued that de Almeida and Dwivedi include the process of pragmatically inferring the content of the empty verb head under the umbrella of processing 'extra VP structure', but they give the impression that this is not what they intend:

By claiming that extra structure accounts for extra processing time, we rely on old yet standard assumptions about sentence processing (Frazier and Fodor 1978): structural computations correlate with processing time.

(De Almeida & Dwivedi, 2008, p. 313)

It is difficult to see how a pragmatic inference could be considered a 'structural computation'. On the one hand, de Almeida and Dwivedi (2008) claim that processing extra syntactic structure is the reason for increased comprehension time in complement coercion structures, but on the other hand, that information present at the discourse level required to fill an empty verb head can reduce this processing time. It is not clear whether they are arguing that increased processing time is due to the fact that an extra syntactic node is present, or whether it is due to the fact that the content of this node must be pragmatically specified. De Almeida and Dwivedi (2008) also introduce Frazier and Fodor's (1978) model as an example of what constitutes a 'structural computation'. This 'sausage machine' model is completely syntactic in nature and does not make any claims about how pragmatic information is processed. This further suggests that they do not mean to label a pragmatic process as part of a structural computation.

Ultimately, it is not clear how de Almeida and Dwivedi (2008) intend the syntactic and pragmatic elements of their account to interact, and as a result it is also not clear what psycholinguistic data would support their argument. However, suppose that de Almeida and Dwivedi (2008) commit to the claim that salient z-information-bearing content can reduce the processing times of complement coercion structures. De Almeida and Dwivedi's (2008) theory still faces a serious problem. They claim that the presence of a likely antecedent for the content of the empty verb head can speed up the processing of the extra syntactic structure. Recall the experiment reported in Traxler et al. (2005), discussed above. In this experiment, Traxler et al. investigated the effect of context on the processing of complement coercion structures and found that the presence of a likely antecedent for the content of the empty verb head did *not* have a measurable effect on processing time unless that antecedent was a fully realised

instantiation of the event sense of the complement coercion structure, which is inconsistent with de Almeida and Dwivedi's (2008) account.

4 Katsika et al. (2012) and a confound in studies of complement coercion

At first glance, Traxler et al.'s (2005) evidence provides a knock-down argument against de Almeida and Dwivedi's (2008) position. However, a recent experiment reported by Katsika et al. (2012) demonstrate convincingly that all previous psycholinguistic studies of complement coercion have overlooked a serious confound in their experiment materials that may have masked any effect of the presence of z-information in context on processing times of complement coercion structures. A full half of the stimuli used in Traxler et al.'s (2005) experiment 1 suffer from this confound. Katsika et al. (2012) argue that the set of verbs used in psycholinguistic experiments on complement coercion consists of at least two main verb subsets. The first such subset contains the aspectual verbs, such as *begin*, *finish*, *start*, and *continue*. The second subset contains verbs that Katsika et al. (2012) call 'psychological' verbs, such as *enjoy*, *resist*, *savour*, and *endure*. They show that aspectual predicates such as *begin* and *finish* do trigger a process of complement coercion when combined with an entity-denoting complement, and exhibit the accompanying increased processing difficulty. However, the psychological verbs do not. This is because although it is possible for the complement of a psychological verb to be event-denoting, this is not actually a selectional *requirement* of psychological verbs: although it is not possible to simply 'begin a book', it *is* possible to simply 'enjoy a book'. This finding has implications for the interpretation of Traxler et al.'s (2005) data on the effects of context on complement coercion processing. Half of the verbs featured in experiment 1 were psychological verbs that do not necessarily trigger a process of complement coercion when paired with an entity-denoting noun. De Almeida and Dwivedi could rightly argue that it would not be expected for z-information-specification to have any effect on the processing of psychological verbs because in actual fact these verb-complement combinations do not feature additional syntactic structure in the first place. A new study of comprehension-to-comprehension priming in complement coercion will now be reported. This study was designed to investigate the issues facing both the lexico-semantic theory and the syntactico-pragmatic theory, and features stimuli that do not suffer from the confound discovered by Katsika et al. (2012).

5 Experiment report

This experiment investigated the effect of different relationships between prime and target sentences on the time taken to comprehend the target sentences, where the target sentences all contained complement coercion structures that featured aspectual verbs. Participants were presented with a prime sentence, then a target sentence, and asked to determine whether there were any errors in *either* the prime or the target. Primes and targets were separated by a fixation cross. The time in between the onset of the target sentence and a participant's response was taken as a measure of how long it took that participant to process the target sentence. An example critical stimulus set is given in Table 1 below:

Table 1 - Example critical item set

Condition	Prime sentences	Target sentence
(i) baseline	The man bought the newspaper.	The girl began the book.
(ii) explicit action	The man was reading the newspaper.	The girl began the book.
(iii) different coerced	The man finished the sandwich.	The girl began the book.
(iv) same coerced	The man began the newspaper.	The girl began the book.

The properties of each experimental condition and how these properties bear on tests of both the lexico-semantic theory and the syntactico-pragmatic theory will now be outlined. In pair (i), the prime sentence features no complement coercion but the target sentence does feature complement coercion. This pair provides a measure of how long it takes to process a complement coercion structure in an environment that should have no effect on this processing time either way. In pair (ii), the action that is implicit in the target complement coercion sentence is explicitly introduced in the prime sentence. This pair is designed to test de Almeida and Dwivedi's (2008) claim that z-information specification should speed up complement coercion processing times. Because the activity that is implicit in the target sentence has just been introduced in the preceding prime, de Almeida and Dwivedi's (2008) prediction would be that reaction times to pair (ii) should be faster than reaction times to pair (i). Note that the prime sentences in pair (ii) do not contain the full event sense required in the interpretation of the complement coercion structure in the accompanying target sentence: the arguments of the verb are not repeated. In pair (iii), the only thing that is held constant across prime and target is the presence of complement coercion. No semantic element from the prime is repeated in the target. This pair is designed to test whether the act of coercing a complement primes another act of complement coercion. Neither theory offers a prediction about whether the process of complement coercion alone should prime a subsequent complement coercion process, and this effect has not been investigated in previous studies. In pair (iv), both prime and target sentences contain complement coercion, and the coercing structure is repeated across sentences. However, the arguments of the verb are not repeated. Traxler et al. (2005) report that complement coercion structures that contain the same verb and require the recovery of the same implicit event exhibit priming. A comparison of pair (iv) reaction times to pair (i) reaction times therefore tests whether this finding is replicable with a different methodology.

It is also possible to compare the acceptance rates of pair (ii) to the other conditions in order to investigate the extent to which complement coercion structures have default interpretations. If complement coercion structures have default interpretations as Traxler et al. (2005) maintain, then there should be no difference between the acceptance rates across critical conditions because participants should simply make use of the default interpretation in all cases. However if, as de Almeida and Dwivedi (2008, p. 322) suggest, isolated complement coercion structures are generally difficult to interpret and do not have default interpretations, then pair (ii) should show higher acceptance rates compared to the baseline condition. This is because in pair (ii), the prime sentence explicitly introduces the action that is left implicit in the target sentence, and participants should be able to make use of this information in order to

help them generate a complete interpretation of the target sentence. In the baseline condition, the prime sentence does not contain any information that could aid in the generation of an interpretation of the target sentence, and so acceptance rates should be lower if no default interpretation is available in some cases.

5.1 Methodology

Participants

32 adult monolingual native speakers of British English drawn from the University College London SONA subject pool with reported normal or corrected-to-normal vision participated in the study. Mean age = 29 (SD = 9.5, range =20-53), 11 males. Subjects gave informed consent and were paid £6 for their participation. In order to maximise the efficiency of data collection, the experiment reported below was run in a block with a separate, unconnected experiment. The order in which the experiments were presented was counterbalanced in order to minimise any potential ordering effects.

Procedure

Participants saw a prime sentence for 3 seconds, and then a fixation cross for 2 seconds. Then participants saw a target sentence. Participants were asked to judge as quickly as possible whether *either* the prime or target sentence contained a violation and respond via a button press. A violation could either be grammatical or semantic. Examples of all types of violation are given in the Materials section below. There were two dependent variables: the time in between the onset of the target sentence and the participant's response, and the rate of acceptance of target sentences. The experimental manipulation was the relationship between the prime and the target in terms of content. The time taken for a participant to respond is taken to be an index of how long it took the participant to fully comprehend the second sentence of the pair. In the critical trials, there were no errors in either prime or target, so the time taken to respond in these trials would not be affected by any error detection or resolution processes. The fact that participants were asked to detect errors in either the prime or the target ensured that participants fully comprehended both sentences. Participants were given a practice session lasting ten trials in order to familiarise them with the task. No data were gathered from the practice trials, and no verbs or nouns used in the practice trials appeared in the main experiment. Every type of error that appeared in the main experiment also appeared in the practice trials, and errors appeared in both prime and target sentences within filler trials. During the practice trials participants were given feedback after each judgement indicating whether they had responded correctly or incorrectly in order to maximise their understanding of the task requirements and to make sure they understood what constituted an error. During the main experiment, participants did not receive feedback. The experiment lasted approximately 15 minutes.

Materials

Sixteen critical target sentences containing a complement coercion structure were generated. These sentences were all of the form [the X verbed the Y] in the simple past tense, and featured an animate agentive subject performing an action on a theme or patient object. For each of these 16 target sentences, four prime sentences were generated. Participants were divided into four groups, and these 64 sentence pairs were divided into four lists such that each participant

only saw any sentence once. A single participant therefore experienced one experimental condition per target sentence and contributed four times to each condition.

Forty filler sentence pairs were generated. These filler sentences also all featured an animate agent performing an action on a patient or theme. The tense and voice of these filler sentences was variable. Twenty eight of the filler items (half of the total trials in a run) contained either a syntactic or semantic violation in either the prime or target sentence. A syntactic violation consisted of a verb subject agreement error or an incorrectly conjugated tense. A semantic violation consisted of an animacy violation in either the subject or object noun phrase head. Care was taken to ensure that violations appeared approximately equally in sentence initial, medial and final positions in order to ensure that participants were motivated to read and fully comprehend the entirety of every sentence. An example of each type of violation is given in Table 2 below:

Table 2 - Examples of filler violations

Violation	Example
Semantic animacy	The dress was ripped by the faith.
Grammatical tense	The policeman will finding the gun.
Grammatical agreement	The butler burn the toast.

A complete list of all the sentence pairs used in the experiment can be found in the Appendix.

5.2 Results

Reaction time data

All critical trials with reaction times above 3500ms were removed on the grounds that these responses were long enough that they were unlikely to reflect automatic sentence processing. Three participants were excluded from reaction time analysis because after removing all trials which were responded to incorrectly, all of the trials from one critical condition were lost. Because the raw reaction time scores were not normally distributed, and parametric statistical tests require normally distributed data, the natural logarithm of all data points was computed and statistical analyses were performed on this transformed dataset (Shapiro-Wilks test on all 4 critical conditions after transformation, all p-values >0.05). Data were analysed by condition and all critical trial data points fell within 2.5 standard deviations of the condition mean, therefore no outliers were removed. It may have been preferable to remove outliers by participant instead of by condition in order to account for variability in natural reading speeds. However, outliers were removed in this way instead of by participant because since a participant only contributed to a condition four times it is unlikely that four data points would provide an accurate measure of what constituted an outlier.

The reaction time data were subjected to a repeated measures one-way ANOVA with independent variable Prime Type (baseline vs same coerced vs different coerced vs explicit action) and dependent variable natural log reaction time. Table 3 below displays the mean natural log reaction times (with standard deviations in parentheses) of each critical condition.

Table 3 - Mean natural log reaction times by condition

Baseline (pair i)	Explicit action (pair ii)	Different Coerced (pair iii)	Same Coerced (pair iv)
7.478 (0.287)	7.408 (0.3)	7.408 (0.3004)	7.374 (0.3)

The ANOVA test revealed no main effect of Prime Type ($F(1, 28) = 0.354$, $p = 0.557$, partial eta squared = 0.12). Given that Traxler et al. (2005) found evidence that repeating a coercing structure sped up processing time, and that this result is not replicated here, the likely interpretation of this result is that the methodology employed in this study is simply not sensitive to these effects.

Acceptance rate data

In the reaction time data analysis, 26% of critical trials were removed because they were judged to contain violations. Recall that de Almeida and Dwivedi (2008) argue that in isolation, a complement coercion structure is underspecified in that a full interpretation is difficult to generate with the information at hand. It is possible that participants were not necessarily judging complement coercion structures as constituting an error, but rather they simply rejected these structures because a full interpretation was difficult or impossible to generate. A Pearson's Chi-Square test was performed in order to determine whether participants were more likely to accept pairs of sentences from one condition over pairs from other conditions. There was a significant difference between prime types in the proportion of trials accepted and rejected (Chi-Square = 33.188, $df = 3$, $p < 0.0001$, Cramer's $V = 0.255$, $p < 0.001$). Subsequent 2x2 table Chi-Square tests were conducted comparing the baseline condition directly to the other three conditions in order to further investigate this effect. Tables 4-6 below display the counts, expected counts and standardised residuals of each pairwise comparison along with the relevant critical statistics.

Table 4 – Chi-Square Test: Baseline vs Same Coerced

Chi-Square = 9.111, $p = 0.004$, Cramer's $V = 0.189$, $p = 0.003$, $df = 1$		Baseline	Same coerced
Accepted	Count	95	72
	Expected count	83.5	83.5
	Standardised residual	1.3	-1.3
Rejected	Count	33	56
	Expected count	44.5	44.5
	Standardised residual	-1.7	1.7

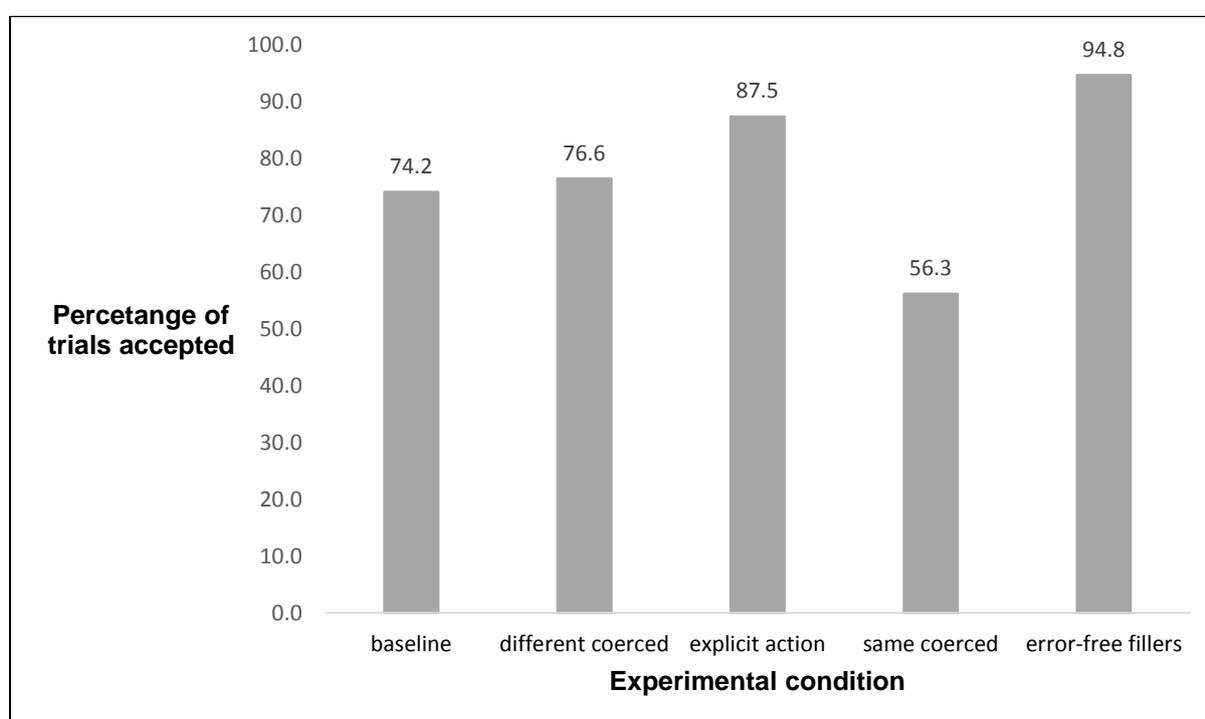
Table 5 – Chi-Square Test: Baseline vs Different Coerced

Chi-Square = 0.189, $p = 0.772$, Cramer's $V = 0.027$, $p = 0.663$, $df = 1$		Baseline	Different Coerced
Accepted	Count	95	98
	Expected count	96.5	96.5
	Standardised residual	-0.2	0.2
Rejected	Count	33	30
	Expected count	31.5	31.5
	Standardised residual	0.3	-0.3

Table 6 – Chi-Squared Test: Baseline vs Explicit Action

Chi-Square = 7.294, p = 0.011, Cramer's V=0.169, p = 0.007, df=1		Baseline	Explicit Action
Accepted	Count	95	112
	Expected count	103.5	103.5
	Standardised residual	-0.8	0.8
Rejected	Count	33	16
	Expected count	24.5	24.5
	Standardised residual	1.7	-1.7

Figure 1 below displays the acceptance rates of trials by condition as percentages of the total number of trials in each condition.

Figure 1 - Acceptance rates of error-free trials

An examination of the relevant statistics in each comparison shows that the Baseline and Different Coerced conditions did not significantly differ in terms of acceptance rates (Chi-Square = 0.189, p = 0.772, Cramer's V = 0.027, df = 1). However, both the Same Coerced condition and the Explicit Action condition did significantly differ from the Baseline condition (Same Coerced: Chi-Square = 9.111, p = 0.004, Cramer's V=0.189, df = 1, Explicit Action: Chi-Square = 7.294, p = 0.011, Cramer's V = 0.169, df = 1). An examination of the counts and expected counts reveals the reason for these effects. Participants were significantly more likely to reject pairs of sentences from the Same Coerced condition than from the Baseline condition. Participants were also significantly *less* likely to reject pairs of sentences from the Explicit Action condition. Therefore, although the nature of the relationship between prime and target sentences did not have an effect on reaction times, there was an effect on how acceptable participants found complement coercion structures.

5.3 Discussion

The reaction time data do not provide support for de Almeida and Dwivedi's (2008) claim that z-information specification should speed up the processing times of subsequent complement coercion structures because reaction times to the Baseline condition did not differ from the reaction times to the Explicit Action condition. However, because no significant effect of prime-target relationship on processing times was found in any condition, the likely explanation is that the methodology employed is not sensitive to reaction time priming effects and therefore it is not safe to draw any conclusions regarding this point. An improvement to the experiment reported here would be to make use of eye-tracking data because this is the most frequently deployed methodology in psycholinguistic research on complement coercion. This would make the results from a study that did not control for the confound identified by Katsika et al. (2012) and the results from a study that did control for this confound would be maximally comparable.

However, an examination of the acceptance rate data provides a novel finding. Previous research has not provided a measure of how context affects whether participants accept a complement coercion structure as grammatical, but only on how fast it takes participants to read a complement coercion structure. Recall that 26% of the critical trials (those trials that contained complement coercion structures) employed in this study were classed as containing violations. Participants were told to reject stimuli if they contained either a semantic or grammatical error. None of the critical trials contained such an error. One possibility is that participants generated a full interpretation of those critical trial pairs that they identified as containing no errors, and rejected those pairs for which a complement coercion structure was not fully interpreted. In the Explicit Action condition, it seems plausible that encountering a potential action in the prime sentence with which to interpret a subsequent underspecified complement coercion target sentence is the reason that participants were more likely to judge a complement coercion structure to be error free. Considering the reaction time data and the acceptance rate data together, these results suggest that although explicitly introducing the action that is left implicit in a complement coercion structure might not speed up processing time, it does aid the process of generating an interpretation. The acceptance data for the Same Coerced condition are more problematic to interpret. Raffray, Pickering, Cai, and Branigan (2013) show that identical coercing expressions show priming from comprehension to production, and Traxler et al. (2005) show that repeating a coercing structure speeds up processing time. Therefore it is somewhat mysterious that repeating a coercing structure across prime and target in the current experiment makes it *less* likely for a target sentence to be accepted. One potential explanation is that because complement coercion structures are difficult to process relative to traditionally compositional structures, participants simply had problems parsing two such structures one after the other in a short space of time while responding to the requirements of the task. However if this were the reason for increased rejection rates, it would be expected that the Different Coerced condition would also show this effect, but the Different Coerced condition did not differ significantly from the Baseline condition. Given its non-intuitiveness, it would be prudent to replicate this finding before committing to any explanation of it. However, these data suggest that, contrary to Traxler et al.'s (2005) claim that complement coercion structures generally have default interpretations, it is relatively difficult to generate a full interpretation of an isolated complement coercion structure.

There are two potential criticisms of the experiment reported here that need to be addressed. Firstly, it is reasonable for de Almeida and Dwivedi to claim that z-information specification requires a richer discursive context than that found in the relationship between disconnected prime and target sentences. Therefore, the fact that the Explicit Action condition

did not show priming relative to the Baseline condition is not necessarily evidence against their theory. This may be true, but it is still clear that compared to other conditions, the Explicit Action condition showed a higher acceptance rate of complement coercion structures and that the most likely reason for this is that the implicit action is specified, allowing for an easier interpretation. Although de Almeida and Dwivedi (2008) do not specifically predict this pattern, it is compatible with their view that the resolution of underspecification accounts for at least some of the difficulty in processing complement coercion structures. However, as argued above, their theory is difficult to apply because it is just not clear how they propose that the syntactic and pragmatic elements of the account interact. Secondly, the number of critical trials in the study was relatively small ($n=16$). Given that Katsika et al. (2012) provide a strong motivation for reducing the size of the class of complement coercers to a handful of semantically related aspectual verbs, this small stimulus set is unavoidable without requiring a large number of verb repetitions across the experiment, which would be potentially problematic because unwanted repetition priming effects would be likely to appear. This issue is therefore unavoidable with the methodology employed here, and is a weakness of the study.

De Almeida and Dwivedi's (2008) syntactico-pragmatic theory has been discounted on the grounds that it is not consistent. It is unclear whether they intend for purely syntactic processing to account for extra comprehension costs, or whether pragmatic inferences also result in comprehension costs. However, a strength of de Almeida and Dwivedi's (2008) account is that it allows for the effects of pragmatic processing on ambiguous structures. The implications of the results reported here for the lexico-semantic theory will now be discussed, and it will be argued that the lexico-semantic theory is lacking a vital pragmatic component. McElree et al. (2001) allow that specific linguistic elements of the discourse context, such as the subject of the verb, can have an effect on how complement coercion structures are interpreted. Traxler et al. (2005, Experiments 3 & 4) show that within a contained narrative, introducing the entire event sense of a fully interpreted complement coercion structure in a context sentence speeds up the processing of a subsequent complement coercion structure in a target sentence. The acceptance rate data reported here show that in relatively artificial, non-narrative contexts, a prime sentence that contains a likely candidate for the implicit action in a target complement coercion structure is enough to aid interpretation even when there is no overt discursive connection between prime and target.

Traxler et al. (2005, p. 18) argue that their results show that 'it is highly unlikely that [the type of ambiguity that features in complement coercion structures] is the source of the observed cost in processing these expressions'. However, despite the fact that their account is confused, de Almeida and Dwivedi (2008) are right to claim that the underspecification seen in complement coercion structures is resolved pragmatically at least some of the time. On the other hand, and contra to de Almeida and Dwivedi's (2008) position, it is highly likely that this pragmatic resolution comes at a cost. It is still uncertain whether Traxler et al.'s (2005) eye-tracking data are valid because it is possible that an eye-tracking paradigm or a self-paced reading paradigm that controlled for Katsika et al.'s (2012) confound would show a greater effect of context on processing times. There are also two arguments against Traxler et al.'s (2005) claim that complement coercion structure ambiguity is resolved on a purely lexical basis, one of which is empirical and one of which is theoretical. Traxler et al. (2005) maintain that it is a strength of their account that it explains how complement coercion structures have default interpretations. However, it is not clear that many of these expressions do prompt 'default' interpretations. In their experiments, Traxler et al. (2005) placed pre-normed complement coercion structures in contexts that were relatively rich in content. In the experiment reported in the current paper, nearly a full quarter of trials containing complement coercion structures were rejected, presumably at least partly because the context in which they appeared did not allow a default interpretation. If a default interpretation was available then

‘correct’ acceptances of Same Coerced trials should have been as high as in the filler trials. However, 94.8% of filler trials that contained no errors and no underspecification were correctly accepted compared to just 56.3% of the Same Coerced trials. By comparison, in the Explicit Action condition, where the prime presumably aided the resolution of the ambiguity present in the target, the acceptance rate was 87.5%. It is also important to stress that the prime sentences in the Explicit Action condition do not contain fully lexicalised instantiations of the complement coercion structure present in the target sentences, as was the case in the Traxler et al. (2005) experiments. The coercing verb is not featured in the prime sentence and the arguments of the verb are not repeated across prime and target either. Although Traxler et al. (2005) may well be right to argue that the effect of context on complement coercion structures is limited in some cases, the evidence presented here suggests that this effect is still slightly larger than they allow.

Additionally, there is an explanatory gap in the lexico-semantic account regarding how the interpretation of an underspecified complement coercion structure is generated. The only way of filling this gap is to accept that pragmatic processing plays a role. Suppose that some complement coercion structures do have a default interpretation, or set of default interpretations, and that these default interpretations are a function of information resembling Pustejovsky’s qualia structure. In the case of relatively transparent complement coercion structures, such as *The author began the book*, this seems plausible. However, there are still questions that this explanation leaves unanswered: What makes it possible for a decision to be made regarding whether the default interpretation is the one a speaker intended to convey? When a default interpretation is incorrect, on what basis does a hearer make this decision and how is the correct interpretation generated? And finally, it is trivial to construct cases where there is obviously no default interpretation of a complement coercion structure, but in which context resolves the ambiguity. For example, consider the following narrative:

“The little girl took out her toy construction set and methodically laid out all the pieces according to their size and colour. After reading the instructions carefully, she began the railway. A few hours later, her mother called her for tea.”

In comparison to cases such as ‘began the book’, it is much less plausible to suggest that the verb phrase ‘began the railway’ has a default interpretation that can be derived from the qualia structure of the noun ‘railway’, or the qualia structure of the noun ‘girl’, which the pronoun ‘she’ refers to in the extract above. The fact that the little girl in this story is understood to be building a toy railway has nothing to do with the lexical representations of the words ‘began’ or ‘railway’, and little girls are not by default associated with the act of railway building. Instead, it is clear that this interpretation is derived from the wider context in which the complement coercion structure appears, and the reader is able to make use of world knowledge about what toy construction sets typically contain and what a little girl might do with one, as well as the intentions of the writer/speaker in communicating the narrative in the first place. As Recanati (2004, p. 93) explains:

[the semantic value of semantically underdeterminate expressions] varies from occurrence to occurrence... yet it varies not as a function of some objective feature of the narrow context but as a function of what the speaker means... the semantic value of the expression – its own contribution to what is said – is a matter of speaker’s meaning, and can only be determined by pragmatic interpretation.

Traxler et al. (2005, p. 18) claim that ‘enriched composition is costly because readers must... [incorporate the action that is left implicit] into the semantic representation of the VP’, but in

many cases this explanation leaves out how it is that comprehenders decide what the implicit action is in the first place. Even in relatively transparent cases such as *The author began the book*, where something like Pustejovsky's nominal qualia structure and telic/agentive roles might plausibly provide the information with which to build the event sense 'began writing the book', a comprehender still has to somehow decide which role is appropriate. In order to do this, a comprehender has to make assumptions about what a speaker intended to mean based on context. Making use of this knowledge is unlikely to be cost-free, and the mental representations involved must be formed at least partly of information that cannot be found at the lexical level. In and of itself, the content of the lexical representations of the words *author*, *book*, and *began* is not sufficient to determine what a speaker intended an utterance of 'The author began the book' to mean.

6 Conclusion

This paper explored two theories of the interpretation of complement coercion structures: the lexico-semantic theory and the syntactico-pragmatic theory. The psycholinguistic literature demonstrates that structures such as 'the girl began the book' engender additional processing costs relative to control structures. The syntactico-pragmatic theory holds that this additional cost is due either to the presence of extra syntactic structure or the pragmatically driven assignment of a value to an empty verb head contained in the extra syntactic structure. However, this account is inconsistent in describing which of these elements engenders the cost, or if both do. The lexico-semantic theory provides an account of complement coercion that is consistent, but incomplete without the allowance that pragmatic inferences are necessary for the interpretation of underspecified complement coercion structures and that these inferences probably do come at a processing cost. The experiment reported here controlled for a confound that renders the findings of previous studies of complement coercion open to question. Although this experiment returned a null result regarding the effect of prime sentences on the time taken to process target complement coercion structures, the acceptance rate data do show an effect. Introducing a likely action with which to complete the interpretation of a complement coercion structure made participants more likely to accept that complement coercion structure as error-free, which is most likely because they were able to generate a full interpretation. This is in contrast to Traxler et al.'s (2005) claims that complement coercion structures have default interpretations and that increased cost is solely due to operations on lexical representations. A subsequent study that employed an eye-tracking paradigm and also controlled for the confound present in previous studies would provide data that was maximally comparable to Traxler et al.'s (2005) results and would clarify the effect of context on the processing of complement coercion structures.

Appendix**Critical Stimuli****Baseline stimuli**

List	Prime	Target
1	The mother lost the card.	The professor finished the letter.
1	The labourer used the sink.	The mechanic finished the car.
1	The pensioner refused the painkiller.	The patient finished the drug.
1	The jeweller examined the stone.	The sculptor continued the statue.
2	The artist dropped the portrait.	The student started the picture.
2	The boy heard the song.	The pianist started the sonata.
2	The soldier destroyed the bridge.	The farmhand completed the henhouse.
2	The worker drafted the blueprint.	The explorer completed the map.
3	The man bought the newspaper.	The girl began the book.
3	The grandmother served the pie.	The chef began the entrée.
3	The aunt washed the clothes.	The seamstress began the quilt.
3	The lecturer explained the problem.	The grandfather continued the crossword.
4	The producer admired the stage.	The architect continued the house.
4	The teenager spilt the juice.	The baby continued the milk.
4	The carpenter broke the table.	The goldsmith started the ring.
4	The medic saw the bullet.	The dentist began the tooth.

Explicit action stimuli

List	Prime	Target
1	The artist was drawing the portrait.	The student started the picture.
1	The boy was playing the song.	The pianist started the sonata.
1	The soldier was building the bridge.	The farmhand completed the henhouse.
1	The worker was sketching the blueprint.	The explorer completed the map.
2	The man was reading the newspaper.	The girl began the book.
2	The grandmother was cooking the pie.	The chef began the entrée.
2	The aunt was sewing the clothes.	The seamstress began the quilt.
2	The lecturer was solving the problem.	The grandfather continued the crossword.
3	The producer was designing the stage.	The architect continued the house.
3	The teenager was drinking the juice.	The baby continued the milk.
3	The carpenter was making the table.	The goldsmith started the ring.
3	The medic was extracting the bullet.	The dentist began the tooth.
4	The mother was writing the card.	The professor finished the letter.
4	The labourer was fixing the sink.	The mechanic finished the car.
4	The pensioner was taking the painkiller.	The patient finished the drug.
4	The jeweller was carving the stone	The goldsmith started the ring.

Same coerced stimuli

List	Prime	Target
1	The producer continued the stage.	The architect continued the house.
1	The teenager continued the juice.	The baby continued the milk.
1	The carpenter started the table.	The goldsmith started the ring.
1	The medic began the bullet.	The dentist began the tooth.
2	The mother finished the card.	The professor finished the letter.
2	The labourer began the sink.	The mechanic began the car.
2	The pensioner finished the painkiller.	The patient finished the drug.
2	The jeweller continued the stone.	The sculptor continued the statue.
3	The artist started the portrait.	The student started the picture.
3	The boy started the song.	The pianist started the sonata.
3	The soldier completed the bridge.	The farmhand completed the henhouse.
3	The worker completed the blueprint.	The explorer completed the map.
4	The man began the newspaper.	The girl began the book.
4	The grandmother began the pie.	The chef began the entrée.
4	The lecturer continued the problem.	The grandfather continued the crossword.
4	The aunt began the clothes.	The seamstress began the quilt.

Different coerced stimuli

List	Prime	Target
1	The man finished the sandwich.	The girl began the book.
1	The grandmother started the coffee.	The chef began the entrée.
1	The aunt started the plates.	The seamstress began the quilt.
1	The lecturer finished the diagram.	The grandfather continued the crossword.
2	The producer began the cigarette.	The architect continued the house.
2	The teenager finished the homework.	The baby continued the milk.
2	The carpenter continued the plans.	The goldsmith started the ring.
2	The medic continued the form.	The dentist began the tooth.
3	The mother completed the lawn.	The professor finished the letter.
3	The labourer completed the list.	The mechanic finished the car.
3	The pensioner began the report.	The patient finished the drug.
3	The jeweller began the inventory.	The sculptor continued the statue.
4	The artist completed the lake.	The student started the picture.
4	The boy continued the puzzle.	The pianist started the sonata.
4	The soldier began the mountain.	The farmhand completed the henhouse.
4	The worker finished the wall.	The explorer completed the map.

Filler Stimuli

Prime	Target	Violation
The fireman saved the pet.	The lifeguard will save the surfer.	no
The manager will file the reference.	The clerk filed the document	no
The breeder trains the dog.	The coach was training the athlete.	no
The gardener was planting the bush.	The climber plants the flag.	no
The janitor was sweeping the hall.	The runner will attempt the marathon.	no
The technician will install the machine.	The officer polishes the medal.	no
The plumber opens the valve.	The guard closes the door.	no
The schoolgirl was watching the sitcom.	The diver found the shipwreck.	no
The mayor awarded the prize.	The instructor will describe the course.	no
The zookeeper watches the lion.	The mugger threatened the victim.	no
The cowboy rode the horse.	The traveller will board the ship.	no
The fisherman caught the salmon.	The doctor will prescribe the treatment.	no
The criminal break the window.	The drunkard broke the vase.	prime, agreement
The musician play the saxophone.	The director pays the contractor.	prime, agreement
The fraudster were destroying the evidence.	The engineer detonates the bomb.	prime, agreement
The author were describing the story.	The crook will hide the weapon.	prime, agreement
The surgeon are installing the pacemaker.	The chemist performs the experiment.	prime, agreement
The politician was visiting the love.	The queen will visit the country.	prime, animacy
The girl patted the fear	The van was driven by the builder.	prime, animacy
The youth was mowing the luck.	The foreman planted the sign.	prime, animacy
The nurse tends the laughter.	The data was processed by the analyst.	prime, animacy
The truth will drive the convertible.	The spy escaped the patrol.	prime, animacy
The watch was sell by the auctioneer.	The gardener was tending the rosebush.	prime, tense
The child throwing the ball.	The word was learned by the baby.	prime, tense
The striker score the goal.	The nanny will pack the suitcase.	prime, tense
The inspector will imposed the fine.	The editor rejected the article.	prime, tense
The baker burnt the cake.	The butler burn the toast.	target, agreement
The mechanic scrapped the car.	The camper were stowing the tent.	target, agreement
The argument was refuted by the philosopher.	The gardener were tending the rosebush.	target, agreement
Prime	Target	Violation

The critic praises the film.	The postman deliver the parcel.	target, agreement
The barber trimmed the beard.	The pupil were learning the language.	target, agreement
The manuscript was ripped by the scribe.	The dress was ripped by the faith.	target, animacy
The president prepares the speech.	The anger cleans the cabin.	target, animacy
The man will lick the stamp.	The frown was watching the game.	target, animacy
The driver ignores the signal.	The teacher grades the camel.	target, animacy
The actress practised the script.	The pupil was practised the exercise.	target, tense
The employee will find the key.	The policeman will finding the gun.	target, tense
The tinker mended the clock.	The thug will steals the money	target, tense
The maid polishes the cabinet.	The uncle clearing the table.	target, tense
The poet composed the haiku.	The pilot has flying the plane.	target, tense

References

- De Almeida, R. (2004). The effect of context on the processing of type-shifting verbs. *Brain and Language*, 90 249-261 DOI:10.1016/S0093-934X(03)00438-3
- De Almeida, R. & Dwivedi, V. (2008). Coercion without lexical decomposition: Type-shifting effects revisited. *Canadian Journal of Linguistics* 50 (2/3) 201-326 DOI: 10.1353/cjl.0.0026
- Frazier, L. & Fodor, J. (1978). The sausage machine: A new two-stage parsing model. *Cognition* 6 291-325 DOI: 10.1016/0010-0277(78)90002-1
- Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press
- Kamp, H. and Partee, B. 1995. Prototype theory and compositionality. *Cognition* 57 (2) 129-191 DOI: 10.1016/0010-0277(94)00659-9
- Katsika, A., Braze, D., Aswini, D., & Piñango, M. (2012). Complement coercion: Distinguishing between type-shifting and pragmatic inferencing. *The Mental Lexicon* 7 (1) 58-76 DOI: 10.1075/ml.7.1.03kat
- McElree, B., Traxler, M., Pickering, M., Seely, R., & Jackendoff, R. (2001). Reading time evidence for enriched composition. *Cognition* 78 B17-B25 DOI: 10.1016/S0010-0277(00)00113-X
- Pickering, M., McElree, B. & Traxler, M. (2005). The difficulty of coercion: A response to de Almeida. *Brain and Language* 93 (1) 1-9 DOI 10.1016/j.bandl.2004.07.005
- Pustejovsky, J. (1995). *The Generative Lexicon*, MIT Press, Cambridge, MA
- Pylkkänen, L., Liinas, R., & McElree, B. (2004). Distinct effects of semantic plausibility and semantic composition in MEG. In E. Halgren, S. Ahlfors, M. Hämäläinen, & D. Cohen (Eds.), *Biomag 2004: Proceedings of the 14th international conference on biomagnetism*. Boston, USA: Biomag.
- Raffray, C., Pickering, M., Cai, Z., & Branigan, H. (2013). The production of coerced expressions: Evidence from priming. *Journal of Memory and Language* 74 91-106 DOI: 10.1016/j.jml.2013.09.004
- Recanati, F. (2004). *Literal Meaning*. Cambridge: Cambridge University Press DOI: 10.1111/j.1933-1592.2007.00088.x
- Traxler, M., Pickering, M., & McElree, B. (2002). Coercion in sentence processing: evidence from eye-movements and self-paced reading. *Journal of Memory and Language* 47 (4) 530-54 DOI: 10.1016/S0749-596X(02)00021-9
- Traxler, M., McElree, B., Williams, R., & Pickering, M. (2005). Context effects in coercion: Evidence from eye movements. *Journal of Memory and Language* 53 (1) 1-25 DOI: 10.1016/j.jml.2005.02.002

Analogical cognition and understanding a word*

Timothy Pritchard

Abstract

Gentner describes a notion of ‘analogical cognition’ that could play a significant role in elucidating what is involved in understanding a word. Gentner’s work has not, though, had much or any impact in linguistics or the philosophy of language. I explain key features of Gentner’s notion, and I argue that it explains how word understanding can be stable, specific, and shared, and how words can contribute to cognition as opposed to just being a means of conveying thoughts.

Keywords: analogical cognition, Gentner, word meaning, understanding, good enough interpretation

1 Introduction

Speakers of a language appear to have a good understanding of most of the words they use. They have an easy and pre-reflective familiarity with individual words, and they draw on this both when speaking and when listening. Even if the thoughts that speakers use combinations of words to convey can be difficult to understand, this difficulty is not typically felt at the level of the individual words themselves. I take these appearances to be well founded and to legitimate the following question: can we find a way to characterize how we understand individual words? The intended focus of the question is not on what we might use a word or words to convey in a given instance, but on our prior understanding of the words, the understanding that we draw on when we communicate. There appears to be little confidence among theorists about how to approach this question; for example, while Recanati (2004) does discuss various options, there is little by way of development of, or debate between, these options.

In this paper, I offer a specific proposal that is based on Gentner’s work on analogical cognition. Her work provides, I believe, resources for making a significant contribution to the debate. T. Bach (2012) has commented that the study of analogical cognition has had little impact on the philosophy of mind and language, and I agree with his opinion that this could be an ‘expensive oversight’. Insight from Gentner’s work helps to motivate and justify the claim that our understanding of particular words is stable, shared, and quite specific. It enables us to delimit a specifically linguistic level of understanding as opposed to more general types of understanding, and it helps to explain how the use of words makes a special contribution to cognition. The account stands critically over against approaches that, inspired by reflection on context sensitivity, have tended to offer indirect, open-ended, and non-specific accounts of the contribution that words bring to utterances.

My preferred terminology is to phrase the debate in terms of understanding rather than in terms of (word) meaning. Notions of meaning often relate to the question of how to represent what a given use of language has conveyed, and this is not always helpful for considering our grasp of individual words. I would prefer to take our understanding of individual words as prior

* I would in particular like to thank Andrew Tolmie for indirectly pointing me towards Gentner’s work. Also, for helpful discussion and corrections, Josie Bowerman, Robyn Carston, Léna Kervran, Nathan Klinedinst, Lewis Pollock.

in order of explanation to accounts of meaning, but nothing turns on this distinction in this paper. Sometimes, though, it will be more natural to speak of word meaning, in particular when this reflects the terminology of other theorists. We might think of a word meaning as the corollary to what is understood in the understanding of a word, though it is useful to note that this takes us away from a more natural emphasis in which ‘meaning’ relates to what is conveyed in the use of language.

So, while understanding is my preferred term, I will also at times speak of word meanings. Either way, the focus is on words considered in a sense prior to their use in utterances. I ignore aspects of understanding that relate specifically to syntactic patterns and morphology, so the focus is more towards what might be described as an underlying or root understanding that may (but need not) recur across different word forms. For example, the same root understanding typically recurs in ‘capitulate’ and ‘capitulation’ but not typically in ‘capitulate’ and ‘recapitulate’.¹

In section 2, I describe two characteristics of language that an account of word understanding should help to explain. In sections 3 and 4, I describe the relevant aspects of Gentner’s notion of analogical cognition, and in section 5, I propose an account of word understanding that is based on this. In section 6, I discuss how the account fits the two characteristics of language described in section 2. In section 7, I consider how the contribution words make to cognition challenges some accounts of the basis for word use.

2 Two characteristics of language

Two characteristics of language that an account of word understanding should help to explain are context sensitivity and the phenomenological immediacy of our experience of meaning when we hear language. There are other important characteristics, some of which I shall mention (such as the ease with which children learn words), others of which I do not directly address (such as compositionality). The two I focus on are particularly pertinent for the account I describe.

Emphasis on a stable, shared, and quite specific understanding (or, stable and specific word meaning) is often seen as problematic. Reflection on language use has suggested to many theorists that we need to treat words’ contributions to utterances in an open-ended and indirect fashion, and that it may even be wrong to think of words as having any specifiable meaning attached to them at all (for overview, see Carston (2012)). These suggestions have arisen from consideration of the flexibility with which words are used, for example as reflected in polysemy or in nuances of interpretation that are sometimes described as involving a modulation of word meanings (Wilson & Carston, 2007). The contribution that words make has been described with terms such as ‘constraint’ (Travis, 2000), ‘clue’ or ‘pointer to conceptual space’ (Carston, 2002), something that points to ‘indefinitely many notions or concepts’ (Sperber and Wilson, 1998), ‘instruction for creating a concept from available mental resources’ (Pietroski, 2005), ‘polysemy complexes’ (Carston, 2016).²

¹ H. Borer (2005) gives an important place to *word* roots (after the pattern of tri-consonantal Hebrew roots), a level of word analysis that is prior to any particular morphological implementation. She does not though speak of a (differently grained) notion of root understanding.

² Borg, who dissents from this general approach, characterizes the view as one that treats word contributions ‘as open-ended, web-like things which stand in need of contextual precisification prior to fixing their input to larger linguistic units’ (2012, xvii). Her own view is that word contributions are ‘discrete, probably atomistic, blob-like things’. The account I propose supports a move away from the open-ended to something more specific.

These suggestions indicate roles, and there has been discussion of what plays these roles. Carston has considered (though now rejects) the view that word meanings could be ‘schemas’ (2002; 2012; 2013). A schema is envisaged as a highly abstract representation that captures merely an outline of the different meanings that a word can be used to express. The schema is, as it were, the skeletal starting point for interpretation, with a fleshing out needed before we attain to an interpretation that will be typical of a given use of a word. Alternatively, word meanings have been characterized as containing more than is required for interpretation. Rather than needing to be fleshed out, a filtering process is required, in which the relevant interpretation is selected (see Recanati (2004), p.140).

More radical suggestions have been proposed, in which word meaning is eschewed altogether. Drawing on Wittgenstein, Recanati (2004) describes (without fully endorsing) a ‘meaning eliminativist’ view. On this view, word use is explained via knowledge of previous uses of a word (source situations), with further use of the word licensed when a novel situation is appropriately similar to those previous uses. Rayo (2014) likewise proposes that rather than thinking in terms of linguistic meanings, it is better to see speakers as associating words with a ‘grab bag’ of mental items: memories, mental images, pieces of encyclopaedic information, and so on. What is conveyed by a word arises from some process that is based on language users having access to this grab bag of items.³

The suggestions reflect the claim that a word is but a pointer, or clue, towards the relevant interpretation of a use of a word. There may or may not be a specifiable meaning to a word, but either way the word acts as an initial input that requires further work before the required interpretation can be recovered. Language users’ grasp of a word does not directly correspond with what a word is used to express on an occasion of use.

There is, though, besides context sensitivity, another and apparently contrasting aspect to our use of language. This is the observation that, on hearing words, understanding arises in such a direct and involuntary manner that it is almost perceptual in character. Fricker (2003) describes how when we hear language, we seem to experience the meaning of the words in the very hearing of the utterance itself; the utterance is experienced as a ‘semantically laden event’. Fricker calls this an understanding *experience*, a quasi-perception. It seems as if the meaning is an intrinsic property of the utterance.

[W]hen one understands an utterance of a sentence, the immediate object of one’s aural experience is not a mere-burst of sound, nor merely a syntactic entity. The phenomenal given in normal language perception is an utterance perceived as syntactically structured and semantically loaded. One hears the words and one hears what they mean—one hears them as expressing a certain proposition. (2003, pp. 337–8)

While there is debate as to whether the parallel with perception is justified, the broad description of the phenomenology is not controversial: at the conscious personal level, meanings seem, at least very often, to directly accompany the words that are used. Further, we cannot choose to avoid uptake of meaning when we hear language. As Pettit (2010) describes it:

Our awareness of the content of speech is typically immediate and unreflective. Indeed, the content of speech in your native language can stubbornly intrude upon

³ Arguably, meaning eliminativist views are merely more thorough-going presentations of the ‘rich’ view. Both are characterized by a need to filter out the relevant use from an array of interpretive material (see Carston (2012), p. 620, on Cohen).

your mental life when you would rather it not, as for example when someone is publicly sharing personal details while talking on the phone. You can plug your ears or try to attend to something else, but you cannot simply block out the content of the speech.

Fricker and Pettit are careful to note that these personal-level observations about the immediacy of meaning may mask complex underlying processes. But it is also reasonable to require that an account of word understanding should help to explain the phenomenon. Prima facie, it is not clear that talking of words as pointers or clues or constraints is best suited for providing an explanation, as they are approaches to words that stress indirectness and inference rather than involuntary directness.

A parallel point can be made with respect to the speed with which words are processed. Keil and Kominsky (2015) express this (in terms of ‘concepts’):

concepts, as manifested through words, must be processed at rates compatible with both producing and understanding speech and the reading of text. When a reader seems to fully comprehend a text at three hundred words a minute, how are hundreds of concepts accessed, deployed, and interrelated each minute? ... [F]or concepts to work in this way, they must impose a very low cognitive load and not require much inferential machinery in real time. It would not be possible to deploy concepts so quickly and effortlessly if they required extensive processing of an elaborate web of beliefs.

If words are pointers or clues, this may imply a cognitive load that is hard to reconcile with the speed of processing. The immediacy and speed of our uptake of meaning may more naturally suggest that individual words have specific and easily understood meanings, which directly, and not via inference, inform our understanding of the utterances in which those words are used. This of course raises the problem of how to explain the evidence that word use is highly flexible and context sensitive.

3 Analogical cognition⁴

A striking feature of human cognition is our ability to handle higher-order relations. Consider the following matching task, which requires a match to be made between the appearance of the base sample (the letter A) and one of the target letters:

Base	Target
A	A B

An ability to evaluate the perceptual similarity between two stimuli has been described as ‘clearly the sine qua non of biological cognition’ (Penn, Holyoak, & Povinelli, 2008), and the ability to match the Base A to the Target sample A is not limited to humans. The matching is based on comparing features of the various tokens. But consider the following different type of matching task:

⁴ I focus on those aspects of Gentner’s work that are particularly relevant to the understanding of individual words. Her work on analogy goes beyond this, including consideration of various learning and reasoning abilities.

Base	Target
AA	BB CD
XYX	ZZW UVU

This is a relational-match-to-sample test, and it is not solved by looking for perceptual similarity between the individual letters in the Base and the Target. Instead, the matching task looks first to a relation that holds between the individual tokens that make up each sample, with matching based on recurrence of that relation. AA matches to BB, and not to CD, because the same relation holds between A and A (namely, sameness) as holds between B and B. Similarly, XYX matches to UVU, not to ZZW, because both display the relation of symmetry.

The examples illustrate the idea of a structure into which individual items enter, structure that, at least for humans, invites a form of cognition that is based on an appreciation of that structure. Even if the items themselves differ considerably, they can still enter into the same relations.

[T]he corresponding objects in the base and target need not resemble each other; what is important is that they hold like roles in the matching systems of relations.

(Gentner, 2010, p. 752)

A railway engine can have the same relation to a carriage as a tractor does to a cart. Humans are sensitive to this shared structure and express it with terms such as ‘pulls’. A concrete block that prevents a car from progressing down a road manifests a situation that shares structure with a book that prevents a marble from rolling further on a table. Humans express this with terms such as ‘obstacle’.

Analogical ability, also called relational cognition, is the ability to perceive like relational structures across different contexts. There is debate as to whether analogical cognition is limited to humans (see Penn et al. (2008)), but there is general agreement that human thought is specially characterized by an ability to engage in this kind of relational matching and that this ability greatly exceeds any relational matching ability that some animals might turn out to possess. It is ubiquitous in human cognition and has been described as the hallmark of human intellectual ability:

Analogical ability ... is a core mechanism of human cognition. The ability to perceive and use purely relational similarity is a major contributor—arguably *the* major contributor—to our species’ remarkable mental powers.

(Gentner & Smith, 2013)

Gentner and her colleagues have shown that this analogical ability is in particular helped by the process of *comparing* different situations or entities (Gentner 2005; 2010). Research shows that when we compare different items, we preferentially highlight not the perceptual similarities but the relational commonalities. According to what Gentner calls structure-mapping theory, this is because when people compare two items or situations, they implicitly seek to find structural correspondences (Gentner & Markman, 1997; Gentner, 2003). This directs attention away from consideration of whether the compared items share an overall (and in particular perceptual) likeness, and towards consideration of shared structure. This ‘relational shift’ gives rise to a form of learning (schema abstraction) in which the common structure is extracted.

This relational shift has been illustrated in young children. In a classic search task,

children watched an experimenter hide a toy in a small model room and had to find another toy hidden ‘in the same place’ in the full-sized room. Children under 3 years of age did very poorly at this task. Building on these results, [a further study] showed (a) if 2.5 year olds are given a similar-scale search task (so that model and room are highly similar) they perform very well (67% correct searches); and (b) when these children are brought back the next day and tested in the standard model-room task (small model and full-sized room), they are far more successful than age-mates who did not receive the highly similar pair first (35% correct). (Gentner & Smith, 2013)

When the model and the room were similar in size, this facilitated the task of looking for an item ‘in the same place’ in both rooms. This also led, by the implicit comparison undertaken of the two rooms, to an appreciation that both rooms shared a structure and that sameness of place is evaluated with respect to that structure (and not, for example, with respect to local perceptual similarities). This is shown in the children displaying improved ability in evaluating sameness of place when confronted with the different sized rooms. Importantly, learning from highly similar things did not lead the children to put an increased emphasis on perceptual similarity as the criterion for sameness of place. Instead, the effect of similarity between the rooms was to make the act of comparing easier. Gentner calls this ‘progressive alignment’, with the comparison of similar things acting as ‘training wheels’ (Gentner, 2010) that help the child to learn of relational matches that can subsequently be used to find samenesses across perceptually dissimilar items.

In another study, children (aged 3.5 years) watched the experimenter hide a toy in a small model room (the Hiding room), and then tried to find the toy hidden ‘in the same place’ in a second model room (the Finding room). The two rooms contained the same types of furniture (bed, table, etc.) in the same configuration, but the specific shapes of the furniture in the Finding room differed from the shapes of the furniture in the Hiding room.

Before engaging in the task, all the children were shown the Hiding room along with another highly similar room (identical except for color). Half the children saw the two rooms together and were encouraged to compare them; the other half discussed each room separately. Children who had compared the rooms were significantly more likely to correctly locate the toy in the Finding room than those who had experienced the rooms separately. Thus, comparing two nearly identical rooms facilitated children’s ability to map their common spatial relational structure to a relationally similar but surface-dissimilar target.

(Gentner, 2010, pp. 759–60)

That is to say, when the children were encouraged to compare the two highly similar rooms, they showed a tendency to pick up on a relational specification of the location of the object (viz., looking to the same relative position) and to be less influenced by the changing shapes of the items of furniture involved.

For adults, likewise, eliciting comparisons helps induce a focus on abstract relational commonalities. Business school students who compared two negotiation scenarios were over twice as likely to transfer the negotiation strategy to an analogous test negotiation as were those who studied the same two scenarios separately (Gentner, 2010, p. 758).

In summary, these and many other studies show that comparing items or situations leads human thinkers to emphasize structural commonalities between those items or situations. Comparison renders the common structure more salient, and leads to a form of learning in which the common structure is extracted. This underpins analogical cognition, which is the

ability to perceive the same structural relations across different situations. Gentner has also shown that analogical cognition is closely related to our use of words.

4 Analogy and language

When discussing what it is that makes humans so smart, Gentner does not just talk about relational cognition; she combines it with our use of language. What makes us smart is ‘a mutual bootstrapping system composed of (a) our exceptional capacity for relational cognition and (b) symbolic systems that augment this capacity’ (Gentner, 2010). Language supports our analogical reasoning abilities.

Experimental work provides plenty of evidence for the important role that a linguistic label itself can play in aiding a shift to relational thinking. For example, Gentner (2003) reports an experiment in which children (age 3 to 5) were shown two groups of three objects (the child’s set and the experimenter’s set) arranged according to size. For example, in a simple case there might be three differently sized cups. Children watched as the experimenter hid a sticker under an object in the experimenter’s triad. The children were told they could find their sticker by looking ‘in the same place’ in their own group of objects. The criterion for the same place was relative size; so if, for example, the experimenter placed the sticker under the largest object, the child should choose the largest object in the child’s group of three objects. The children were indirectly taught about this criterion in 14 trial attempts. In each trial, the children would suggest an answer and then be shown the correct response.

The children were particularly unsuccessful in the task when the two groups of objects contained the same items but differently relationally placed. For example, suppose that the largest item (under which the sticker is placed) in the experimenter’s triad was a cup, and the children’s group of three objects had a cup but it occurred in the middle or smallest position. With this type of condition, instead of choosing the largest item in their own group of objects, children would often choose the cup, answering on the basis of perceptual similarity. In other words, in spite of repeated trials in which the relational criterion was used, the children were, under certain conditions, quite easily side-tracked away from a focus on this criterion and towards perceptual similarity.

This dramatically changed when the objects in the groups received linguistic labels that related to the relational dimension. For example, the labels could be ‘big’, ‘little’, ‘tiny’. With this change to the conditions of the experiment, there was a massive shift to relational thinking and away from the use of perceptual similarity as a guide. The effect persisted over time, with the label-trained children performing well on future tests even when the labels were no longer used. In other words, use of the linguistic labels helped to shift evaluation of what it is to be ‘in the same place’. Gentner suggests that children learn to become analogical thinkers, with language facilitating this relational shift.

Similar results have been found for spatial terms such as ‘top’, ‘middle’, ‘bottom’ (Loewenstein & Gentner, 2004). When the position of an item is described as on ‘top’ of a box (for example), children very much more readily recognized that the relational position was the relevant criterion when asked to locate a test item that is in the ‘same’ position with respect to a different box. When the position of the item was described simply as ‘here’, children were less ready to recognize or act on the relational criterion.

Christie and Gentner (2014) report an experiment in which children (average age 2.5 years) were tested on a relational-match-to-sample test. Initially, even after feedback, the children did not pick up on the relational criterion (answering at only chance levels). But when the children were given a novel linguistic label for the standard (a ‘truffett’), and told to apply that label to the choice of alternatives, there was a shift to the relational criterion. This effect is

described as ‘quite remarkable’ given that the word was novel and the children in this condition received no training with the novel word or feedback. It appears that when a matching task is presented in terms of applying a term, this facilitates access to a relational criterion amongst very young children for whom that criterion is otherwise not particularly accessible.

Similarly, Gentner (2005) reports an experiment that used the novel label ‘blick’. In the first condition, children were told that ‘The knife is the blick for the watermelon. The axe is the blick for the tree’. They were then asked, ‘What would be the blick for the paper?’, and were offered three options: a thematic response (pencil), a perceptually similar response (another piece of paper), and a relational response (scissors). In a second condition, instead of using this novel ‘blick’ label, the children were told that the knife goes with the watermelon in ‘the same way’ that the axe goes with the tree. They were then asked what would go with the paper in ‘the same way’. For 4 and 6 year olds, the relational response was more likely in the blick condition than in this second condition.

An inviting explanation of this is that linguistic terms, being applicable to multiple different particular situations, are naturally aligned with the relational commonalities that comparison of situations encourages. Labelling a situation will thereby prompt a search for some relevant structural relation (Gentner, 2005). This gives rise to a process of abstraction, the use of a label helping to ‘de-situate’ the relational structure from the initial context (Gentner, 2003). Understanding of this structure underpins subsequent use of the term, usage which is not based on expectations of perceptual similarity.

Use of a linguistic label also helps preserve the abstraction in memory and renders it more accessible for future use (Gentner calls this ‘reification’). “The most obvious instance of this increased accessibility is that a named pattern can be re-invoked by using the name” (Gentner & Christie, 2010). This may seem a trivial point, but it is highly significant. As I will briefly discuss in section 7, there is evidence that without the use of linguistic labels, these relational patterns are not as accessible to cognition as more obvious perceptual similarities. Language is, *par excellence*, a device in which labels are re-used across multiple items and situations that may show little or nothing by way of perceptual similarity. Rather than simply being a medium through which analogical cognition is recorded, the suggestion here is that language itself plays a constitutive role in facilitating this type of cognition.

5 A hypothesis for word understanding

The evidence supports the following hypothesis: at least for many words, for a word to be understood is for the word to trigger a focus on a structural relation. This focus underpins the matching capacity that is pervasively illustrated in our use of a word. The sheer complexity of word use will always challenge any simply drawn suggestion. But the hypothesis is well motivated and has various interesting characteristics, which I discuss in the rest of the paper.

The account dissociates the basis for word use from whether or not items/situations share properties, in particular perceptual likenesses (‘surface’ similarities, as Gentner sometimes puts it). This is massively confirmed by how we use words: we effortlessly, and without thinking of the dissimilarities, apply a single term across situations that can differ widely in the particular features that make up those items or situations. We do not think of the differences when we describe an elastic band, an electric device, an internal combustion engine, part of a biological organism, or even a person, as a ‘motor’. Use of the single word ‘motor’ focuses our attention

on a shared structural relation (roughly, something that imparts motion).⁵ We can recognize when a negotiation is occurring, whether that is between a mother and her child, a buyer and seller, between the members of a committee, or between nations. These situations differ in many ways, but they are the same in respect of manifesting a given structure (roughly, a discussion aimed at reaching an agreement). The challenge is not to think up such examples, but to think up examples where word use is based on something like perceptual similarity instead.

Gentner and Kurtz (2005) observe that a relational status is plausibly attributed to a great number of the nouns that we use. They contrast relational categories, whose members ‘cohere on the basis of a core relationship’, with what they call entity categories, whose members ‘share many intrinsic properties’ (which will often be perceptual, see Gentner (2005)).⁶ Words such as ‘tulip’ and ‘camel’, they suggest, name entity categories. Gentner and Kurtz do not, though, distinguish the weaker claim that some words are principally used of items that happen to display overall likeness from the stronger and more relevant claim that the understanding of some words is constituted in terms of a grasp of a likeness that must hold between members of the category. Just because members of a category happen to be perceptually similar, for example, it does not follow that this similarity plays a constitutive role in our understanding of the word for that category. For example, books used to be far more uniform in appearance and methods of construction than they are today; but modern production techniques, and in particular the shift to electronic books, show that these uniformities were not actually determinative for how the word ‘book’ was understood. Further, some of Gentner and Kurtz’s proposed examples of terms for entity categories, such as ‘household appliance’ and ‘musical instrument’, are clearly inadequate: neither are used to name items that need share overall likeness, and both have plausible structural explanations (pianos and triangles both play a role in enabling music to be produced; kettles and irons are both manufactured products used for particular household operations). The important point, though, independently of what we think of the proposed entity categories, is that for at least very many words in regular use a structural explanation is plausible.

This claim is consistent with the observation that, in practice, surface similarities can be used to help guide our use of words (in particular for categories whose members do display general overall likeness). Many studies have shown the importance of shape as a cue for choosing the appropriate item to label with a given noun.⁷ This is adequately explained, on the current hypothesis, as due to a correlation that often holds between shape and the relations into which items enter. This is particularly clear with the case of function. If it looks like a hammer or a bicycle, an item can probably play the roles of a hammer or bicycle. But this does not show that the understanding of the label is constituted in terms of the shape of an item. Abilities to sort, and make inferences concerning, a given category go beyond the input from word understanding. The input from word understanding provides (by hypothesis) only a limited basis for sorting and inferential abilities; these restricted abilities will draw on consequences that follow from the specific structural feature that a word picks up on. While I may use shape to identify that an item is a motor, this ability does not arise from my understanding of the word

⁵ I take this rough description from the OED. Lexicographers have often provided a wealth of suggestions as to how to characterize structural relations.

⁶ Use of the qualifier ‘intrinsic’ is meant to indicate a contrast with ‘relational’, but it is an unfortunate term to use. It is intrinsic to a barrier to be an item that prevents something from moving. I take the contrast to be between the claim that members are alike in some overall fashion and the claim that membership is based on a structural relation. As I point out, these are not mutually exclusive.

⁷ On the shape-bias effect, see for example the special issue *Developmental Science*, 11(2), 2008.

‘motor’, which, on the current proposal, is limited to the relation of (roughly) something that imparts motion. Knowledge of shape, and other similar types of knowledge, belong to general knowledge about motors, and are not part of linguistic understanding.⁸

Use of the word ‘mouse’ for the computer accessory was initially motivated, according to its inventor Doug Englebart, on grounds of similarity in appearance to a mouse. Interestingly, this is a case of a word acquiring a new meaning (homonymy) rather than an extension based on what the word ‘mouse’ originally meant. Further, while shape was the motivating factor for using this label, shape is not the basis for further use of the computer word ‘mouse’. Computer mice often no longer have tails (the wire), but that is not important for the extension of the word, which is based on a relation that the item enters into between a user and a computer.

By providing a limited scope to word understanding, the account helps to explain why children can learn so many words with such apparent ease and rapidity. Word understanding is based on a quite specific appreciation of a particular structure, one that is available even to young children. By hypothesis, a child can have as good a linguistic understanding of the word ‘motor’ as an expert mechanic does.

Before considering further characteristics of the approach, I will finish this section with consideration of what is probably going to be regarded as a key difficulty for the account.

Wittgenstein (1953, §66) famously suggested that for at least some words, such as ‘Spiel’ (game, that which is played)⁹, there is no common factor that explains the use of the word across the range of instances to which it is applied. Wittgenstein’s example is indeed highly suggestive, and it is not an easy task to think up an account that accurately reflects or predicts the range of situations to which we apply the term and also the range of situations to which we do not apply the term. In English usage, ‘game’ applies to football and cricket, but not (it seems) to athletics events; it applies to Trivial Pursuits but perhaps not to pub quizzes; it may apply to some types of training exercise (War games) but not to others. A particularly useful feature of the example is that there is no compelling reason to appeal either to homonymy, or even to polysemy,¹⁰ when explaining at least the majority of our uses of ‘game’. So, according to the style of account that I am proposing, there ought to be some relatively simple relation that is characteristic of the instances to which we apply the term ‘game’.

While Wittgenstein’s example does show the difficulty of the task, he does not show that it is futile to look for such an account. Three points can be made.

First, he doesn’t actually show that ‘game’ has no unifying explanation for application. He merely lists a range of suggestions as to what the commonality (or commonalities) might be—such as being amusing, or involving winning and losing—and after a brief attempt concludes that no such commonality is to be found. Arguably, Wittgenstein was looking for some bundle of individual features that might hold of all and only games, rather than to a structural relation that is characteristic of games as such.¹¹

Second, part of Wittgenstein’s own positive proposal for how terms are applied (insofar as it is such a proposal) appeals to ‘family resemblances’. It is significant that this idea was

⁸ In agreement with Fodor (1998; 2008), the account indicates that abilities are not constitutive of concepts (in my terms, word understandings), though, unlike Fodor, some abilities follow directly from the limited sphere of word understanding. Fodor empties the lexicon of any such sphere of understanding, and reduces the connection between lexical concept possession and abilities to an association that holds by virtue of lexical concepts acting as pointers to mental files, in which all the epistemically relevant material is found.

⁹ See Hoyningen-Huens (2015). ‘Spiel’ has a wider range of application than the English word ‘game’.

¹⁰ Polysemy does not have a clear definition, but my point is that a vast number of uses of ‘game’ do not need to be qualified as showing variations of sense, even if the games it applies to vary much in characteristics.

¹¹ Hoyningen-Huens (2015) suggests that Wittgenstein wasn’t looking at the appropriate level of abstraction, and he thinks that an account of ‘Spiel’ can in fact be given.

illustrated by a largely perceptual example (the build and features of persons in a family, (1953, §67)), and the focus is on comparing some appropriate number of features of different items. As we have seen, there is compelling evidence that looking to structural relations is important for at least many words, and this is a focus that removes attention from making a comparison between the features of items.

Third, the context of Wittgenstein's discussion is that of showing that we cannot explain word use by appeal to a putative meaning that somehow determines, by itself, the items or situations of which the word can be truly used. This is an important philosophical point, but it is not one that the appeal to analogical cognition need be seen as inconsistent with. The focus with analogical cognition is on a commonality of understanding that can be specified at the linguistic level; the focus is not on attempting to delineate sets of defining marks that would have the role of dividing up items into fixed extensions. Usage indicates that a marble can be correctly described as a barrier, and correctly described as not a barrier. This is predicted by the analogical approach: whether the structural relation is realized will depend on what enters into the relation with the marble.¹²

Even if appeal to analogical cognition is broadly correct, there are reasons for thinking that it may often prove difficult to pin down, in explicit terms, our understanding of a word. In practice, word use is influenced by how neighbouring words are used. Language aids cognition not only by providing individual labels, but by providing a multitude of related labels, creating a complex multi-dependent framework of word usage that both induces a fineness of grain in our understanding of a given word and also leads to contingencies of usage that may be beyond the explanatory reach of theory. Use of terms such as 'athletics' (for track and field events), or 'swimming' (for pool events), seems to block the use of 'game' for these events, whereas there is no such block for football, cricket, and hockey. On the other hand, we also speak of the Olympic Games (a point noted by Hoyningen-Huens (2015)), which illustrates the contingencies that can be displayed in word use. In terms of structural relations between user and computer, a track pad on a laptop need hardly be distinguished from a mouse. But we do make a distinction, and in this case it seems plausible to suppose that appreciation of the difference in how the items are manipulated enters into our understanding of 'mouse' and 'track pad'.

Related words, therefore, may lead to usage of a given word being influenced by a potentially complex array of factors. This is consistent, though, with the claim that looking to structural relations provides an explanation for the primary basis on which word understanding is founded.

6 Understanding and context sensitivity

One response to the phenomenon of context sensitivity has been to suggest that word meanings are 'underspecified'. That is to say, what a word means does not itself fully determine the content that a use of the word may come to convey. Rather, the word meaning provides some sort of starting point, from which a particular interpretation may be constructed.

An approach to word understanding based on analogical cognition has an element of underspecification, and in this respect is well able to deal with at least some types of context sensitivity. Because the emphasis is on a structural relation, and not the specific features of the items that manifest this relation, nothing is determined in the account as to how a particular

¹² The point is anticipated in Plato's discussion of Forms. Forms can be understood in terms of structure (see Prior (1983)), and this helps explain the phenomenon of the 'compresence of opposites', in which a predicate both applies, and does not apply, to a particular item (see Pritchard (forthcoming)).

item need be instantiated in a given instance. Understanding the word ‘obstacle’ or ‘motor’ does not involve any determination for the type of obstacle or motor it is or for any of their features—beyond that they will help satisfy a given relational role. The way in which something is an obstacle, for example, or the perspective from which it is an obstacle, is not specified at all.

So, on hearing the utterance ‘The motor pulled the obstacle out of the path’, our linguistic understanding presents us with a claim in which an item that causes movement stands in the pulling relation to an item that plays a preventing role on a terrain that enables movement between points (a rough characterization of ‘path’).

But if so much is left unspecified, can we explain the phenomenology, in which meanings appear to be directly apprehended in the hearing of words? The answer, I believe, is yes, and relates both to the fact that the account does suggest a *specific* meaning (as opposed to some vague pointer to conceptual space) and on the observation that understanding is not an all or nothing affair—language users often appear to make do with merely ‘good enough’ understandings. We need to allow for the possibility that the kind of understanding that is immediately experienced on hearing words has no more than the specificity contained in the semantics of the words used themselves, even if further levels of representation are always possible.

Of particular interest in this respect is evidence that language users do not seek for fully specified interpretations of utterances, but make do with ‘good enough’ representations for the purposes at hand (see Ferreira, Bailey, and Ferraro (2002); Stanford and Surt (2002), from which the following examples are taken). For example, consider the pronouns in the following sentences:

- (1) Mary bought a brand new Hitachi radio.
- (2) It was in Selfridge’s window.
- (3) Later, when Joan saw it, she too decided it would be a good purchase.

The referent of *it* in (2) is not fully specified. Did Mary buy the particular radio that was physically in the window or just one of the same type? The interpretation of *it* in (3) offers further possibilities: perhaps Mary’s radio, perhaps the radio in the window, perhaps some other window or collection altogether. The point of the example is that there is no requirement for a language user to resolve these details. Processing might not and need not occur to such a fine grain. Similarly, scope ambiguities need not always be resolved in order for the import of an utterance to be clear. From ‘At least one problem preoccupies every politician’, along with ‘John is a politician’, one can infer that ‘At least one problem preoccupies John’. This inference can be made without committing to one or other scope ordering of the original sentence.

These examples suggest that language users often leave details unresolved when there is no requirement for making a decision. The point can be extended beyond reference and quantifier scope resolution. Frisson (2009) surveys experimental work that appears to show a similar effect at work with polysemy. For example, the word ‘school’ can be used in ways that would invite, if required, subtly different dimensions of implementation. Jocelyn walked to the school (emphasis on the physical place); her mother talked to the school (this is manifested by talking to representative members of staff); the school won the match (an action that is implemented by members of the school); and so on. Frisson speaks of these different implementations as interpretations or senses that the word ‘school’ can take on in different uses.

Experimental evidence shows that these interpretations are not accessed in the same way that the differing meanings of homonymous words are accessed. With homonyms (such as ‘coach’: vehicle, trainer), experimental work shows that all the possible meanings are initially

activated, with access largely determined by the frequency of the individual meaning (a low frequency meaning will be activated but might not reach the threshold required for access). Context can reorder access by making a subordinate meaning more accessible. Resolution of one or other readings typically occurs without delay. For example, on hearing ‘The coach argued with the players’, both meanings of ‘coach’ will initially be activated but the vehicle meaning will quickly be dropped. Sometimes the wrong choice will be made, and this delays interpretation further down the line. We might have: ‘The coach argued with the players; it refused to start when they were jumping around inside’.

By contrast, with the ‘school’ examples (and other similar examples), experimental evidence shows that relative frequency of the various possible senses has no effect and no extra processing effort was found when a less frequent sense was intended. While a possible account is that all possible senses are initially activated (independent of frequency), Frisson and Pickering (1999) argue that there is good reason to rule this out, on grounds of being uneconomical (words often have lots of different senses) and of there being no evidence for competition between the senses (which would be expected if all senses were activated and would lead to longer reading times).¹³ Further, if all senses were activated, there would be a big problem with sense selection (far more so, given the number of senses, than with typical homonymy): if the preceding context does not indicate which the appropriate sense is, we would predict that revisions would frequently be needed if a particular sense was in fact selected.

Frisson and Pickering (1999) put forward, therefore, an alternative hypothesis: what is initially activated is a single, semantically underspecified meaning. This meaning is meant to encompass all the semantically related interpretations of a word.

This abstract meaning is supposed to be the same for all established senses of a word, that is, the same underspecified meaning encompasses all semantically related interpretations of a word that are known to the reader. (Frisson, 2009)

Frisson combines this suggestion with the claim that language users make do with ‘good enough’ interpretations, and suggests that language users will often not bother to resolve the interpretation of ‘school’ beyond this initial underspecified abstract meaning. All options are left open, and the actual processing that goes on may be shallow: ‘a GE [good enough] approach seems to be most compatible with the initial activation of rather shallow lexical-semantic information, which is implied in the idea of semantic underspecification’ (Frisson, 2009).

However, while Frisson’s work helps show how the understanding that we immediately pick up from words need not show depth, there are aspects of his approach to underspecification that can be questioned. The ‘school’ type of example is presented in the following way:

- The different possible readings are called ‘interpretations’ or ‘senses’.
- The underspecified meaning is the ‘same’ for all these senses and ‘encompasses’ them all.

This pushes us in the direction of thinking that ‘school’ has a wide array of interpretations (e.g., covering both physical buildings and people), and that all these interpretations are somehow

¹³ On the contrary, lexical decision tasks for words with multiple senses tend to show faster response times than for words with fewer senses.

present in embryo, or otherwise ‘encompassed’, in a single meaning of ‘school’. This single meaning does not itself fully specify which interpretation will be relevant on an occasion.

Frisson himself is not committed to what form this underspecified meaning will take (see Frisson (2009), 121f.).¹⁴ One option is that the meaning underspecifies in the sense of providing an over-rich semantic representation, from which some sort of selection is required for achieving the interpretation required for a particular occasion of use (Pustejovsky, 1995) presents an ‘over-rich’ approach to lexical semantic representation). An over-rich representation contains more information than is required for a given interpretation, and does not in itself specify which aspect of that over-rich representation will be relevant in a given use. A second option is to go in the opposite direction and to think of a highly schematic or impoverished type of meaning that contains some core commonality that holds across all interpretations. Rather than requiring a process of selection, a schematic meaning needs to be ‘filled out’ if we are to get to a specific interpretation. We might think of a single skeleton that can be fleshed out in different ways on different occasions.

While Frisson is non-committal, I wish to focus on the second (‘impoverished’) option. In part this is because I think that simpler, more skeletal, accounts of lexical knowledge are more plausible than accounts that propose more complex types of representation. But also, it is difficult to understand how a rich semantic specification fits Frisson’s general approach. He writes that what is initially accessed in interpretation is “not a full-fledged, specific interpretation”, but rather something that “will serve as the starting point for a more detailed specific interpretation” (Frisson 2009, p. 122). Further, in line with the ‘good enough’ emphasis, what is initially activated is “rather shallow lexical-semantic information” (Frisson 2009, p. 123). It is not clear how a rich lexical entry can be characterized as requiring ‘more detail’ or as being ‘rather shallow’.

However, it is also problematic to understand how an impoverished meaning can in fact play the kind of role that is required by the example that Frisson provides. First, there is no clear sense as to what this impoverished representation that is shared across all interpretations will look like. With the term ‘school’, we seem to be positing an abstraction that captures a commonality shared by physical buildings, sports teams, administrators, institutions, and other possible nuances of interpretation. It is hard to imagine what this commonality could be.

Carston (2016), criticizing ‘schema’ approaches in general, raises further concerns. One concern is that the proposed abstraction doesn’t explain the various interpretations but is merely a result of those interpretations. If a new interpretation arises, the abstracted meaning will have to change in some way to accommodate it.¹⁵ Explanation starts with the interpretations, and the proposed meaning is nothing but a by-product of this—it is an explanatory idle wheel. Further, the proposed abstraction does not seem to have any relevance to what children learn when they learn words. A child does not seem to learn the abstraction, but would appear to start with one or other of the various interpretations.

The account of word meaning in terms of analogical cognition provides a response to these concerns, as well as challenging aspects of Frisson’s discussion of the ‘school’ example. Rather than saying that the meaning of ‘school’ is underspecified, or that it encompasses all the interpretations, the claim is that the meaning of ‘school’ is completely specific. That is to say, speakers of English share a specific understanding, an understanding that grasps the structural relation that is describable roughly as ‘institution that gives teaching’. Nothing

¹⁴ Thanks to Josie Bowerman and Robyn Carston for pointing this out.

¹⁵ This point is noted by Frisson and Pickering themselves: ‘The underspecified meaning is only applicable for established senses. Because the underspecified meaning is an abstraction over the features of specific senses, a novel interpretation of a word cannot be captured by the underspecified meaning’ (Frisson and Pickering, 2001).

further is specified by the understanding; but in its own terms this understanding is specific, and by hypothesis this specific notion of school occurs as a clearly articulated constituent of thought that is triggered when the word is used.

More detailed consideration of what the various scenarios involve might lead to a focus on specific aspects of the institution; but this does not undermine the claim that all that the word is responsible for is the structural relation that is triggered in thought. More detailed consideration might lead us to conclude that Jocelyn's mother spoke to a member of staff. But this doesn't need to be encompassed by the meaning of 'school', in the sense of requiring a meaning that is somehow general across this and other implementations. Sainsbury (2001) argues that there is a tendency for the claim that there are distinct ways in which one and the same meaning could be true to be confused for the claim that there are distinct meanings. The meaning of the word need not be responsible for any of the (endless) possible details concerning how the scenarios are implemented. In particular, notice that we may sometimes deliberately speak in terms (for example) of having contacted 'the school' when we do not wish to focus on the specific person through whom this contact was made. Our point may simply be one of indicating that the institution was informed. And we may just have used a loudspeaker. On the current hypothesis, the word 'school' is not itself responsible for any particular implementation.

This revised approach indicates that the 'good enough' reading is not to be understood in terms of a mysterious abstraction, but in terms of a quite straightforward understanding that we are talking about a school, viz., an institution that provides education. This will be good enough for many purposes, and if more information is needed it will be elicited (e.g., 'Who did you speak to?').

We can, therefore, make a distinction between plausible and implausible abstractions. The plausible abstraction relates to the kind of structural relation that Gentner has studied. The implausible abstraction is the type of abstraction that is meant to provide a core commonality that embraces the potentially many different accounts of precisely which part of the institution was directly relevant in a given instance. Notice in particular that the range of interpretations that Frisson has in mind relate to particular features of schools, hence *not* to the overall relational perspective. The focus is more on a kind of general knowledge about schools (such as who one speaks to when one contacts a school), but it is just this kind of general knowledge that the account based on analogical cognition indicates is not relevant to word understanding.

The account provides a response to Carston's idle wheel criticism. Rather than being a putative abstraction that results (perhaps implausibly) from a wide and possibly growing array of interpretations, the proposal is for a psychologically realistic level of understanding that plays a central explanatory role in the judgements that are in play when words are used. Carston (2016) supports her idle wheel criticism by arguing that when children first learn a word, what they learn are precisely the senses that are grasped in communication, not the proposed schematic abstractions. But we have seen that there is good evidence, even from a very young age, that word use induces a focus on structural relations—these being plausible abstractions—and that subsequent use of a word triggers the corresponding understanding. Further, the account is not suggesting that word understanding starts at a level that is abstracted away from normal everyday thought. On the contrary, the triggered element of thought can be exactly what is being conveyed. In Carston's terms, the account is fully conceptual.

By appealing to specific, highly accessible, and psychologically realistic understandings of individual words, we have a plausible explanation for the immediacy with which understandings are triggered on hearing language. Details of the features of items or situations are left unspecified, but language users make do with 'good enough' representations. Not all

aspects of interpretation, such as particular implementations for the school scenarios, need to be accounted for in a theory of word understanding.¹⁶

7 A contribution to cognition

The role of language is sometimes envisaged as a medium through which the conceptual realm can be accessed and communicated. A. Clark has described this as the ‘translation view of language’, in which encountered language ‘merely serves to activate the complexes of internal states or representations that are the real cognitive workhorses’ (Clark, 2006). Once the translation has been accomplished, the linguistic form can be thrown away, leaving just the ‘essence’, a meaning that has been ‘fully extracted and rendered in some alternative inner format’.

Gentner’s work supports an alternative view: words are tools that contribute aspects of cognition that are either not found or are under-represented at the more basic biological level of thought: ‘One function of language may be to augment natural modes of cognition with an alternative representational scheme that permits abstract cognition’ (Gentner, 2003). In particular, words may contribute by providing memorizable and context *insensitive* building blocks (as Clark (2006) puts it), and this is possible by virtue of the shift to analogical cognition that the use of linguistic form facilitates. From this perspective, to talk of words as ‘pointers’ to complex arrays of conceptual material is misguided. Rather than pointing somewhere, words trigger specific structural understandings, which help take our thought to a different level.

Much experimental work shows that if people, in a given situation, are asked to think of situations that are similar to the one they are in, their decision ‘is typically driven largely by surface similarities, such as similar objects and contexts, rather than by similarities in relational structure’ (Gentner & Smith, 2013). In other words, the kind of similarity that promotes memory retrieval tends to contrast with the kind of retrieval that is facilitated by language.

The claim that ... language aids analogical retrieval is important, because analogical retrieval is generally quite poor. People routinely fail to be reminded of past experiences that are relationally similar to current experiences.

(Gentner & Smith, 2013)

This suggests that judgements of similarity, when not guided by use of a given word, will not group situations in the way a typical word of language would. The word ‘open’ groups together widely disparate items and situations that do not share surface similarities (open books, bottles, rooms...). By contrast, a judgement of similarity, not guided by the word ‘open’, would, in general, be expected to group an open bottle with a surface similar item such as a vase, an open book with a neat sheaf of papers, and so on. This is a problem for accounts of words, such as Recanati’s sketch of ‘meaning eliminativism’ (2004), that appeal to judgements of similarity—between a situation in which a word has already been used and a novel situation—as the basis

¹⁶ Nathan Klinedinst has pointed out to me that if we take ‘school’ to have a single meaning of the type I suggest, we are implicitly loosening selectional restrictions for other predicates. For example, we say ‘The school is made of wood’, though an institution that provides teaching, being partially abstract, cannot itself be the sort of thing that is wooden. Further, we have the problem of explaining why certain inferences seem to be blocked. From ‘The school is made of wood’, and ‘The school won the match’, we don’t want to infer ‘Something made of wood won the match’. Perhaps a partial explanation is that when further details or inferences are elicited, aspects of the ‘good enough’ interpretation can get challenged and thereby need to be modified. But these modifications are only required when circumstances demand it.

for explaining when a word is re-used. Arguably, such accounts will make the wrong predictions about word usage. Rather than a term being used of situations that we are naturally inclined to regard as similar, it seems as if the term itself plays a central role in facilitating and directing the relational focus that is manifested in the typical array of uses found with words.

8 A way forward?

In summary, human cognition arises in part ‘from learned symbol systems that facilitate the apprehension of relational structure’ (Gentner, 2003). Language provides a structure that is immediately apprehended and that interacts with our cognitive abilities, resulting in the kind of thought that humans are capable of. Each individual word plays a small part, contributing a specific piece of structure.

It remains to be seen whether this approach is consistent with the various phenomena (grouped under headings such as polysemy, metonymy, metaphor) that have been studied by theorists. By looking to a basic underlying understanding of a word, the hypothesis has considered only a limited, albeit important, part of the puzzle. In practice, words appear together in syntactic structures, and this contributes another major component. But in any explanation, it is important to have some idea of the division of labour (what different elements of language can and cannot contribute), and of the kinds of representations that need explaining. An approach to words based on analogical cognition helps to provide some insight into these issues.

References

- Bach, T. (2012). Analogical cognition: applications in epistemology and the philosophy of mind and language. *Philosophy Compass*, 7(5), 348–360.
- Borer, H. (2005). *Structuring sense volume 1: In name only*. Oxford: Oxford University Press.
- Borg, E. (2012). *Pursuing meaning*. Oxford: Oxford University Press.
- Carston, R. (2002). *Thoughts and utterances*. Oxford: Blackwell.
- Carston, R. (2012). Word meaning and concept expressed. *The Linguistic Review*, 29(4), 607–623.
- Carston, R. (2013). Word meaning, what is said and explicature. In C. Penco & F. Domaneschi (Eds.), *What is said and what is not*. CSLI Publications: Stanford, CA, USA.
- Carston, R. (2016). The heterogeneity of procedural meaning. *Lingua*, 175–176.
- Christie, S. & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science*, 38, 383–397.
- Clark, A. (2006). Material symbols. *Philosophical Psychology*, 19(3), 291–307.
- Ferreira, F., Bailey, K., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11, 11–15.
- Fodor, J. (1998). *Concepts*. Oxford: Clarendon Press.
- Fodor, J. (2008). *LOT 2*. Oxford: Oxford University Press.
- Fricker, E. (2003). Understanding and knowledge of what is said. In A. Barber (Ed.) *Epistemology of language* (pp. 325–366). Oxford: Oxford University Press.
- Frisson, S. (2009). Semantic underspecification in language processing. *Language and Linguistics Compass*, 3(1), 111–127.
- Frisson, S. & Pickering, M. (1999). The processing of metonymy: evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(6), 1366–1383.
- Gentner, D. (2003). Why we’re so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: advances in the study of language and thought* (pp. 195–235). Cambridge, MA: MIT Press.
- Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time* (pp. 245–275). Hillsdale, NJ: Erlbaum.
- Gentner, D. (2010). Bootstrapping the mind: analogical processes and symbol systems. *Cognitive Science*, 34, 752–775.

- Gentner, D. & Christie, S. (2010). Mutual bootstrapping between language and analogical processing. *Language and Cognition*, 2(2), 261–283.
- Gentner, D. & Kurtz, K. (2005). Relational categories. In W. Ahn, R. Goldstone, B. Love, A. Markman, & P. Wolff (Eds.), *Categorization inside and outside the laboratory* (pp. 151–175). Washington, DC: APA.
- Gentner, D. & Markman, A. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52(1), 45–56.
- Gentner, D. & Smith, L. (2013). Analogical learning and reasoning. In D. Reisberg (Ed.), *The Oxford Handbook of Cognitive Psychology* (pp. 668–681). New York: Oxford University Press.
- Hoyningen-Huene, P. (2015). A note on the concept of game (or rather *Spiel*). In G. Betz, D. Koppelberg, D. Löwenstein, & A. Wehofsits (Eds.), *Weiter denken – über Philosophie, Wissenschaft und Religion* (pp. 205–210). Berlin: de Gruyter.
- Keil, F. & Kominsky, J. (2015). Grounding concepts. In E. Margolis & S. Laurence (Eds.), *The conceptual mind* (pp. 677–692). MIT Press.
- Loewenstein, J. & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, 50, 315–353.
- Penn, D., Holyoak, K., & Povinelli, D. (2008). Darwin’s mistake: explaining the discontinuity between human and nonhuman minds. *Brain and Behavioral Sciences*, 31, 109–178.
- Pettit, D. (2010). On the epistemology and psychology of speech comprehension. *The Baltic International Yearbook of Cognition, Logic and Communication* (Vol. 5, *Meaning, Understanding and Knowledge*, pp. 1–43).
- Pietroski, P. (2005). Meaning before truth. In G. Preyer & G. Peter (Eds.), *Contextualism in philosophy* (pp. 255–302). Oxford: Clarendon Press.
- Prior, W. (1983). The concept of paradigm in Plato’s theory of Forms. *Apeiron*, 17(1), 33–42.
- Pritchard, T. (forthcoming). Is there something in common? Forms and the theory of word meaning. *European Journal of Philosophy*.
- Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.
- Recanati, F. (2004). *Literal meaning*. Cambridge University Press.
- Rayo, A. (2014). A plea for semantic localism. *Nous*, 47(4), 647–679.
- Sainsbury, M. (2001). Two ways to smoke a cigarette. *Ratio*, 14(4), 386–406.
- Sanford, A. & Sturt, P. (2002). Depth of processing in language comprehension: not noticing the evidence. *Trends in Cognitive Sciences*, 6(9), 382–386.
- Sperber, D. & Wilson, D. (1998). The mapping between the mental and the public lexicon. In P. Carruthers & J. Boucher (Eds.), *Language and thought: interdisciplinary themes* (pp. 184–200). Cambridge: CUP.
- Travis, C. (2000). *Unshadowed thought*. Cambridge, MA: Harvard University Press.
- Wilson, D. & Carston, R. (2007). A unitary approach to lexical pragmatics: relevance, inference, and ad hoc concepts. In N. Burton-Roberts (Ed.), *Pragmatics* (pp. 230–259). Palgrave: London.
- Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.

A Theta-Gamma Neural Code for Feature Set Composition with Phase-Entrained Delta Nestings*

Elliot Murphy

Abstract

Continuing a recent line of research into the functional role of neural oscillations in language and cognition, I propose that a central computational system (implemented via gamma, theta and delta interactions) interfaces with external interpretive and attentional systems (implemented via beta and alpha modulations) to derive certain core features of phrasal construction. These interactions are achieved through the coupling of the phase and amplitude of distinct rhythms. A theta-gamma code generates the initial feature-set while thalamic and occipital alpha is used to inhibit irrelevant information sources. This code is itself embedded within left-cortical delta rhythms while an increasing beta rhythm, which finds its source within a cortico-basal ganglia-thalamo-cortical loop, is used to maintain the existing cognitive set in memory. Neuroethological arguments are presented in favour of the species-specific nature of this neural code, and it is argued that the unique hierarchies exhibited in natural language emerge from this human-specific rhythmic profile.

Keywords: Neural oscillations, theta, gamma, delta, labeling hypothesis, oscillome

1 Introduction

In what follows I will adopt the Labeling Hypothesis put forward in Murphy (2015a), that labeling – and not Merge – constitutes the exclusive content of ‘narrow syntax’ and hence should be the focus of any evolutionary scenario and neuroethological investigation. This is in contrast to standard ‘Merge-centric’ evolutionary hypotheses (Berwick & Chomsky, 2016; Hauser, Yang, Berwick, Tattersall, Ryan, Watumull, Chomsky, & Lewontin 2014; Mukherji, 2010; Nóbrega & Miyagawa, 2015, and much other work). Boeckx and Theofanopoulou (2015) noted that the concept of labeling in Murphy (2015a) was not formulated at a fine enough level of granularity to permit linking hypotheses to be drawn up between linguistic theory and neurobiology. This was later achieved in Murphy (2015b), in which the elementary syntactic sub-operations of set-formation, labeling, and cyclic transfer were proposed to arise from the well-preserved hierarchy of mammalian brain rhythms and their interactions, following seminal work by Giraud and Poeppel (2012) on the dynamics of phonological computation. Through outlining a program of Dynamic Cognomics, a number of hypotheses were presented concerning the neurobiological basis of phrase structure building, feature valuation, and other properties of language. In brief, this research program attempts to attribute to neural oscillations (at the level of the ‘oscillome’, or oscillatory brain activity ascribed certain computational roles) particular computational and representational properties which were previously attributed only at the level of the ‘cognome’, the set of computational operations the human nervous system could execute (Poeppel, 2012).

Oscillations enable the construction of coherently organised neuronal assemblies through establishing transitory temporal correlations, and are increasingly being implicated in

*My thanks go to Daniel Bush, Karl Friston, John O’Keefe, Andrea E. Martin and the participants of the Eighth Annual Meeting of the Society for the Neurobiology of Language at the UCL Institute of Education and the Cell Symposia: Big Questions in Neuroscience conference in San Diego for helpful discussion in developing the ideas presented here. This work was supported by an ESRC scholarship (1474910).

numerous cognitive processes (see, for instance, Kaplan, Adhikari, Hindriks, Mantini, Murayama, Logothetis, & Deco (2016)). They are typically measured through electroencephalography (EEG) and magnetoencephalography (MEG) and reflect synchronised fluctuations in neuronal excitability and are grouped by frequency, and include delta (δ : ~0.5-4Hz), theta (θ : ~4-8Hz), alpha (α : ~8-12Hz), beta (β : ~12-30Hz) and gamma (γ : ~30-100Hz). These are generated by cortical and subcortical structures and form hierarchical, nested relationships such that slower rhythms can modulate the amplitude of faster rhythms.

Contrary to what might be expected from a program declaring interest in mind-brain commensurability, Dynamic Cognomics (which I will here refer to simply as oscillomics) is purely an oscillation-based methodology set against a backdrop of Marrian and evo-devo assumptions, with no stipulations about what form linguistics-neuroscience linking hypotheses should necessarily take. Oscillomics has the potential to reduce what Tomalin (2006, p. 188) calls the ‘attendant logico-philosophico-mathematical baggage’ of generative grammar by re-embedding linguistic combinatorics within a framework of brain dynamics.

Language itself is argued by some to be the origin of complex thought and imagination. Tattersall (2016) writes:

This symbolic capacity of ours resonates in every realm of our experience, and in combination with the complex ancestral foundation of intuitive intelligence and association-making on which it is superimposed ... it is the fount not only of our imagination and creativity, but of all those other cognitive peculiarities that set us apart from even our closest living relatives in nature.

Throughout the 20th century and into the 21st, artificial intelligence research has repeatedly moved in and out of synch with biology, at one point being closely aligned with current developments only to diverge slightly at another. Linguistics, on the other hand, has often claimed to be in synch with biology, but has never in fact achieved this goal. The relative lack of interest in biology exhibited by many linguists becomes even more curious when we see how interested in language many biologists are: While perhaps novel to linguists and cognitive scientists, the potential existence of a relationship between language and neural oscillations was discussed as early as 1978 in O’Keefe and Nadal’s monumental work *The Hippocampus as a Cognitive Map* (O’Keefe & Nadal, 1978). Here, the authors discuss Chomsky and Jackendoff’s work on syntax and the semantic representation of space, attempting to draw links between how the hippocampus appears to represent vectors and how the language system appears to represent space. O’Keefe and Nadal attempted to show how their theory of cognitive maps could account for certain properties of linguistic ‘deep structure’ and semantic long term memory. In effect, the present contribution is an attempt to do the opposite of O’Keefe and Nadal, broadening both the neurophysiological and computational landscape that these authors set down and discussing a wide number of cortical and subcortical structures and how they might implement elementary linguistic operations.

Computational models of syntactic, semantic and phonological knowledge are likely not going to be reduced fully to neural tissue, but George Box’s famous saying that ‘all models are wrong but some are useful’ should remind us that the ultimately ‘incorrect’ nature of neurocomputational models can at least provide a useful function in directing future research. The study of language should take advantage of whatever sciences could enhance its hypotheses and methodologies. In the Enlightenment period, developments in what would today be termed philosophy of mind contributed to an understanding of some fundamental linguistic concepts like personhood. In the 1950s, developments in meta-mathematics contributed to an understanding of language’s grammatical architecture. Today, with the

proliferation of the brain sciences, it is likely that other features of language can now be explored.

Yet this proliferation can come with certain risks, often not acknowledged by researchers. Jonas and Kording (2016) used the methods of neuroscience to try and understand the MOS 6502 microchip (the processor in, amongst other things, the Commodore 64), which contains 3510 transistors as is able to run only the most primitive, vintage video games. Their results should generate a fair degree of humility: They discovered only that the chip has a clock and is able to read and write to memory. Nothing else was uncovered about it via the standard methods of neuroscience ('lesioning' transistors, analysing individual transistors and local field potentials, performing Granger causality analysis, and so forth). Jonas and Kording discovered general-purpose operations which transistors could perform, but they found no 'Donkey Kong transistor' or 'Space Invaders transistor', i.e. transistors essential and exclusive to a given game. And while the chip is purely deterministic, neurons can exhibit random behaviour. The neurolinguistic lessons are clear, with Jonas and Kording's work serving as a further motivation for abandoning classical 'cartographic' models of language comprehension, which propose one-to-one mappings between brain regions and (often fairly complex) cognitive operations, fixating on Broca's and Wernicke's areas. We are often told by neurolinguists that certain linguistic operations 'take place' at a given region, or are 'interpreted' along a particular pathway, yet the story of what exactly the brain is doing to derive these localised interpretations is left unanswered.

Even a cursory look at the regions implicated in recent oscillatory studies of language comprehension (see Lam, Schoffelen, Udden, Hulten and Hagoort (2016) for a review) indicates that the classical Broca-Wernicke model is far from adequate. Language comprehension generates multiple, large-scale oscillatory changes in a number of distant regions, and so limiting the 'language regions' to Broca's and Wernicke's areas is unjustified, and the dynamic functioning of the brain at once suggests that a simple mapping between a given region and a particular linguistic representation or operation is not going to be forthcoming. The fact that certain regions implicated in syntactic, semantic and phonological processes are spatially overlapping suggests that a different system of segregation will be required than standard cartographic approaches; namely, frequency-based segregation.

The same criticisms apply to Boeckx's (2016) hypothesis that since both the fronto-parietal and fronto-temporal networks process sequences, the strengthened connections between them found in the human brain 'could allow for the processing of sequences of sequences, of the sort attested in grammatical constructions'. How exactly these strengthened connections could lead to recursion is not explained, and no further details – algorithmic or implementational – are presented.

In what follows I will consequently approach the problem of implementing language in the brain with a number of crucial caveats. The models and hypotheses put forward may not ultimately produce a degree of understanding that Marr or Lazbnick would appreciate, but scientific inquiry needs to have its limits, and being able to recognize these limits does not as a result invalidate a given neuroscientific enterprise. Neurolinguistics has been flooded with data over recent years but it has also seen a remarkable lack of conceptual or theoretical innovation to account for this data. So severe has the flooding been that major neurolinguists continue to apply the intuitive cartographic mindset to newly emerging modes of analysis, with Bastiaansen and Hagoort (2015) coming to the neurobiologically implausible conclusion that γ rhythms are responsible for something as complex as 'semantic unification' simply because they detected γ increases during semantically well-formed structures. Notice that this 'theory' is in fact purely a re-description of the data, and an explanation for why we see γ increases is absent. Brown (2014) summarizes the current situation well in a brief report:

If we really care about the question of how the brain works, we must not delude ourselves into thinking that simply collecting more empirical results will automatically tell us how the brain works any more than measuring the heat coming from computer parts will tell us how the computer works. Instead, our experiments should address the questions of what mechanisms might account for an effect, and how to test and falsify specific mechanistic hypotheses.

The physical sciences place a great deal of emphasis on the importance of theoretical physics, and not just, for instance, experimental particle physics. There is no reason why neuroscientists should not afford the same respect to theoretical neurobiology, yet the drive for experimental innovation is currently by far the dominant force in the field. Filling the gap between the flood of data and neurobiological theory, I will argue, is a particular neurocomputational multiplexing algorithm supported by an understanding of how the brain dynamically operates in real time.

2 Beyond Functional Mapping: A Neural Code for Language

As Jackendoff presciently noted, ‘even if we know *where* a structure is localized in the brain – the sort of information that neural imaging can provide – we do not know *how* the brain initiates the structure’ (2007, p. 13). A neurolinguistic theory is ‘incomplete if it does not offer genuine solutions to the problems of combinatoriality, structural hierarchy, and binding among structures’ (2007, p. 15). An understanding of how the brain operates – and not simply what it looks like under a scanner – is provided by brain dynamics and the study of oscillatory neural activity. These are produced by cortical and subcortical structures and form a hierarchy such that slow rhythms phase-modulate the power of faster rhythms (a relationship explained in Figure 2, below). Slower rhythms synchronize distant brain regions while faster rhythms activate local neuronal assemblies (Yan & Li, 2013; Murphy, 2015b).

Brain dynamics and neurochemistry can constrain the types of operations performed by the nervous system, but they cannot easily reveal *what* operations are performed (Seung, 2012). For this, we need to turn to linguistic theory. In what follows I will present some basic operations required for phrasal construction and interpretation and attempt to re-embed them within a particular oscillomic framework.

2.1 Merge

The set-forming operation Merge constructs a new syntactic object out of two already formed (Chomsky, 1995). Merge(*the,book*) would form {*the, book*}. We can define the operation as follows:

- (1) *Merge*
 Select two lexical items α and β and form the set $\{\alpha,\beta\}$:
 $\mathbf{M}(\alpha,\beta) = \{\alpha,\beta\}$

In Murphy (2015b) it was suggested that since lower frequencies are known to synchronise distant cortical regions, this may represent the oscillomic implementation of Merge. I assumed that thalamic and occipital α embeds cross-cortical γ to generate, or ‘lexicalize’, individual representations before these γ -individuated units were themselves embedded within parahippocampal θ . The decoupling of γ from α would be achieved through the activity of the thalamic reticular nucleus.

Maintaining mentally constructed visual objects in memory (i.e. combining a small number of basic shapes into a larger unit), in contrast to ‘whole’ objects, leads to greater fronto-

parietal θ synchronisation (Ewerdwalbesloh, Palva, Rösler, & Khader 2016), and so this rhythm seems well-suited to the maintenance of complex objects generated in fronto-parietal circuits. In addition, θ is involved in engaging functionally distinct sub-regions of the medial prefrontal cortex, integrating information from the output of different cognitive operations (Mas-Herrero & Marco-Pallarés, 2016); a capacity presumably well-suited to transferring representational information from distinct linguistic sub-systems (e.g. phonology, semantics).

2.2 Labeling

Labeling is the operation selecting which lexical features determine the phrasal category, ensuring that $\{A, N\}$ forms an NP, not an AP. This process leads to the asymmetric hierarchies found in phrase structures and appears to be human-specific. Labeling is also monotonic since the identity of a labeled set is sustained when embedded inside another set. Tomalin (2007, p. 1784) notes in his account of the development of the theory of recursive functions that ‘even if a label-free system is proposed, the essential constructional process remains the same’; procedures of set-formation, property attribution and maintenance are required for the construction of ‘a potentially infinite set of hierarchical structures’. Labeling permits the brain to generate items which can be freely deployed (i.e. maintained, retrieved and recombined with other structures), partly independent of perception.

I will now briefly outline some existing proposals, refining and expanding on some of the central ideas. It was suggested in Murphy (2015b) that labeling arises from the slowing down of the above γ -itemized clusters to β followed by $\{\alpha(\beta)\}$ coupling, involving a cortico-basal ganglia-thalamo-cortical loop. The involvement of this loop would be a result of it being a major source of the brain’s core timing system, with a feature of labeling being rhythmicity (Bartolo, Prado, & Merchant, 2014; Bartolo & Merchant, 2015). However, for reasons expanded on below, it now seems that β more likely nests within left-cortical δ while remaining coupled with parahippocampal θ . This is partly due to δ being implicated in phrasal processing in recent experimental work (e.g. Ding, Melloni, Zhang, Tian, & Poeppel, 2016), but there are other reasons to suppose the importance of $\{\delta\{\theta(\beta)\}\}$ coupling, returned to in Section 2.5. For now, I will assume that the $\{\theta(\gamma)\}$ feature-sets ultimately shift to $\{\delta\{\theta(\beta)\}\}$ phrasal/labeled units. An important question, which must be addressed in future experimental work, is how this shift occurs. There appear to be two options: Either the γ -itemized clusters first slow to β before the $\{\theta(\beta)\}$ complex is nested within δ , or the multiplexed $\{\theta(\gamma)\}$ units are embedded first and the items slow to β at a later point. Fewer β -itemized clusters would be able to phase-lock to θ due to the size of each β cycle (around three β cycles per θ cycle), potentially accounting for the unequal ratio between feature-sets responsible for determining phrasal status and feature-sets which perform other roles with respect to syntactic function (e.g. tense features) and content (e.g. lexical features).

Regardless of which route the language system takes to achieve phrasal nesting, the above model seems to be a particularly robust one, from both conceptual and empirical standpoints.¹ For instance, accruing evidence suggests that β holds objects while γ simply generates them (Martin & Ravel, 2014). Dean, Hagan and Pesaran (2012) demonstrate that β is capable of comparing old and new information from distinct modalities due to its wide temporal windows; in a linguistic context, it would compare recently generated but stored phrasal structures with newly-merged non-phrasal elements such as complements.

¹ Indeed, both routes may be taken under different scenarios. In the case of an ambiguously labeled phrase which could constitute either a DP or an NP, it may be that γ is not slowed to β until after phase-entrainment to δ has taken place when the precise interpretation is fixed.

Alternatively, it may be that only some γ -itemized clusters slow to β , namely those clusters responsible for storing the features determining the phrasal/labeled status of a given δ -entrained set. Generative grammar has long suggested that what determines the label of a given set of lexical items is only a subset of the entire feature complex, and so it is possible that the language system is composed of two simultaneously entrained δ and θ complexes; β labeling features and γ non-labeling features:

LABELING FEATURES: $\{\delta\{\theta(\beta)\}\}$
 NON-LABELING FEATURES: $\{\delta\{\theta(\gamma)\}\}$

The cortico-basal ganglia-thalamo-cortical loop is especially appealing with respect to its potential role in labeling given the findings that different ‘stations’ on the loop can be responsible for information integration from distinct cortical sources: convergence appears to occur at the pallidum (Yelnik, Percheron, & Francois, 1984), subthalamic nucleus (Haynes & Haber, 2013) and thalamus (Theyel, Llano, & Sherman, 2010). These stations also display a degree of functional specificity, and it is likely that different features of linguistic representations (phonological, semantic, and so forth) are integrated into the phrase structure-building process in a procedural manner. The anterior thalamus, for instance, is already being considered a major site of memory formation (Sweeney-Reed, Zaehle, Voges, Schmitt, Buentjen, Kopitzki, Hinrichs, Heinze, Rugg, Knight, & Richardson-Klavehn, 2015), going against standard cortico-centric and hippocampal-centric models. The fact that lesions to Broca’s area do not always result in syntactic impairments or modulations in sentential processing competence (Mohr, Pessin, Finkelstein, Funkenstein, Duncan, & Davis, 1978) also suggests that other, possibly non-cortical circuits are (or at least can be) implicated in phrase structure building.²

2.3 A Multiplexing Algorithm for Feature-Set Composition

In Murphy (2016a) the above model for Merge and labeling was significantly expanded in an attempt to explore the algorithmic properties of the oscillatory interactions invoked to explain the human capacity for phrasal construction and interpretation.³ Phase-aligned oscillatory activity permits multiplexing, or the encoding and decoding of multiple information streams (Akam & Kullmann, 2014) – precisely what is required to bind the representationally distinct syntactic, semantic and phonological streams of information which constitute any given lexical item. Multiplexing allows the reconfiguration of connectivity and information types stored and extracted from a given neural network, and this was applied in Murphy (2016a) to phrasal construction.

Under the present model, parahippocampal regions are the site of initial $\{\theta(\gamma)\}$ coupling. The hippocampus seems particularly suited to facilitate these cross-modular interactions since

² Friederici’s (2016) claim that BA 44 and its dorsal fiber connections to the temporal cortex ‘support the processing of structural hierarchy in humans’ is therefore not incorrect, but is rather simply a piece of a larger system found in the language-ready brain. Indeed, neurolinguists should acknowledge more readily that functional, structural hypotheses regarding language comprehension and production are necessarily limited by technology: fMRI may (and in fact appears to) implicate distinct structures from, for instance, MEG. Relying solely on fMRI to build a cartographic model of the language system will lead to hypotheses ignoring the dynamical nature of the brain’s activity.

³ The model in Murphy (2015b) was referred to as the ‘Basic Label’ model. The expanded version in Murphy (2016a) was termed a multiplexing algorithm and here I will subsume the Basic Label model under this more general term, part of the broader ‘oscillomic’ approach to language advocated throughout.

it has been argued to play a role in integrating ‘what’ and ‘where’ information in the perirhinal cortex through the lateral entorhinal cortex and the postrhinal cortex through the medial entorhinal cortex (Fernández-Ruiz & Oliva, 2016). Lisman and Jensen (2013) suggested that these γ and θ rhythms form a code for representing multiple, ordered items. Since each θ cycle contains four to eight nested γ cycles, different representations (e.g. a series of short-term memory sequences) can be sequentially coordinated. This also fits the revision of working memory from the classic five-nine items to four items (Cowan, 2001). This may constrain the number of lexical features able to be transferred in a given derivational cycle, with this idea being implemented in Murphy (2016a) through the algorithm in Figure 1.

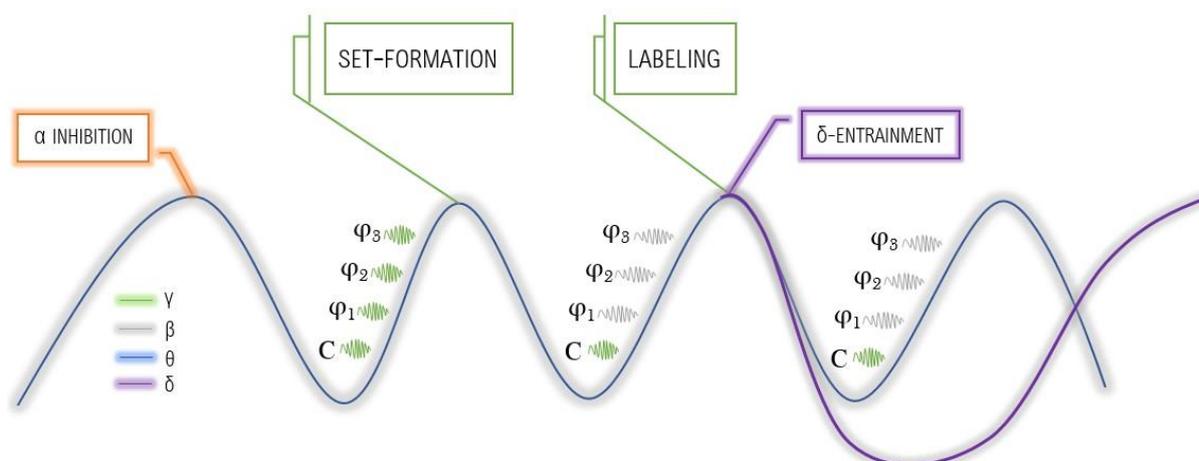


Figure 1: A multiplexing algorithm for feature-set construction. ‘C’ denotes Case feature, ‘ ϕ ’ refers to ϕ -features (Person, Number, Gender). During the encoding of the feature-set, the initial features are represented at the most preferred phase of θ (the trough), resulting in cross-frequency coupling.

In Figure 1, after inhibition reduces over the θ cycle the most excitable representation would be itemized through low-middle γ , followed sequentially by other, less excitable clusters (the right-hand side of the figure detailing the role of δ will be returned to below).⁴ Certain of these γ clusters would then slow to β to be maintained as the ongoing structure’s phrasal identity. After the feature-sets are constructed, phase resetting θ would induce a pause in γ and β activity and, as a result, a γ and β phase reset (Teschke & Karhu, 2000). This would determine the composition of the feature-set, completed after θ phase resetting. The process of phase resetting involves the re-alignment of ongoing oscillatory phases in relation to a reference point. This reference is either endogenous or exogenous (Voloh & Womelsdorf, 2016), and so in the case of generating (sets of) feature-sets during, for instance, internal monologue, the ‘cue’ itself would be endogenous (see also Murphy 2016c for a similar proposal applied to the domain of pragmatics).

During the construction of each set, a ‘check list’ of items would be sent downstream to regions oscillating at slower rhythms to be interpreted in a strict, linear sequence. This model presupposes that a given cluster (or indeed a given neuron within a cluster) will only fire once during the slow cycle, something which at least approximates neurobiological plausibility in that spikes and bursts are followed by strong after-hyperpolarization currents which are

⁴ I should stress that there will certainly still be a need to invoke processes such as cultural evolution to account for the ultimate complexity of the grammatical system acquired by speakers, but this algorithmic model can at least act as the foundation of this linguistic knowledge.

normally a consequence of Ca-activated K currents which last ~100ms, hence reducing the chances of multiple within-cycle firings (Storm, 1990).

As mentioned above, certain of these γ clusters would slow to β to determine how the feature-set is stored in the existing 'cognitive set'. Reconciling any cartographic neurolinguistic model with the present oscillomic algorithm could take a number of forms, but perhaps the most parsimonious and empirically adequate direction would be to assume that after a phrase has been constructed via cross-frequency coupling feature-set construction (implicating temporal regions of the ventral stream like the medial temporal lobe), regions along the dorsal stream firing at β (likely regions closer towards the inferior frontal gyrus) would be able to synchronise with the cortico-basal ganglia-thalamic loop via phase-phase coupling and transfer the categorised set to be maintained through an increasing β rhythm. Regions in the left inferior frontal cortex typically seen as the 'seat of syntax' would consequently become buffering zones where particular multiplexed cycles are interpreted, hence their increased activity during fMRI scanning in response to hierarchical linguistic structures.

With the exception of this claim about β - β coupling, the present oscillomic model, and the views expressed in the neurolinguistics literature more generally, do not make any empirical claims about the function of phase coupling between identical rhythms. For instance, Kaplan, Bush, Bonnefond, Bandettini, Barnes, Doeller and Burgess (2014) discovered θ phase coupling between medial prefrontal cortex and medial temporal lobe during spatial memory recall, and θ - θ phase coupling would also be a possibility, for instance, between parahippocampal regions and the prefrontal cortex, although the precise functional role of this form of coupling remains an open question.

The property-attribution feature of labeling (i.e. attributing to a constructed set a particular computational identity, such as a DP or VP) can be explained through the above $\{\delta\{\theta(\gamma/\beta)\}\}$ code. But what about the other core feature, cyclicity? I suggest that this is achieved through β increases as phrases are constructed and labeled, with each 'old', already-constructed phrase being maintained by β as the newer phrases are constructed by the θ - γ code and selective γ decreases to β .

The particular kind of coupling relationship which could most effectively implement the above model is found in Figure 2: Phase-frequency cross-frequency coupling (D in Figure 2) ensures that the number of cycles of the faster rhythm correlates with the specific phase of the slower rhythm, while phase-power coupling ensures a correlation between the amplitude of the faster rhythm and the phase of the slower rhythm. Both of these coupling types could implement the current oscillomic model. As claimed above, cross-frequency coupling could coordinate activity across distant cortical areas as a way of integrating information from different representational domains, giving rise to the combinatoric power of linguistic generativity.

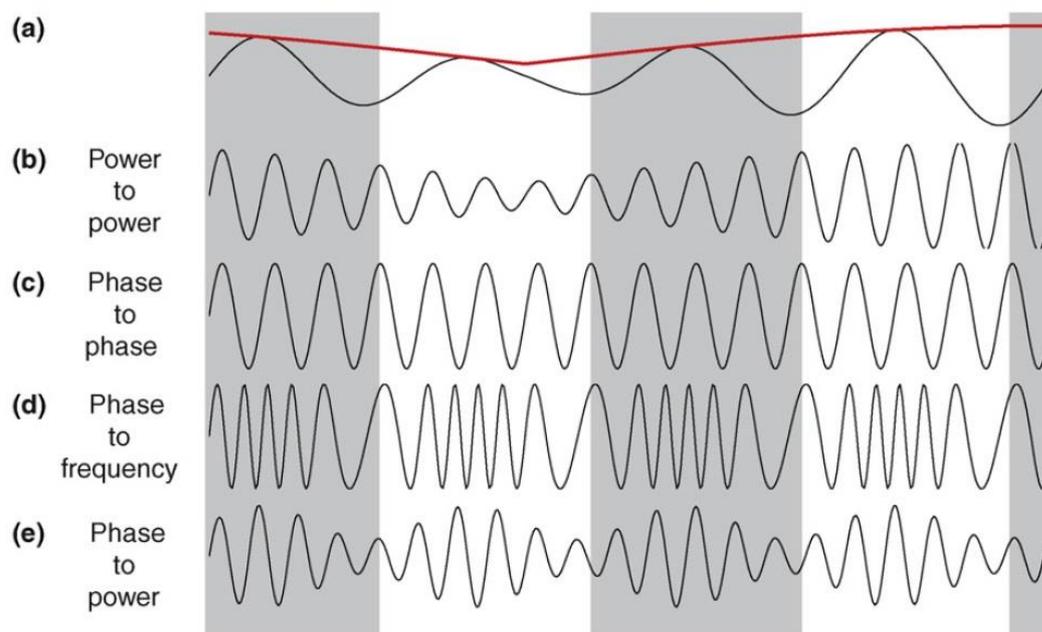


Figure 2: A representation of the different types of interplay slower and faster rhythms exhibit. (A) A theta oscillation and its power denoted by the red line. (B) Power-power: The amplitude of the slower and faster rhythms correlates. (C) Phase-phase: A set number of faster cycles within a given slower phase. (D) Phase-frequency: The faster rhythms correlate with the specific phase of the slower rhythm. (E) Phase-power: The power of the faster rhythm correlates with the phase of the slower rhythm. Reproduced from Jensen and Colgin (2007).

Early research by Sternberg (1966) suggested that short-term memory items are recalled at a rate of one per 30ms, and this appears to be the approximate temporal separation of γ cycles – possibly placing constraints on the number of linguistic features able to be retrieved. Given that top-down, higher-level processes involved in language comprehension begin to occur around 200-600ms post-utterance (Skeide & Friederici, 2016), this presumably presents these γ rhythms with enough time to extract a number of features before top-down processes (such as lexical-semantic categorisation and phrase structure reconstruction, coordinated by slower rhythms) end and bind them into feature-sets. This permits a natural transition to the well-established findings of the cartographic literature, which have shown that 200-400ms post-stimulus lexical information is sent from the left anterior superior temporal gyrus and superior temporal sulcus to BA 45 and BA 47 via the fronto-occipital fasciculus, at which point semantic relations between lexical feature-sets is determined via interactions with the inferior parietal cortex (Binder, Desai, Graves, & Conant, 2009).

A potentially interesting topic for research surrounds whether each stored γ item/representation within the slower rhythm is maintained with equal accuracy, or whether a degree of decay occurs depending on the circuits involved, and indeed whether this would influence the retrievability and interpretability of a given feature in comparison with its neighbours. Increases in γ , nested within θ , have been found across working memory maintenance tasks, including visual (Roux & Uhlhaas, 2014), auditory (Kaiser, Rahm, & Lutzenberger, 2009) and somatosensory (Haegens, Osipova, Oostenveld, & Jensen, 2010), suggesting that this general mechanism can be implemented via distinct neural codes triggering domain-specific representations.

The next issue which needs to be addressed concerns the regions involved in γ -itemisation during feature retrieval. Motivations for assuming that the stored γ -individuated features are widely distributed across the cortex come, most recently, from Keene, Bladon,

McKenzie, Liu, O'Keefe, and Eichenbaum (2016). This study of object-context associations in rats showed that all task dimensions (object position, identity and context) were encoded in every parahippocampal processing stream they investigated: the medial entorhinal cortex, the lateral entorhinal cortex and the perirhinal cortex. This degree of representational diversity across the hippocampus was present even at the single-cell level; neurons traditionally assumed to be 'spatial' cells (e.g. grid cells) were often involved in object identity, in the same way that the perhaps ill-named 'place cells' can also be involved in object encoding. It is therefore likely that semantic features triggered during sentence comprehension are also widely distributed and not specified to a particular language-relevant cortical or subcortical region. The subcellular format of the γ -itemisation process remains to be explored, however recent work by Kastellakis, Silva, and Poirazi (2016) provides something of a first step in developing an algorithmic model of associative memory formation by probing into the mechanisms underlying the linking of memories of different strengths. Such memories appear to exhibit synaptic co-clustering within the dendrites of neurons common to each memory.

$\{\theta(\gamma)\}$ coupling in entorhino-hippocampal regions has been shown to be crucial for memory recall (Schomburg, Fernandez-Ruiz, Mizuseki, Berenyi, Anastassiou, Koch, & Buzsaki, 2014), but reasons to believe that the $\{\theta(\gamma)\}$ code is causally implicated in memory retrieval and maintenance (and does not simply correlate with some experimental manipulation) come from Vosskuhl, Huster, and Herrmann's (2015) use of transcranial alternating current stimulation (tACS) to decrease participant's θ such that the $\theta:\gamma$ ratio changed and an abnormally large number of γ cycles could be nested within θ . This resulted in enhanced short-term memory performance (i.e. the storage of information, but not the manipulation of a memorised set of items), however working memory operations themselves (i.e. manipulation of items) were not affected, neither during nor after stimulation (see Figure 3).

This suggests that the cross-frequency coupling documented in memory tasks is not epiphenomenal but is rather representative of a coding scheme and the physical limitations of cognition. Given what has been reviewed above, the reason why working memory operations were not affected by θ modulation in Vosskuhl et al. (2015) may be a result of the fact that this process does not influence the structure of γ -itemization nor does it change the topology of the derived feature-set, it merely expands it.

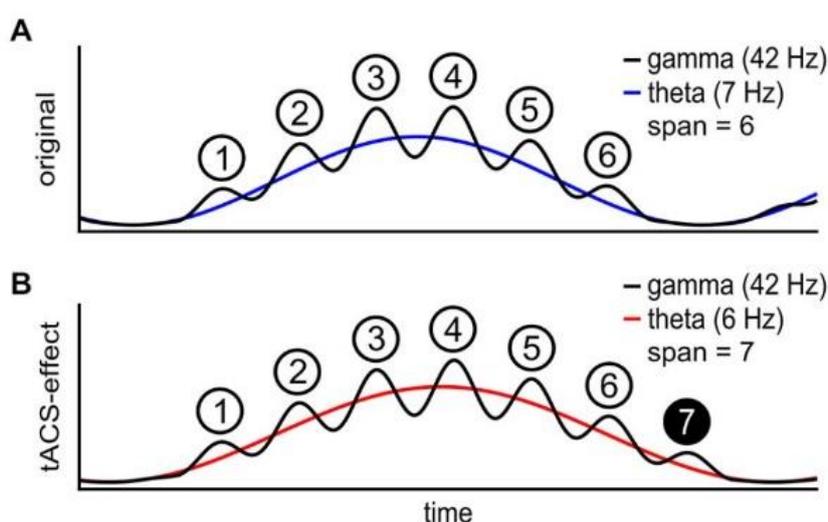


Figure 3: The theta-gamma neural code for working memory (A) and the slowing of theta via transcranial alternating current stimulation (tACS) permitting the maintenance of an additional item (B). Reproduced from Vosskuhl et al. (2015).

Finally, a recent study concerning episodic memory also made a small step towards the types of goals for neurolinguistics outlined here. Heusser, Poeppel, Ezzyat, and Davachi (2016) showed that episodic sequence memory is supported by a θ - γ phase code, such that ‘elements within an experience are represented by neural cell assemblies firing at higher frequencies (gamma) and sequential order is encoded by the specific timing of firing with respect to a lower frequency oscillation (theta)’. This should surely boost the confidence of language scientists trying to show that some form of phase-amplitude coupling is responsible for basic aspects of syntactic and semantic feature-set composition. Given that baboons have recently been shown to be capable of accurately storing the linear position of single episodes (Noser & Byrne, 2015), a comparative oscillomic approach to episodic memory could easily be adopted (for instance, of the kind proposed in Murphy 2016b).

2.4 α -Inhibition

The roles of δ , θ , β and γ in language have been discussed so far. But what of α ? Due to the inherently inhibitory nature of this rhythm (generated, for instance, via pulvinar connectivity with early/visual parietal cortex), it is also likely that it inhibits a given region permitting only the most excitable representations to be triggered. Along with the role attributed to it above of ‘generating’ the initial γ -itemization process, the inhibition and information gating capacities of α could contribute to the efficient coordination of the $\{\theta(\gamma)\}$ coupling model. The α band would therefore act to shield the ongoing concatenation of features from irrelevant information, otherwise excitable (a form of ‘protection’, for Roux and Uhlhaas (2014)). This would constitute a particular implementation of the ‘inhibition-timing hypothesis’ of Klimesch, Sauseng, and Hanslmayr (2007), according to which α can inhibit task-irrelevant neural circuits, increasing in amplitude over irrelevant regions. For instance, Friese, Köster, Hassler, Martens, Trujillo-Barreto, and Gruber (2013) discovered that successful memory encoding not only yielded enhanced $\{\theta(\gamma)\}$ cross-frequency coupling, but also decreased prefrontal and occipital α .⁵ Likewise, Michelmann, Bowman, and Hanslmayr (2016) found that when subjects mentally replayed a short sound or video clip, α decreases were found in sensory-specific regions. Using simulated neural networks, Gips, van der Eerden, and Jensen (2016) also showed that inhibitory α modulations coupled to γ serve to temporally segment visual information, preventing an overload of information. Finally, α decreases at right fronto-temporal sites were found when clear syllables were temporally expected, with a longer foreperiod duration (Wilsch, Henry, Herrmann, Maess, & Obleser, 2015). These expectancy effects also appear in language, with increased semantic predictability leading to reduced parieto-occipital α (Wöstmann, Herrmann, Wilsch, & Obleser 2015).

The finding that α is also generated in the thalamus and hippocampus (Buffalo, Fries, Landman, Buschman, & Desimone, 2011) lends further support to the role of subcortical structures in language comprehension, and the spatial proximity of thalamic and hippocampal α to θ and β sources suggests that the present oscillomic model could be implemented highly efficiently, with minimal conduction delays (see also Kleen, Testorf, Roberts, Scott, Jobst, Holmes, & Lenck-Santini (2016)). The susceptibility of particular circuits to synchronise with α is modulated by cholinergic and serotonergic mechanisms alongside glutamatergic afferents acting via metabotropic receptors (Uhlhaas, Haenschel, Nikolic, & Singer, 2008). Both metabotropic glutamate receptors (mGluR) and muscarinic acetylcholin receptors (mAChR)

⁵ See also Kaplan et al. (2014) for evidence of the role of $\{\theta(\gamma)\}$ cross-frequency coupling in spatial memory retrieval, with their θ source being the medial temporal lobe.

generate thalamo-cortical α , with this distinction being of particular functional relevance: Vijayan and Kopell (2012) tested a conductance-based thalamic model of awake α which demonstrated that mAChR-generated α supports information processing during tasks, while mGluR-generated α performs the above discussed role of shutting out interfering information.

Thalamic α has also been implicated in modulating cortical γ power, and is also suited to synchronize distinct cortical regions oscillating at α (Gips et al., 2016), increasing the likelihood that it plays an important role in phrase structure building and semantic composition. Finally, a recent verbal generation task by Wojtecki, Elben, Vesper and Schnitzler (2016) using EEG found θ - α power increases (i.e. 6-12Hz) and enhanced θ - α coherence between the subthalamic nucleus and frontal sites, with power increasing with performance on the task.

Thalamic, hippocampal and frontal α and θ rhythms may consequently act as, respectively, inhibitory and control processes which modulate γ -related processes involved in the retrieval and activation of language-relevant features, with β then being employed in the maintenance of existing feature-sets in memory. Altogether, this suggests that linguistic communication and the interpretation of speech/gesture/sign results in multiple brains effectively being coupled together, their oscillatory activity in some degree of synchrony.

Other studies can shed further light on the oscillatory dynamics employed by the language system. Hanslmayr, Staresina, and Bowman (2016) review the experimental literature from rodents and humans and suggest a general trend. While hippocampal $\{\theta(\gamma)\}$ synchronisation mediates the binding of distinct episodic memory representations, the desynchronisation of slower neocortical rhythms (α and β) also appears to mediate the encoding of episodic memories. Once again, we find the potentially inhibitory role of α in aiding successful memory encoding and retrieval by other regions of the brain. The hippocampal synchronisation system therefore appears to bind information, while the neocortical desynchronisation system stores the representational content.

2.5 Recursively Nesting Feature-Sets

One of the human-unique aspects of language is its featurally rich lexical representations, which cannot be reduced to simplex or binary structures but are rather composed of a cluster of distinct syntactic, semantic and phonological features. This level of representational complexity could possibly and feasibly be generated by more widespread cross-frequency coupling yielding greater levels of representational information than that permitted in the brains of other primates. Moreover, there appear to be human-unique levels of cross-frequency coupling diversity in the neocortex (Maris, Fries, & van Ede, 2016), with both phase and frequency being modulated to transfer information stored in local ensembles across distributed networks. In this sense, the computational properties of brain rhythms appear able to implement some (perhaps all) of the core syntactic and semantic operations of language.

Although phase-amplitude coupling is most commonly found between θ and γ , recent work suggests that θ can also entrain to δ in the human brain (Miller, Hermes, Honey, Sharma, Rao, den Nijs, Fetz, Sejnowski, Hebb, Ojemann, Makeig, & Leuthardt, 2010; Maris, van Vugt, & Kahana, 2011). Given experimental findings that δ can entrain to the full range of hierarchical linguistic structures from syllables to phrases and sentences (Ding et al., 2016), the discovery of human-specific $\{\delta(\theta)\}$ phase-amplitude coupling is potentially of great significance. I would like to suggest that the above $\{\theta(\gamma)\}$ feature-set multiplexing algorithm can itself be embedded within δ cycles originating within regions shown to entrain to hierarchical linguistic structures and which can also oscillate at slow δ frequencies; namely, the superior temporal gyrus and left inferior frontal cortex (see Figure 1). This additional layer of nesting would then give rise to the additional layer of hierarchy seen in human syntax, which

goes beyond both the phonological syntax of birdsong and the semantically atomic structures of primate calls and gestures:

LINGUISTIC HIERARCHY: {PHRASE{LEXICAL ITEM(FEATURES)}}
 OSCILLOMIC HIERARCHY: { δ { θ (γ)}}

Indeed, phase-amplitude coupling is itself not unique to language, let alone humans, being involved in perception, decision-making and navigation (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2005; Kepecs, Uchida, & Mainen, 2006), and so the interactions between distinct types of cross-frequency couplings and the representations they manipulate are instead the likely source of language-specific combinatorial capacities. There is a growing consensus that phase-amplitude coupling can support the phase coding of neural representations through, for instance, the categorisation of visual objects (Watrous, Deuker, Fell, & Axmacher, 2015), and given the inherently semantic nature of the categorisation processes documented by Watrous et al. – faces, tools, houses and scenes – it would be somewhat surprising if this form of cross-frequency coupling is not also involved in the interpretation of linguistic structures. As the authors of the study put it, exploring cross-frequency coupling ‘may provide new avenues for decoding the human representational system’.

The emergence of these species-specific oscillatory nestings would constitute the exclusive content of the ‘Small Bang’, as it was termed in Murphy (2015a); the neural alterations required to bring about modern *homo sapiens* and what linguists refer to as narrow syntax, the basis of Tattersall’s (2016) ‘symbolic’ capacity. The Labeling Hypothesis is therefore given renewed support.

A topic for future research concerns the physiological properties of { δ (θ)} and { θ (γ)} coupling, their developmental profile, and whether their emergence can be explained within certain evolutionary frameworks, such as the ‘globularity’ hypothesis (Benítez-Burraco & Murphy, 2016; Murphy & Benítez-Burraco, 2016) or perspectives oriented around human-specific myelination rates (Murphy, 2015b). Parsing research by Momma (2016) investigating the time course of syntactic priming effects has even suggested that structure-building units during production are very small, going against much of the production literature which claims that pre-formed templates are large. The above basic oscillomic schema (accounting as it does only for labeled units) may therefore carry considerable explanatory reach both in terms of production and comprehension.

To detail the oscillomic process of hierarchical phrase structure building in a more familiar way, consider Figure 4. In the four-step procedure outlined here, a determiner and noun are encountered and their respective feature-sets are triggered by the above α -supported and θ -coupled algorithm. The γ -itemized clusters responsible for identifying the phrase as a DP then slow to β , and the entire complex is nested within δ . The phrase is then maintained in memory (a necessary feature for labeling for occur) via an increasing β rhythm. The same procedure occurs for the next phrase, ν P, and the two β -maintained identities are stored together and interpreted in anterior portions of Broca’s area as the larger multi-phrasal unit, TP.

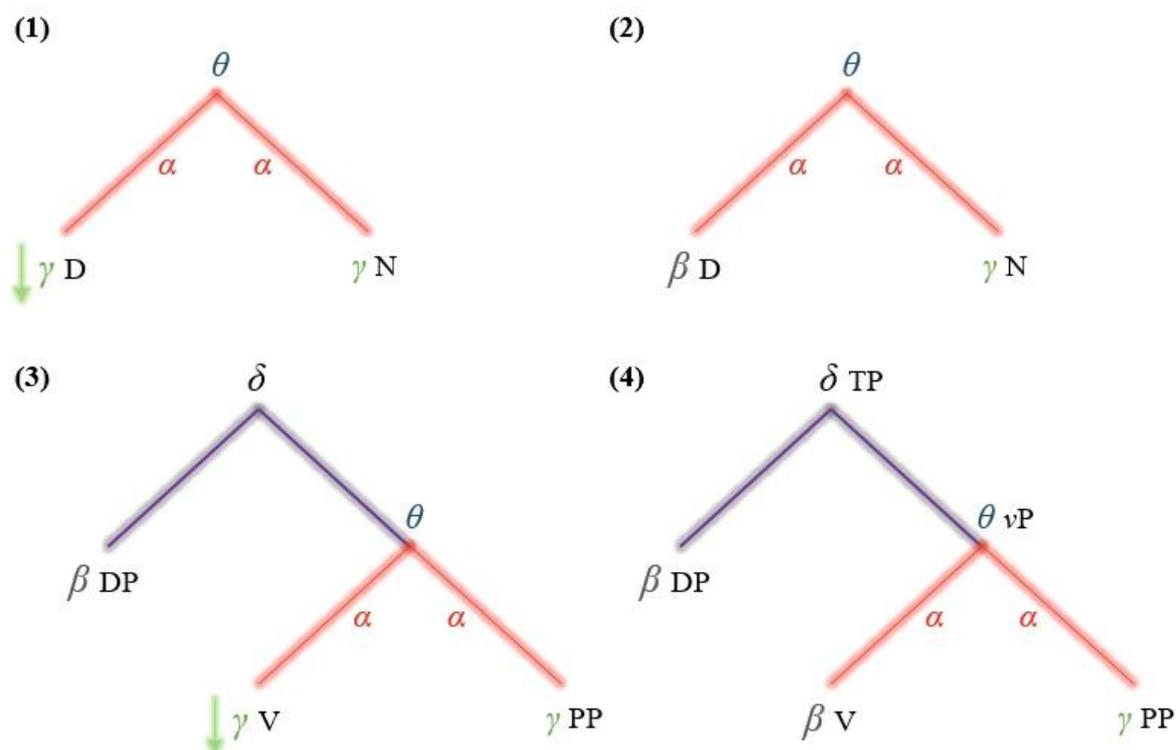


Figure 4: An oscillomic tree representing the putative rhythms responsible for particular lexical and phrasal structures according to the present multiplexing algorithm. ‘TP’ denotes Tense Phrase, ‘vP’ denotes Verb Phrase (e.g. ‘swam in the river’), ‘DP’ denotes Determiner Phrase (e.g. ‘the man’), and ‘PP’ denotes Prepositional Phrase (e.g. ‘in the river’). Regions of generation: Delta = left inferior frontal regions, anterior and middle temporal regions. Theta = parahippocampal regions. Alpha = thalamic and occipital regions. Beta = anterior regions of Broca’s area and the basal ganglia. Gamma = cross-cortical sites storing features in long-term memory. Following the findings in Ding et al. (2016), it is likely that the delta peak in (3) is at 2Hz due to it being a simple phrase, whereas the peak in (4) would shift to 1Hz due to its sentential nature.

If coupling across cortical and subcortical oscillations is responsible for the hierarchical combination of computations at the syllabic and phonemic levels, ‘restoring the natural arrangement of phonemes within syllables’ (Hyafil, Fontolan, Kabdebon, Gutkin, & Giraud, 2015), and if the present multiplexing algorithm is an accurate representation of syntactic and semantic feature composition, then we can conclude that oscillations are also responsible for restoring the natural arrangement of features within words.⁶

As discussed, I am assuming that α plays an important inhibitory role, but I additionally want to suggest that it is functionally closely related to θ cycles such that α can coordinate feature-sets organized by θ . Since α is a crucial part of domain-general attentional mechanisms, I will adopt the claim in Jensen, Gips, Bergmann and Bonnefond (2014) that direct attention is allocated only to the first items in a given θ - or α -constructed sequence, and that the final (and hence least excitable) items are processed pre-attentively. Along with explaining a number of visual phenomena (see Jensen et al., 2014), it is possible that this mechanism is responsible for

⁶ Oscillomic research into the development of attentional mechanisms will be crucial here, with recent work showing a direct correlation between the emergence of endogenous attentional mechanisms (mind-internal directed attention, in contrast to exogenous attention, directed at the environment) and the ability to track non-adjacent morphosyntactic dependencies, both emerging at around 12-15 months (de Diego-Balaguer, Martinez-Alvarez, & Pons, 2016).

certain facts about language, such that, quite often, only certain features of a given word or phrase come to attention during sentence comprehension and are made ‘prominent’. The firing of particular neural clusters representing a given feature X could also engage fast GABAergic inhibition (a result of γ activity) and inhibit the firing of clusters representing feature Y . Indeed, it has been found that patients with attention deficit hyperactivity disorder exhibit impaired control of posterior α (Mazaheri, Coffey-Corina, Mangun, Bekker, Berry, & Corbett, 2010), and it is possible that these sorts of oscillatory connectopathies could explain deficits in cognitive functions reliant to some extent on attention.

These assumptions are supported by a recent self-paced reading study which monitored oscillatory dynamics. Vignali, Himmelstoss, Hawelka, Richlan, and Hutzler (2016) had participants read syntactically and semantically well-formed sentences, sentences containing a semantic violation, and word lists. Fixations at semantically unrelated words elicited lower β desynchronisation, while γ power increased as well-formed sentences were processed. This γ effect was not found with word lists. In addition, θ power increased in the 300-800ms window after sentence onset in well-formed sentences, but not during word lists, lending support to the present hypothesized role of θ in memory retrieval and syntactic ‘chunking’ operations.⁷

If the present hypothesis about human-specific forms of δ -entrainment is correct, then non-human primates should exhibit markedly distinct oscillatory profiles. As reviewed in Murphy (2016b), the current experimental literature suggests that in both humans and monkeys, the construction of an internal sequence leads to β increases, maintenance failure/disruption leads to β decreases, and sequence execution results in subsequent γ increases. Yet as Figure 5 indicates, there also appears to be a less broad range of oscillatory couplings available in monkeys with respect to the crucial representational properties implicated in the forms of cognitive operations required for interpreting and producing call sequences.

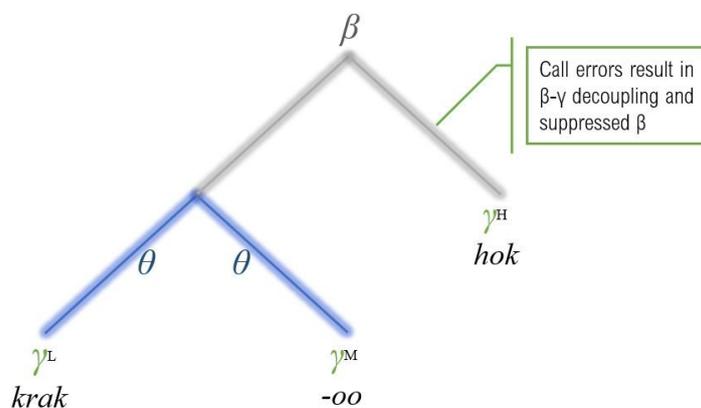


Figure 5: An oscillomic tree of the rhythms responsible for monkey call sequences (from Murphy 2016b). The superscript next to each γ cycle denotes low, medium and high γ (sharp wave-ripples), with γ power scaling with the number of calls held in memory. Each γ -generated item would couple with hippocampal θ (see Lee, Simpson, Logothetis and Rainer (2005) and Jutras, Fries and Buffalo (2013) for the role of this rhythm in monkey working memory) in order to achieve the binding of morphological elements ('krak-oo') before coupling with basal ganglia-

⁷ A further advantage of the present model is that it points to the compatibility between dynamical systems approaches to the brain and more traditional computational models which treat the brain as a digital computer. These perspectives are not incommensurable if we assume that particular discrete computational operations are executed via dynamical properties of neural oscillations.

initiated and striatal β , which would in turn increase in amplitude until the event of either an erroneous action sequence (call production) or the termination of the call series. This hypothesis is supported by studies claiming that β activity operates as a general coupling mechanism of assembly activity across brain structures (Fujioka, Trainor, Large, & Ross, 2012). Note that monkey oscillomic trees lack labels, being composed of call pair-combinations ('krak hok') and morphological elements ('-oo').

Notable support for the model in Figure 5 comes from Ramirez-Villegas, Logothetis, and Besserve's (2015) study of sharp wave-ripples in the macaque hippocampal CA3-CA1 network, which credited a role for γ rhythms in memory reactivation, transfer and consolidation (findings echoed in Kaplan et al., 2016). As mentioned, it seems that language-relevant δ -entrainment is not easily achieved in non-humans. For instance, while the rat globus pallidus can nest β and γ rhythms within a slower δ rhythm, Dejean, Arbuthnott, Wickens, Le Moine, Boraud, and Hyland's (2011) electrode recordings of local field potentials indicated that when such nesting did occur β and γ were negatively correlated and at anti-phase, with the phase positions of both the β and γ clusters being significantly different. The rat oscillome and its intrinsic rhythmic hierarchies may exhibit reduced nesting capabilities relative to the human oscillome – and, indeed, relative to the primate oscillome more generally, given Charvet, Hof, Raghanti, van der Kouwe, Sherwood, and Takahashi's (2016) findings of expanded anterior to posterior cortico-cortical tracts in primates relative to other mammals. These species-specific oscillomic interactions may lay the foundations of what Hauser and Watumull (2016) term the 'Universal Generative Faculty' common to language, mathematics, morality and music; a combinatorial system constructing a potentially infinite array of hierarchically organised sequences interfacing with distinct representational domains.

Connectivity issues would also be crucial to this account. It may ultimately be found that the $\{\delta\{\theta(\gamma)\}\}$ code is not strictly human-specific, but rather that the regions involved in long-range cross- and sub-cortical nesting are unique in their rich, cross-modular connections. Since frequency bands themselves fulfil various functional roles depending on their region of generation, it is possible that the present neural code can be found in some form or another across a range of mammalian species but that the particular regions involved in humans are widespread enough to produce the forms of rich cross-modularity exhibited by language.

In any event, these findings lead to direct predictions for the monkey oscillome given what is known about the combinatorial capacities of monkeys. Schlenker, Chemla, Schel, Fuller, Gautier, Kuhn, Veselinović, Arnold, Căsar, Keenan, Lemasson, Ouattara, Ryder, and Zuberbühler (2016), for instance, show that monkeys can combine two word-like elements, but cannot then combine this set with another atomic element. This would require a concatenation operation, a representational base ('lexicon'), and finally a temporary workspace where the combined set can be stored in memory. Rizzi (2016) calls this a 2-merge system (word-phrase merger), while monkeys seem limited to 1-merge systems (word-word) – although they may be capable of 2-merge systems if one defines a *krak-oo* conjunction followed by another distinct call as a type of phrase-word combination. Human language, in contrast, is a 3-merge system, permitting not just word-phrase merger, but phrase-phrase merger, requiring a secondary workspace in addition to the temporary workspace required by 2-merge systems. Nested rhythmic activity of the kind discussed here might act as the foundation of this secondary workspace, with a greater number of nested couplings seemingly available to the human oscillome yielding the greater number of stored representations in memory, although the precise details of this dual memory buffer (both with respect to anatomic connectivity and oscillatory dynamics) currently remain elusive.

interaction between these two layers. Displacement and agreement, for instance, connect items from across the two layers:

(5) What did you eat ___?

In this instance, a lexical item has been displaced to the functional layer, facilitating novel conceptual combinations and hierarchical relations. Cross-frequency interactions between the structures responsible for storing representations from each layer may give rise to this L-F-L-F pattern.

There does, however, seem to be a certain degree of disagreement concerning this oscillomic approach to language. Goucha, Zaccarella and Friederici (2016), for instance, make the following claim:

Alternative mechanisms based on brain oscillations have been proposed as a crucial element for the emergence of language (Murphy, 2015b) ... However, those mechanisms seem to already be in place in other species. For example, despite the crucial brain expansion that took place in primates and especially humans compared to other mammals, the rhythmical hierarchy of oscillations is mainly kept unchanged (Buzsáki et al., 2013). As Friederici and Singer (2015) pinpoint, the basic neural mechanisms behind cognition through the hierarchical embedding of oscillations are transversal across animals.

This is a perhaps unwarranted criticism since in the paper they discuss I cite the same authors as Goucha, Zaccarella and Friederici, pointing out that while the hierarchy of brain rhythms themselves may be preserved, it is crucially their cross-frequency coupling relations which are human-specific. The authors also make no attempt to explore the various functional roles of the ‘rhythmical hierarchy of oscillations’ in human and animal cognition, and it seems as if simply noting that Friederici and Singer (2015) ‘pinpoint’ something which is in fact well-known within neuroethology is a sufficient reason to reject any forthcoming linking hypotheses concerning linguistic computation and neural oscillations. Notice how in their claim that ‘the rhythmical hierarchy of oscillations is mainly kept unchanged’, one particular word is operative: *mainly*. No attempt is made by Goucha et al. to explore oscillatory features which are in fact human-specific, and what their functional implications may be for language and cognition.

Lastly, there are various ways to experimentally test the present model. Along with using MEG to compare word lists to acceptable phrases and comparing phrases of legal words to phrases of non-words (which would probe into the neural correlates of labeling), there are also particular stimuli types involving semantic rather than syntactic manipulations which could be of use. For instance, there are certain sentences permitting copredication which involve nouns being modified by multiple, semantically distinct adjectives (e.g. a *man* can be tall [PHYSICAL] but also happy [MENTAL STATE], and a *book* can be interesting [INFORMATION] but also blue [PHYSICAL]):

- (5) Distinct senses; legal noun: The tall and happy **man** left the building.
- (6) Related senses; legal noun: The tall and skinny **man** left the building.
- (7) Distinct senses; illegal noun: The tall and happy **mulg** left the building.
- (8) Related senses; illegal noun: The tall and skinny **mulg** left the building.

This stimuli set would reveal the oscillatory dynamics behind two major processes: (i) The maintenance in working memory of an ongoing sentential structure, and (ii) the attribution

of distinct semantic representations to a single nominal. Performing an analysis at the point of the word *left* will also be revealing, since this is the point when subjects will need to syntactically ‘reanalyse’ or ‘re-label’ the ongoing construction from a simple DP to a more complex TP.

3 Conclusion

Accruing evidence from independent domains in neuroscience and broader areas of biology converge on the suggestion that human-specific hierarchical interactions between brain rhythms may have brought about the emergence of language. Overall, the computational roles ascribed here are supported by a range of experimental findings, such as the general claims in Ketz, Jensen, and O’Reilly (2015) that θ is related to recollective/episodic memory processes, β is related to familiarity and executive control, and α is related to the gating of sensory information into higher order systems. Even something seemingly as simple as Merge or semantic composition requires complex interactions between a range of oscillations and regions. A number of research directions open up at this point, taking either the route of theoretical neurobiology or experimental linguistics, through which the neurocomputational basis of language may well reveal itself.

References

- Akam, T., & Kullmann, D.M. (2014). Oscillatory multiplexing of population codes for selective communication in the mammalian brain. *Nature Reviews Neuroscience*, *15*, 111-122.
- Bartolo, R., & Merchant, H. (2015). β oscillations are linked to the initiation of sensory-cued movement sequences and the internal guidance of regular tapping in the monkey. *Journal of Neuroscience*, *35*(11), 4635-4640.
- Bartolo, R., Prado, L., & Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *Journal of Neuroscience*, *34*, 3910-3923.
- Bastiaansen, M., & Hagoort, P. (2015). Frequency-based segregation of syntactic and semantic unification during online sentence level language comprehension. *Journal of Cognitive Neuroscience*, *27*(1), 2095-2107.
- Benítez-Burraco, A., & Murphy, E. (2016). The oscillopathic nature of language deficits in autism: from genes to language evolution. *Frontiers in Human Neuroscience*, *10*, 120.
- Berwick, R., & Chomsky, N. (2016). *Why Only Us: Language and Evolution*. Cambridge, MA: MIT Press.
- Binder, J.R., Desai, R.H., Graves, W.W., & Conant, L.L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*, 2767-2796.
- Boeckx, C. (2016). A conjecture about the neural basis of recursion in light of descent with modification. *Journal of Neurolinguistics* <http://dx.doi.org/10.1016/j.jneuroling.2016.08.003>.
- Boeckx, C., & Theofanopoulou, C. (2015). Commentary on: Labels, cognomes, and cyclic computation: an ethological perspective. *Frontiers in Psychology*, *6*, 784.
- Brown, J.W. (2014). The tale of the neuroscientists and the computer: why mechanistic theory matters. *Frontiers in Neuroscience*, *8*, 349.
- Buffalo, E.A., Fries, P., Landman, R., Buschman, T.J., & Desimone, R. (2011). Laminar differences in gamma and alpha coherence in the ventral stream. *PNAS*, *108*, 11262-11267.
- Buzsáki, G., Logothetis, N., & Singer, W. (2013). Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. *Neuron*, *80*(3), 751-764.
- Charvet, C.J., Hof, P.R., Raghanti, M.A., van der Kouwe, A.J., Sherwood, C.C., & Takahashi, E. (2016). Combining diffusion MR tractography with stereology highlights increased cross-cortical integration in primates. *Journal of Comparative Neurology* doi:10.1002/cne.24115.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87-114.
- De Diego-Balaguer, R., Martinez-Alvarez, A., & Pons, F. (2016). Temporal attention as a scaffold for language development. *Frontiers in Psychology*, *7*, 44.

- Dean, H.L., Hagan, M.A., & Pesaran, B. (2012). Only coherent spiking in posterior parietal cortex coordinates looking and reaching. *Neuron*, *73*(4), 829-841.
- Dejean, C., Arbutnot, G., Wickens, J.R., Le Moine, C., Boraud, T. & Hyland, B.I. (2011). Power fluctuations in beta and gamma frequencies in rat globus pallidus: association with specific phases of slow oscillations and differential modulation by dopamine D1 and D2 receptors. *Journal of Neuroscience*, *31*, 6098-6107.
- Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*, 158-164.
- Ewerdtwalbesloh, J.A., Palva, S., Rösler, F., & Khader, P.H. (2016). Neural correlates of maintaining generated images in visual working memory. *Human Brain Mapping* doi:10.1002/hbm.23313.
- Fernández-Ruiz, A., & Oliva, A. (2016). Distributed representation of 'what' and 'where' information in the parahippocampal region. *Journal of Neuroscience*, *36*(32), 8286-8288.
- Friederici, A.D. (2016). Evolution of the neural language network. *Psychonomic Bulletin & Review* doi: 10.3758/s13423-016-1090-x.
- Friederici, A. D., & Singer, W. (2015). Grounding language processing on basic neurophysiological principles. *Trends in Cognitive Sciences*, *19*(6), 329-338.
- Friese, U., Köster, M., Hassler, U., Martens, U., Trujillo-Barreto, N., & Gruber, T. (2013). Successful memory encoding is associated with increased cross-frequency coupling between frontal theta and posterior gamma oscillations in human scalp-recorded EEG. *NeuroImage*, *66*, 642-647.
- Fujioka, T., Trainor, L.J., Large, E.W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *Journal of Neuroscience*, *32*, 1791-1802.
- Gips, B., van der Eerden, J.P.J.M., & Jensen, O. (2016). A biologically plausible mechanism for neuronal coding organized by the phase of alpha oscillations. *European Journal of Neuroscience* doi:10.1111/ejn.13318.
- Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511-517.
- Goucha, T., Zaccarella, E., & Friederici, A.D. (2016). A revival of the *homo loquens* as a builder of labeled structures: neurocognitive considerations. Ms. Max Planck Institute for Human Cognitive and Brain Sciences, Germany.
- Hanslmayr, S., Staresina, B.P., & Bowman, H. (2016). Oscillations and episodic memory: addressing the synchronization/desynchronization conundrum. *Trends in Neurosciences*, *39*(1), 16-25.
- Haegens, S., Osipova, D., Oostenveld, R., & Jensen, O. (2010). Somatosensory working memory performance in humans depends on both engagement and disengagement of regions in a distributed network. *Human Brain Mapping*, *31*(1), 26-35.
- Hauser, M., & Watumull, G. (2016). The universal generative faculty: the source of our expressive power in language, mathematics, morality, and music. *Journal of Neurolinguistics* <http://dx.doi.org/10.1016/j.jneuroling.2016.10.005>.
- Hauser, M., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., Chomsky, N., & Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in Psychology*, *5*, 401.
- Haynes, W.I.A., & Haber, S.N. (2013). The organization of prefrontal-subthalamic inputs in primates provides an anatomical substrate for both functional specificity and integration: implications for basal ganglia models and deep brain stimulation. *Journal of Neuroscience*, *33*(11), 4804-4814.
- Heusser, A.C., Poeppel, D., Ezzyat, Y., & Davachi, L. (2016). Episodic sequence memory is supported by a theta-gamma phase code. *Nature Neuroscience* doi:10.1038/nn.4374.
- Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., & Giraud, A.-L. (2015). Speech encoding by coupled cortical theta and gamma oscillations. *eLife*, *10*, 7554.
- Jackendoff, R. (2007). *Language, Consciousness, Culture: Essays on Mental Structure*. Cambridge, MA: MIT Press.
- Jensen, O., & Colgin, L.L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences*, *11*, 267-269
- Jensen, O., Gips, B., Bergmann, T.O., & Bonnefond, M. (2014). Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends in Neurosciences*, *37*(7), 357-369.
- Jonas, E., & Kording, K. (2016). Could a neuroscientist understand a microprocessor? *bioRxiv* <http://dx.doi.org/10.1101/055624>.
- Jutras, M.J., Fries, P., & Buffalo, E.A. (2013). Oscillatory activity in the monkey hippocampus during visual exploration and memory formation. *PNAS*, *110*(32), 13144-13149.
- Kaiser, J., Rahm, B., & Lutzenberger, W. (2009). Temporal dynamics of stimulus-specific gamma-band activity components during auditory short-term memory. *NeuroImage*, *44*(1), 257-264.
- Kaplan, R., Adhikari, M.H., Hindriks, R., Mantini, D., Murayama, Y., Logothetis, N.K., & Deco, G. (2016). Hippocampal sharp-wave ripples influence selective activation of the default mode network. *Current Biology*, *26*, 1-6.

- Kaplan, R., Bush, D., Bonnefond, M., Bandettini, P.A., Barnes, G.R., Doeller, C.F., & Burgess, N. (2014). Medial prefrontal theta phase coupling during spatial memory retrieval. *Hippocampus*, *24*(6), 656-665.
- Kastellakis, G., Silva, A.J., & Poirazi, P. (2016). Linking memories across time via neuronal and dendritic overlaps in model neurons with active dendrites. *Cell Reports*, *17*, 1491-1504.
- Keene, C.S., Bladon, J., McKenzie, S., Liu, C.D., O'Keefe, J., Eichenbaum, H. (2016). Complementary functional organization of neuronal activity patterns in the perirhinal, lateral entorhinal, and medial entorhinal cortices. *Journal of Neuroscience*, *36*, 3660-3675.
- Kepecs, A., Uchida, N., & Mainen, Z.F. (2006). The sniff as a unit of olfactory processing. *Chemical Senses*, *31*, 167-179.
- Ketz, N.A., Jensen, O., & O'Reilly, R.C. (2015). Thalamic pathways underlying prefrontal cortex–medial temporal lobe oscillatory interactions. *Trends in Neurosciences*, *38*(1), 3-12.
- Kleen, J.K., Testorf, M.E., Roberts, D.W., Scott, R.C., Jobst, B.J., Holmes, G.L., & Lenck-Santini, P-P. (2016). Oscillation phase locking and late ERP components of intracranial hippocampal recordings correlate to patient performance in a working memory task. *Frontiers in Human Neuroscience*, *10*, 287.
- Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Research Reviews*, *53*, 63-88.
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C.E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, *320*(5872), 110-113.
- Lam, N.H.L., Schoffelen, J.-M., Udden, J., Hulst, A., & Hagoort, P. (2016). Neural activity during sentence processing as reflected in theta, alpha, beta, and gamma oscillations. *NeuroImage* <http://dx.doi.org/10.1016/j.neuroimage.2016.03.007>.
- Lee, H., Simpson, G.V., Logothetis, N.K., & Rainer, G. (2005). Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron*, *45*, 147-156.
- Lisman, J.E., & Jensen, O. (2013). The theta–gamma neural code. *Neuron*, *77*, 1002-1016.
- Maris, E., Fries, P., & van Ede, F. (2016). Diverse phase relations among neuronal rhythms and their potential function. *Trends in Neurosciences*, *39*(2), 86-99.
- Maris, E., van Vugt, M., & Kahana, M. (2011). Spatially distributed patterns of oscillatory coupling between high-frequency amplitudes and low-frequency phases in human iEEG. *NeuroImage*, *54*, 836-850.
- Martin, C., & Ravel, N. (2014). Beta and gamma oscillatory activities associated with olfactory memory tasks: Different rhythms for different functional networks? *Frontiers in Behavioral Neuroscience*, *8*, 218.
- Mas-Herrero, E., & Marco-Pallarés, J. (2016). Theta oscillations integrate functionally segregated sub-regions of the medial prefrontal cortex. *NeuroImage* doi:10.1016/j.neuroimage.2016.08.024.
- Mazaheri, A., Coffey-Corina, S., Mangun, G.R., Bekker, E.M., Berry, A.S., & Corbett, B.A. (2010). Functional disconnection of frontal cortex and visual cortex in attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *67*, 617-623.
- Michelmann, S., Bowman, H., & Hanslmayr, S. (2016). The temporal signature of memories: identification of a general mechanism for dynamic memory replay in humans. *PLoS Biology*, *14*(8), e1002528.
- Miller, K.J., Hermes, D., Honey, C.J., Sharma, M., Rao, R.P., den Nijs, M., Fetz, E.E., Sejnowski, T.J., Hebb, A.O., Ojemann, J.G., Makeig, S., & Leuthardt, E.C. (2010). Dynamic modulation of local population activity by rhythm phase in human occipital cortex during a visual search task. *Frontiers in Human Neuroscience*, *4*, 197.
- Miyagawa, S., Berwick, R.C., & Okanoya, K. (2013). The emergence of hierarchical structure in human language. *Frontiers in Psychology*, *4*, 71.
- Mohr, J.P., Pessin, M.S., Finkelstein, S., Funkenstein, H.H., Duncan, G.W., & Davis, K.R. (1978). Broca aphasia: pathologic and clinical. *Neurology* *28*: 311-324.
- Momma, S. (2016). Parsing, generation, and grammar. PhD dissertation. University of Maryland.
- Moro, A. (2006). Copular sentences. In M. Everaert & H. van Riemsdijk (Eds), *The Blackwell Companion to Syntax II* (pp. 1-23). Oxford: Blackwell.
- Mukherji, N. (2010). *The Primacy of Grammar*. Cambridge, MA: MIT Press.
- Murphy, E. (2015a). Labels, cognomes and cyclic computation: An ethological perspective. *Frontiers in Psychology* *6*: 715.
- Murphy, E. (2015b). The brain dynamics of linguistic computation. *Frontiers in Psychology*, *6*, 1515.
- Murphy, E. (2016a). The human oscillome and its explanatory potential. *Biolinguistics*, *10*, 6-20.
- Murphy, E. (2016b). Evolutionary monkey oscillomics: Generating linking hypotheses from preserved brain rhythms. *Theoretical Linguistics*, *42*(1-2), 117-137.
- Murphy, E. (2016c). A pragmatic oscillome: aligning visual attentional mechanisms with language comprehension. *Frontiers in Systems Neuroscience*, *10*, 72.
- Murphy, E., & Benítez-Burraco, A. (2016). Language deficits in schizophrenia and autism as related oscillatory connectopathies: an evolutionary account. *Neuroscience & Biobehavioral Reviews* <http://dx.doi.org/10.1016/j.neubiorev.2016.07.029>.

- Narita, H. (2014). $\ast\{t,t\}$. Poster presented at the 32nd West-Coast Conference on Formal Linguistics (WCCFL 32), University of Southern California.
- Nóbrega, V., & Miyagawa, S. (2015). The precedence of syntax in the rapid emergence of human language in evolution as defined by the integration hypothesis. *Frontiers in Psychology*, 6, 271.
- Noser, R., & Byrne, R.W. (2015). Wild chacma baboons (*Papio ursinus*) remember single foraging episodes. *Animal Cognition*, 18(4), 921-929.
- O'Keefe, J., & Nadal, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford: Oxford University Press.
- Poeppel, D. (2012). The maps problem and the mapping problem: two challenges for a cognitive neuroscience of speech and language. *Cognitive Neuropsychology*, 29, 34-55.
- Ramirez-Villegas, J.F., Logothetis, N.K., & Besserve, M. (2015). Sharp wave-ripple complexes in a reduced model of the hippocampal CA3-CA1 network of the macaque monkey. *BMC Neuroscience*, 16(Suppl 1), P15.
- Reuland, E. (2011). *Anaphora and Language Design*. Cambridge, MA: MIT Press.
- Richards, N. (2010). *Uttering Trees*. Cambridge, MA: MIT Press.
- Rizzi, L. (2016). Monkey morpho-syntax and merge-based systems. *Theoretical Linguistics*, 42(1-2), 139-145.
- Roux, F., & Uhlhaas, P.J. (2014). Working memory and neural oscillations: alpha-gamma versus theta-gamma codes for distinct WM information? *Trends in Cognitive Sciences*, 18, 16-25.
- Schlenker, P., Chemla, E., Schel, A.M., Fuller, J., Gautier, J.-P., Kuhn, J., Veselinović, D., Arnold, K., Căsar, C., Keenan, S., Lemasson, A., Ouattara, K., Ryder, R., & Zuberbühler, K. (2016). Formal monkey linguistics. *Theoretical Linguistics*, 42(1-2), 1-90.
- Schomburg, E.W., Fernandez-Ruiz, A., Mizuseki, K., Berenyi, A., Anastassiou, C.A., Koch, C., Buzsaki, G. (2014). Theta phase segregation of input-specific gamma patterns in entorhinal-hippocampal networks. *Neuron*, 84, 470-485.
- Seung, S. (2012). *Connectome: How the brain's wiring makes us who we are*. Boston: Houghton, Mifflin, Harcourt.
- Skeide, M.A., & Friederici, A.D. (2016). The ontogeny of the cortical language network. *Nature Reviews Neuroscience*, 17, 323-332.
- Sternberg S. (1966). High-speed scanning in human memory. *Science*, 153(736), 652-654.
- Storm, J.F. (1990). Potassium currents in hippocampal pyramidal cells. *Progress in Brain Research*, 83, 161-187.
- Sweeney-Reed, C.M., Zaehle, T., Voges, J., Schmitt, F.C., Buentjen, L., Kopitzki, K., Hinrichs, H., Heinze, H.-J., Rugg, M.D., Knight, R.T., & Richardson-Klavehn, A. (2015). Thalamic theta phase alignment predicts human memory formation and anterior thalamic cross-frequency coupling. *eLife*, 4, e07578.
- Tattersall, I. (2016). How can we detect when language emerged? *Psychonomic Bulletin & Review* doi:10.3758/s13423-016-1075-9.
- Tesche, C.D., & Karhu, J., (2000). Theta oscillations index human hippocampal activation during a working memory task. *PNAS USA*, 97, 919-924.
- Theyel, B.B., Llano, D.A., & Sherman, S.M. (2010). The corticothalamocortical circuit drives higher-order cortex in the mouse. *Nature Neuroscience*, 13(1), 84-88.
- Tomalin, M. (2006). *Linguistics and the Formal Sciences: The Origins of Generative Grammar*. Cambridge: Cambridge University Press.
- Tomalin, M. (2007). Reconsidering recursion in syntactic theory. *Lingua*, 117, 1784-800
- Uhlhaas, P.J., Haenschel, C., Nikolic, D., & Singer, W. (2008). The role of oscillations and synchrony in cortical networks and their putative relevance for the pathophysiology of schizophrenia. *Schizophrenia Bulletin*, 34, 927-943.
- Vignali, L., Himmelstoss, N.A., Hawelka, S., Richlan, F., & Hutzler, F. (2016). Oscillatory brain dynamics during sentence reading: a fixation-related spectral perturbation analysis. *Frontiers in Human Neuroscience*, 10, 191.
- Vijayan, S., & Kopell, N.J. (2012). Thalamic model of awake alpha oscillations and implications for stimulus processing. *PNAS U.S.A.*, 109(45), 18553-18558.
- Voloh, B., & Womelsdorf, T. (2016). A role of phase-resetting in coordinating large scale neural networks during attention and goal-directed behavior. *Frontiers in Systems Neuroscience*, 10, 18. <http://dx.doi.org/10.3389/fnsys.2016.00018>
- Voskuhl, J., Huster, R.J., & Herrmann, C.S. (2015). Increase in short-term memory capacity induced by down-regulating individual theta frequency via transcranial alternating current stimulation. *Frontiers Human Neuroscience* 9, 257. doi: 10.3389/fnhum.2015.00257.
- Watrous, A.J., Deuker, L., Fell, J., & Axmacher, N. (2015). Phase-amplitude coupling supports phase coding in human ECoG. *eLife*, 4, e07886.
- Wilsch, A., Henry, M.J., Herrmann, B., Maess, B., & Obleser, J. (2015). Alpha oscillatory dynamics index temporal expectation benefits in working memory. *Cerebral Cortex*, 25(7), 1938-1946.
- Wojtecki, L., Elben, S., Vesper, J., & Schnitzler, A. (2016). The rhythm of the executive gate of speech: subthalamic low-frequency oscillations increase during verbal generation. *European Journal of Neuroscience* doi: 10.1111/ejn.13429.

- Wöstmann, M., Herrmann, B., Wilsch, A., & Obleser, J. (2015). Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. *Journal of Neuroscience*, 35(4), 1458-1467.
- Yan, B., & Li, P. (2013). The emergence of abnormal hypersynchronization in the anatomical structural network of human brain. *NeuroImage*, 65, 34-51.
- Yelnik, J., Percheron, G., & Francois, C. (1984). A Golgi analysis of the primate globus pallidus. II. Quantitative morphology and spatial orientation of dendritic arborizations. *Journal of Comparative Neurology*, 227(2), 200-213.

Joint Interventions in Autism Spectrum Disorder: Relating Oscillopathies and Syntactic Deficits*

Charlotte Wilkinson, Elliot Murphy

Abstract

We put forward a dual clinical intervention strategy for people exhibiting language deficits in autism. This combines an emerging understanding of the ‘oscillopathic’ brain – how an abnormal profile of neural oscillations can give rise to syntactic and semantic deficits – with a more traditional approach centred on a course of speech therapy targeting specific linguistic deficits. It is our intention that a joint intervention focusing on abnormal oscillations alongside deficient language comprehension will also serve to re-align the general focus of inquiry away from cartographic neuroimaging issues and towards dynamic, oscillation-based analyses.

Keywords: Autism spectrum disorder, TMS/tDCS, SLT, neural oscillations, syntax, semantics

1 Introduction

Autism spectrum disorders (ASD) are typically defined as neurodevelopmental disorders involving a number of cognitive, social and communicative deficits. People with ASD exhibit repetitive behaviour alongside communicative problems such as a lack of eye contact and an inability to read paralinguistic cues (Bailey, Phillips, & Rutter, 1996). Systematic causal connections have been drawn between ASD and particular linguistic deficits (Crespi & Badcock, 2008), developments which we would here like to relate in an explicitly programmatic fashion to recent work suggesting that such deficits arise from an abnormally synchronised brain (Benítez-Burraco & Murphy, 2016; Murphy & Benítez-Burraco, 2016a). We believe that our developing understanding of syntactic and semantic deficits in ASD can complement these investigations into the neural oscillations putatively responsible for them. We will propose a joint intervention strategy through which a therapeutic course targeting syntactic deficits in ASD will be delivered alongside a series of repetitive transcranial magnetic stimulation (rTMS) or transcranial direct current stimulation (tDCS) interventions used to modulate the oscillations responsible for the abnormal linguistic ASD profile. Similar approaches have already been taken to improve performance on visual tasks in patients with schizophrenia (Farzan, Barr, Sun, Fitzgerald, & Daskalakis, 2012; Barr, Farzan, Rajji, Voineskos, Blumberger, & Arenovich, 2013), and we hope that the evidence presented here will convince researchers and therapists to consider the viability of a dual syntactic-oscillatory intervention in the ASD population.

*We would like to thank the patients and staff of the Central London Community Healthcare NHS Trust for their support and guidance in developing the ideas presented here, and also the speech and language therapy department at Woodfield School, London. Our thanks go to the attenders of the 2016 Nijmegen Lectures at the Max Planck Institute for Psycholinguistics for fruitful discussion. This work was supported by an ESRC scholarship (1474910) to EM.

2 Language Deficits in ASD

Emerging research indicates that at the neurological level ASD is primarily characterised by asynchronous neural oscillations (Tierney, Gabard-Durnam, Vogel-Farley, Tager-Flusberg, & Nelson, 2012). Because children can progress from one disorder sub-type to another during development, we believe that a longitudinal approach to research regarding children with ASD involving perennial ‘oscillopathic’ monitoring and manipulation/stimulation could generate a more robust diagnosis and prognosis. But in order to decide which brain regions and rhythms should be targeted, an understanding of what the precise language-related deficits in ASD are is needed.

Language deficits in ASD are no longer seen as being confined to pragmatic or world knowledge impairments, which are also widely compromised amongst other cognitively impaired children (Abbeduto & Hesketh, 1997). Indeed, ASD is often comorbid with language disorders such as Specific Language Impairment (Tager-Flusberg, 2006). While a core language deficit may exist in ASD, this is hard to pinpoint empirically due to variability in the ASD linguistic phenotype and concomitant linguistic profiles. Nevertheless, it is by now clear that certain linguistic constructions are generally difficult for ASD individuals to interpret.

It has been shown, for instance, that individuals with ASD integrate semantic information differently when interpreting phrases (Eigsti, Bennetto, & Dadlani, 2007) and that semantic knowledge is consolidated differently, being less primed by semantically-related words. Language delay also occurs with syntactic knowledge (Geurts & Embrechts, 2008; Philofsky & Fidler, 2007; Saldaña & Frith, 2007). Relatedly, problems with syntax suggest an impairment of procedural memory.

Language impairments in ASD involve difficulties with comprehending relative clauses, *wh*-questions, raising and passives (Perovic & Janke, 2013; Perovic, Modyanova, & Wexler, 2007; Wada, 2015). Children with autism also exhibit markedly degraded performance when interpreting reflexive pronouns (Fortunato-Tavares, Andrade, Befi-Lopes, Limongi, Fernandes, & Schwartz, 2015; Perovic, Modyanova, & Wexler, 2013). When presented with a battery of dative expressions exhibiting either syntactic alternations or restrictions (either in syntactic constructions or in marked semantic differences) and asked to identify the indirect object of the verb, children with autism performed worse than typically developing controls when the dative could syntactically alternate compared to when it was restricted (Stockbridge, Happé, & White, 2014). Children with autism consequently exhibit severe limitations in their ability to interpret alternating syntactic constructions, and we would like to suggest that this lack of flexibility may emerge from restrictions in oscillatory interactions. Since syntactic capacities directly impact cognition (albeit in ways which are still under debate) the neural basis of these linguistic deficits may shed some light on the biological basis of ASD more broadly.

3 Oscillopathies in ASD

Many of the nerve tracks and regions which differ in the ASD brain from neurotypical individuals are also implicated in language comprehension (Benítez-Burraco & Murphy, 2016), but the idiosyncratic nature of these regional differences (as documented in Hahamy, Behrmann, & Malach (2015)) leads us to suggest that neural oscillations provide a more robust scale of neurolinguistic analysis – in particular given that oscillations are increasingly being shown to be responsible for a number of higher cognitive capacities. For instance, brain dynamics can provide plausible candidates for linguistic computation. Following the model outlined in Murphy (2015, 2016), we will assume that set-formation (the taking of two objects,

X and Y, and forming the set $[X, Y]$) is implemented via α inhibition ‘shielding’ the embedding of a series of cross-cortical γ rhythms within parahippocampal θ , and that maintaining these sets in memory (creating linguistic phrases) requires the slowing down of certain γ clusters to β . The β band has independently been shown to be implicated in maintaining mentally constructed objects in memory (Engel & Fries, 2010). Finally, we will assume that agreement relations are implemented via cross-cortical evoked γ , with this rhythm being involved in attention and perceptual ‘feature binding’ (Sohal, Zhang, Yizhar, & Deisseroth, 2009).

With this background, we can turn to the oscillopathic ASD brain and use it as a way to approach possible electromagnetic remedial interventions. Early models of the neurobiological foundations of ASD placed much emphasis on anatomical abnormalities, in particular the volume of certain regions like the amygdala and cerebellum (Amaral, Schumann, & Nordahl, 2008). However, the varying neuropathology of the condition points to more global systematic differences (Murphy & Benítez-Burraco, 2016a).

While studies of the linguistic ASD ‘oscillome’ (neural oscillations ascribed particular computational properties) remain in their infancy, we believe enough has been learned to be able to construct a model of intervention targets. Kikuchi, Shitamichi, Yoshimura, Ueno, Hiraishi, Hirose, Munosue, Nakatani, Tsubokawa, Haruta et al. (2013) and Rojas, Maharajh, Teale, and Rogers (2008) reported significantly increased γ power for ASD individuals, with Kikuchi et al. finding in addition reduced cross-cortical θ , α and β . Even slight modulations in the phase of these slower rhythms could lead to the inability of external memory and interpretive systems to ‘read off’ the content of the feature-sets constructed by the above process of cross-cortical γ nesting within θ . Indeed, we believe that the notion of weak central coherence used by Happé and Frith (2006) to account for language deficits in autism (a reduced ability to distil generalizable information from diverse input) is a direct consequence of disturbed lower-level oscillatory coherence; see, for instance, the communication-through-coherence hypothesis in Fries (2005, 2015) which claims that oscillatory synchronisation modulates the efficacy of anatomical connections.

Inhibiting γ -oscillating regions and stimulating sub-cortical structures oscillating at the slower θ – β ranges during these and similar semantic memory tasks would likely prove conducive to increased syntactic and semantic competence. Bangel, Batty, Ye, Meaux, Taylor, and Doesburg (2014) found reduced β during a number estimation task in ASD, and broader disruptions in rhythmic coordination have been frequently documented (Benítez-Burraco and Murphy, 2016). The ASD brain also appears to involve reduced θ during tasks necessitating inter-regional synchronisation (Doesburg, Vidal, & Taylor, 2013). If parahippocampal θ is a major site of such inter-regional synchronisation (as argued by Fernández-Ruiz and Oliva (2016)) with respect to varying representational properties of language (i.e. semantic, phonological and syntactic representations), then this finding would contribute to an explanation for why people with ASD often have difficulties with this cross-modular feature of language.

Unusually long-lasting prefrontal and central γ has also been documented in ASD individuals during the interpretation of semantic incongruity (Braeutigam, Swithenby, & Bailey, 2008), possibly indicating the recruitment of a general search mechanism (fast γ) to replace the normal rhythmic processes (slow γ) responsible for retrieving and comparing semantic representations. This speaks to the legitimacy of an rTMS or tDCS intervention, as do the findings that picture-naming tasks result in reduced γ and β in the left inferior frontal gyrus in ASD participants (Buard, Rogers, Hepburn, Kronberg, & Rojas, 2013). In addition, abnormal γ oscillations may play a role in the pathogenesis of the inability to properly interpret reflexive pronouns, discussed above. Jochaut, Lehongre, Saitovitch, Devauchelle, Oulasagasti, Chabane, Zilbovicius, and Giraud (2015) also found altered oscillatory connectivity between auditory and language cortices in ASD subjects, but the precise phasal properties of this

oscillopathy – and how these properties could relate to and inform an oscillopathic model of language deficits in autism – remains an open topic.

Finally, in a recent MEG study of ASD subjects, Ghuman, van den Honert, Huppert, Wallace, and Martin (2016) found local hypersynchrony in lateral occipitotemporal θ , and long-range hyposynchrony in the α band that was most severe in circuits implicated in social processing. This α hyposynchrony correlated with social symptom severity, and we can infer from this that linguistic competence would also likely be impaired due to the strong relation between these two faculties. These findings point us in a very clear direction for targeted interventions.

4 Therapeutic Interventions in ASD

Focusing rTMS/tDCS interventions on the oscillations responsible for syntactic deficits in ASD brings with it much promise, with TMS recently being used to relieve ASD symptoms such as irritability and repetitive behaviours while enhancing hand-eye coordination, information processing and social skills (Sokhadze, El-Baz, Sears, Opris, & Casanova, 2014; Oberman, Enticott, Casanova, Rotenberg, Pascual-Leone, & McCracken, 2016; Casanova, Hensley, Sokhadze, El-Baz, Wang, Li, & Sears, 2014). Moreover, Oberman et al. (2016, p. 184) claim that rTMS ‘may represent a novel treatment strategy for reducing some of the core and associated ASD symptoms’. This suggests that rTMS interventions targeting the presently discussed regions would likely contribute to the improvement of linguistic processing.

ASD is a spectral disorder, and so an individual’s linguistic abilities are naturally idiosyncratic. The specific type of speech and language therapy (SLT) intervention, to be delivered alongside any rTMS/tDCS interventions, which would be most efficacious consequently depends upon particular deficits and the targets set by the patient and their SLT practitioner after initial evaluation. As noted above, there is an emerging consensus regarding the syntactic and semantic deficits exhibited by individuals with autism (e.g. difficulties with reflexive pronouns, *wh*-questions and certain dative expressions). Crucially, Perovic et al. (2013) present reasons to believe that such deficits arise directly from grammatical competence, and are not the consequence of pragmatic deficiencies. Our present joint intervention proposal is thereby granted a strong degree of alignment between phenotypic syntactic deficits and the oscillatory abnormalities claimed here to be responsible for them. As a result, we are confident that any rTMS/tDCS intervention will produce the desired remedial outcomes and will not simply increase performance on ancillary, language-external processes which happen to be tied, on way or another, to the core linguistic deficits.

Typically, SLT targets communication, social interaction and life skills (Bartlett, Gharani, Millonig, & Brzustowicz, 2005; for a systematic overview see Goldstein (2002)). Any intervention would be dependent on the individual consistently demonstrating competent verbal communication; severely delayed speech onset and/or mutism has long been observed in ASD (Kanner, 1943; Lovaas, 1977). For younger patients, SLT approaches similar to the Early Start Denver Model (Dawson, Rogers, Munson, Smith, Winter, Greenon, Donaldson, & Varley, 2010) or pre-school Autism Communication Therapy (Green, Green, Charman, McConachie, Aldred, Slonims, Howlin, Couteur, Leadbitter, Hudry, Byford et al., 2010) would be appropriate, and it would be essential for practitioners to find novel ways of integrating the training of syntactic competence with more traditional developmental pragmatic approaches which involve naturalistic interactions and collaborative work. Most of the therapeutic interventions developed in recent years have focused specifically on pragmatic impairments, and we leave it to future work to develop robust interventions for syntactic and semantic deficits.

Some developmental concerns naturally arise at this point. As Goswami (2016) points out, the oscillatory profile responsible for language comprehension appears to go through a number of stages throughout childhood and into young adulthood, and so a more careful and considered analysis of the oscillatory basis of language deficits in children and young adults with autism would be needed before a joint intervention of the kind proposed here can reasonably be expected to yield remedial outcomes.

Independently, both rTMS/tDCS and SLT interventions have provided positive results for individuals with ASD, and we believe that a combination of these techniques will result in boosted outcomes, with long-lasting effects. Since the noted linguistic deficits in ASD all require the recruitment of procedural and semantic memory resources, we believe that rTMS/tDCS stimulation/inhibition (depending on the oscillopathy) of the presently documented rhythmic sources will lead to a significant improvement in the language capacities of individuals with ASD. Recently, Iaccarino, Singer, Martorell, Rudenko, Gao, Gillingham, Mathys, Seo, Kritskiy, Abdurrob et al. (2016) restored defective γ waves in mice with Alzheimer's disease, which resulted in remedial effects through microglia modifications, and this approach could easily be applied to people with autism.

In addition, since ASD patients have clear language-related difficulties with working memory, which ultimately impede language processing, it is sensible to ask how the brain performs working memory tasks, and use this knowledge to construct a model of the rhythms and regions which should be attended to during rTMS/tDCS intervention. Lisman and Jensen (2013), for instance, argue that coupled γ and θ oscillations form a code for representing multiple, sequenced items in memory. These rhythms are generated in the cortex (particularly occipital regions) and hippocampus, pointing ASD interventions in a clear direction.

5 Conclusion

To conclude, we believe that not only do oscillatory signatures yield new endophenotypes of ASD, but they can also point therapists and clinicians towards a clear set of task-dependent electrophysiological interventions. We have suggested that this program should be linked to more traditional speech therapy interventions, targeting (i) the oscillopathic signatures implicated in linguistic – and, more specifically, syntactic – deficits in ASD via transcranial magnetic stimulation and transcranial direct-current stimulation, and (ii) the specific syntactic difficulties documented in ASD (e.g. the interpretation of reflexive pronouns, as in Fortunato-Tavares et al., 2015). A course of speech therapy targeting these linguistic structures, alongside the more direct oscillopathic intervention, would be a novel joint program in its comprehensiveness and intensity, particularly if as little as one hour a week of SLT can make a significant difference (Vismara, Colombi, & Rogers, 2009). Finally, given the identification of similar oscillopathies in schizophrenia (Murphy & Benítez-Burraco, 2016b), rTMS could also be used as a therapeutic intervention to modulate the oscillations responsible for the abnormal linguistic profile documented in the brains of individuals with schizophrenia – a topic for future, and promising, research.

References

- Abbeduto, L., Hesketh, L.J. (1997). Pragmatic development in individuals with mental retardation: learning to use language in social interactions. *Mental Retardation and Developmental Disabilities Research Reviews*, 3, 323–333.
- Amaral, D.G., Schumann, C.M., & Nordahl, C.W. (2008). Neuroanatomy of autism. *Trends in Neurosciences*, 31, 137–145.

- Bailey, A., Phillips, W., & Rutter, M. (1996). Autism: towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives. *Journal of Child Psychology and Psychiatry*, 37(1), 89–126.
- Bangel, K.A., Batty, M., Ye, A.X., Meaux, E., Taylor, M.J., & Doesburg S.M. (2014). Reduced beta band connectivity during number estimation in autism. *NeuroImage: Clinical*, 6, 202–213.
- Barr, M.S., Farzan, F., Rajji, T.K., Voineskos, A.N., Blumberger, D.M., Arenovich, T., et al. (2013). Can repetitive magnetic stimulation improve cognition in schizophrenia? Pilot data from a randomized controlled trial. *Biological Psychiatry*, 73, 510–517.
- Bartlett, C.W., Gharani, N., Millonig, J.H., & Brzustowicz, L.M. (2005). Three autism candidate genes: a synthesis of human genetic analysis with other disciplines. *International Journal of Developmental Psychology*, 23, 221–234.
- Benítez-Burraco, A., and Murphy, E. (2016). The oscillopathic nature of language deficits in autism: from genes to language evolution. *Frontiers in Human Neuroscience*, 10, 120. doi:10.3389/fnhum.2016.00120.
- Braeutigam, S., Swithenby, S.J., & Bailey, A.J. (2008). Contextual integration the unusual way: a magnetoencephalographic study of responses to semantic violation in individuals with autism spectrum disorder. *European Journal of Neuroscience*, 27, 1026–1036.
- Buard, I., Rogers, S.J., Hepburn, S., Kronberg, E., & Rojas, D.C. (2013). Altered oscillation patterns and connectivity during picture naming in autism. *Frontiers in Human Neuroscience*, 7, 742. doi: 10.3389/fnhum.2013.00742.
- Casanova, M.F., Hensley, M.K., Sokhadze, E.M., El-Baz, A.S., Wang, Y., Li, X., & Sears, L. (2014). Effects of weekly low-frequency rTMS on autonomic measures in children with autism spectrum disorder. *Frontiers in Human Neuroscience*, 8, 851. doi: 10.3389/fnhum.2014.00851.
- Crespi, B., & Badcock, C. (2008). Psychosis and autism as diametrical disorders of the social brain. *Behavioral and Brain Sciences*, 31(3), 241–320.
- Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., Donaldson, A., & Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. *Pediatrics*, 125(1), e17–23.
- Doesburg, S.M., Vidal, J., & Taylor, M.J. (2013). Reduced theta connectivity during set-shifting in children with autism. *Frontiers in Human Neuroscience*, 7, 785. doi: 10.3389/fnhum.2013.00785.
- Eigsti, I.M., Bennetto, L., & Dadlani, M. (2007). Beyond pragmatics: morphosyntactic development in autism. *Journal of Autism and Developmental Disorders* 37, 1007–1023.
- Engel, A.K., & Fries, P. (2010). Beta-band oscillations – signalling the status quo? *Current Opinion in Neurobiology*, 20, 156–165.
- Farzan, F., Barr, M.S., Sun, Y., Fitzgerald, P.B., & Daskalakis, Z.J. (2012). Transcranial magnetic stimulation on the modulation of gamma oscillations in schizophrenia. *Annals of the New York Academy of Science*, 1265, 25–35.
- Fernández-Ruiz, A., & Oliva, A. (2016). Distributed representation of ‘what’ and ‘where’ information in the parahippocampal region. *Journal of Neuroscience*, 36(32), 8286–8288.
- Fortunato-Tavares, T., Andrade, C.R.F., Befi-Lopes, D., Limongi, S.O., Fernandes, F.D.M., & Schwartz, R.G. (2015). Syntactic comprehension and working memory in children with specific language impairment, autism or Down syndrome. *Clinical Linguistics & Phonetics*, 29, 499–522.
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9, 474–480.
- Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron*, 88, 220–235.
- Geurts, H., & Embrechts, M. (2008). Language profiles in ASD, SLI, and ADHD. *Journal of Autism and Developmental Disorders*, 38, 1931–1943.
- Ghuman, A.S., van den Honert, R.N., Huppert, T.J., Wallace, G.L., & Martin, A. (2016). Aberrant oscillatory synchrony is biased towards specific frequencies and processing domains in the autistic brain. *Biological Psychiatry* <http://dx.doi.org/10.1016/j.bpsc.2016.07.006>.
- Goldstein, H. (2002). Communication intervention for children with autism: a review of treatment efficacy. *Journal of Autism and Developmental Disorders*, 32(5), 373–396.
- Goswami, U. (2016). Educational neuroscience: neural structure-mapping and the promise of oscillations. *Current Opinion in Behavioural Sciences* 10: 89–96.
- Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, P., Couteur, A.L., Leadbitter, K., Hudry, K., Byford, S., Barrett, B., Temple, K., Macdonald, W., Pickles, A., & PACT Consortium. (2010). Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. *Lancet*, 375(9732): 2152–2160.
- Hahamy, A., Behrmann, M., & Malach, R. (2015). The idiosyncratic brain: distortion of spontaneous connectivity patterns in autism spectrum disorder. *Nature Neuroscience*, 18, 302–309.

- Happé, F., & Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36, 5–25.
- Iaccarino, H.F., Singer, A.C., Martorell, A.J., Rudenko, A., Gao, F., Gillingham, T.Z., Mathys, H., Seo, J., Kritskiy, O., Abdurrob, F., Adaikkan, C., Canter, R.G., Rueda, R., Brown, E.N., Boyden, E.S., & Tsai, L.H. (2016). Gamma frequency entrainment attenuates amyloid load and modifies microglia. *Nature*, 540, 230–235.
- Jochaut, D., Lehongre, K., Saitovitch, A., Devauchelle, A-D., Olasagasti, I., Chabane, N., Zilbovicius, M., & Giraud, A-L. (2015). Atypical coordination of cortical oscillations in response to speech in autism. *Frontiers in Human Neuroscience*, 9, 171. doi: 10.3389/fnhum.2015.00171.
- Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, P., Le Couteur, A, Leadbitter, K., Hudry, K., Byford, S., Barrett, B., Temple, K., & Macdonald, W. (2000). Parent-mediated communication-focused treatment in children with autism (PACT): a randomised control trial. *Lancet*, 375(9732), 2152–2160.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2(3), 217–250.
- Kikuchi, M., Shitamichi, K., Yoshimura, Y., Ueno, S., Hiraishi, H., Hirotsawa, T., Munesue, T., Nakatani, H., Tsubokawa, T., Haruta, Y., Oi, M., Niida, Y., Remijn, G.B., Takahashi, T., Suzuki, M., Higashida, H., & Minabe, Y. (2013). Altered brain connectivity in 3-to 7-year-old children with autism spectrum disorder. *NeuroImage: Clinical*, 2, 394–401.
- Lovaas, O.I. (1977). *The Autistic Child: Language Development through Behavior Modification*. New York: Irvington.
- Oberman, L.M., Enticott, P.G., Casanova, M.F., Rotenberg, A., Pascual-Leone, A., & McCracken, J.T. (2016). Transcranial magnetic stimulation in autism spectrum disorder: challenges, promise, and roadmap for future research. *Autism Research*, 9(2), 184–203.
- Murphy, E. (2015). The brain dynamics of linguistic computation. *Frontiers in Psychology*, 6, 1515. doi: 10.3389/fpsyg.2015.01515.
- Murphy, E. (2016). The human oscillome and its explanatory potential. *Biolinguistics*, 10, 6–20.
- Murphy, E., & Benítez-Burraco, A. (2016a). Language deficits in schizophrenia and autism as related oscillatory connectopathies: an evolutionary account. *Neuroscience & Biobehavioral Reviews* <http://dx.doi.org/10.1016/j.neubiorev.2016.07.029>.
- Murphy, E., & Benítez-Burraco, A. (2016b). Bridging the gap between genes and language deficits in schizophrenia: an oscillopathic approach. *Frontiers in Human Neuroscience*, 10, 422. doi: 10.3389/fnhum.2016.00422.
- Perovic, A., & Janke, V. (2013). Issues in the acquisition of binding, control and raising in high-functioning children with autism. *UCL Working Papers in Linguistics*, 25, 131–143.
- Perovic, A., Modyanova, N., & Wexler, K. (2007). Knowledge of c-command and A-movement in children and adolescents with autism and with Asperger syndrome. Paper presented at *GALA 2007: Generative Approaches to Language Acquisition*, Barcelona, 6–8 September.
- Perovic, A., Modyanova, N., & Wexler, K. (2013). Comprehension of reflexive and personal pronouns in children with autism: a syntactic or pragmatic deficit? *Applied Psycholinguistics*, 34(4), 813–835.
- Philofsky, A., & Fidler, D. (2007). Pragmatic language profiles of school-age children with autism spectrum disorders and Williams syndrome. *American Journal of Speech-language Pathology*, 16, 368–380.
- Rojas, D.C., Maharajh, K., Teale, P., & Rogers, S.J. (2008). Reduced neural synchronization of gamma-band MEG oscillations in first-degree relatives of children with autism. *BMC Psychiatry*, 8, 66.
- Saldaña, D., & Frith, U. (2007). Do readers with autism make bridging inferences from world knowledge? *Journal of Experimental Child Psychology*, 96, 310–319.
- Sokhadze, E.M., El-Baz, A.S., Sears, L.L., Opris, I., & Casanova, M.F. (2014). rTMS neuromodulation improves electrocortical functional measures of information processing and behavioral responses in autism. *Frontiers in Systems Neuroscience*, 8, 134. doi: 10.3389/fnsys.2014.00134.
- Sohal, V.S., Zhang, F., Yizhar, O., & Deisseroth, K. (2009). Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature*, 459, 698–702.
- Stockbridge, M.D., Happé, F.G.E., & White, S.J. (2014). Impaired comprehension of alternating syntactic constructions in autism. *Autism Research*, 7(3), 314–321.
- Tager-Flusberg, H. (2006). Defining language phenotypes in autism. *Clinical Neuroscience Research*, 6, 219–224.
- Tierney, A.L., Gabard-Durnam, L., Vogel-Farley, V., Tager-Flusberg, H., & Nelson, C.A. (2012). Developmental trajectories of resting EEG power: an endophenotype of autism spectrum disorder. *PLoS One*, 7, e39127.
- Vismara, L.A., Colombi, C., & Rogers, S.J. (2009). Can one hour per week of therapy lead to lasting changes in young children with autism? *Autism*, 13(1), 93–115.
- Wada, R. (2015). Clinician recasts and production of complex syntax by children with and without specific language impairment. *All Graduate Theses and Dissertations*. Paper 4277.