<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *Journal of Semantics*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Pagin, P. (2019)

Belief Sentences and Compositionality. Notional Part

Journal of Semantics, 36(2): 241-284

<https://doi.org/10.1093/jos/ffy018>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-173059>

Belief sentences and compositionality. Notional part.*

Peter Pagin

December 12, 2018

Abstract

This paper presents an account of notional belief attributions, that is, belief attributions where the belief content is fully specified. The proposal combines a Hintikka style possible-worlds semantics for the belief operator and a structured meanings approach for giving a structured mode of presentation of the belief content. The semantics is not standard compositional, but it satisfies a more general notion of compositionality, explained in the paper. This notion, *general compositionality*, allows *semantic switching*: the semantic function relevant for an embedded term is distinct from that which applies to the term in which it is embedded. The general relation between compositionality and hyperintensional contexts is discussed in detail. In the first section, it is argued that we need to combine structured and unstructured meanings.

*The ideas of this paper were presented the first time at the LOGOS seminar at University of Barcelona in 2008, and after that on a number of occasions: in 2010 in Paris, at the third SPE (*Semantics and Philosophy in Europe*) conference, and at the Institut Jean-Nicod, in 2013, at the *Linguistics & Philosophy* workshop at the Goethe-Universität Frankfurt, and at the *Frege Puzzles* conference in Umeå, in 2014 at a guest lecture at the Ruhr-Universität Bochum, and at a guest lecture at the Pompeu Fabra University in Barcelona, in 2017 at the Computational Linguistics Conference *CompLing2017*, at Stockholm University. One reason publication took so long was that I wanted to add the treatment of *relational attributions* (*de re* and quantifying-in). In the end, the material was too much for one paper, and so only the notional part appears here.

I am indebted for comments to many people in the audiences on these occasions, and cannot do justice to all of them. With apologies to those not mentioned, I would like to thank Dave Chalmers, Lucas Champollion, Imogen Dickie, Paul Egré, Graeme Forbes, Manuel García-Carpintero, Justyna Grudzinska, Magdalena Kaufmann, Max Kölbel, Josep Macià, Louise McNally, Genoveva Martí, Olga Poller, Paul Portner, Francois Recanati, Dave Ripley, Sven Rosenkranz, Bob Stalnaker, Isidora Stojanovic, Markus Werning, Clas Weber, Dan Zeman, and Ede Zimmerman.

I am also indebted to two anonymous referees, to the editors of the journal, and very much, for discussions on many occasions through the years, to Kathrin Glüer-Pagin and Dag Westerståhl.

This research has been funded by a grant from The Swedish Foundation for Humanities and Social Sciences (Riksbankens Jubileumsfond), for the project *Interpretational Complexity*. The project has also received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement no. 675415

1. Introduction: the basic belief tension

Most of this paper is devoted to presenting a formal semantics for belief sentences, or belief attributions,¹ i.e. for sentences of the form

- (1) x believes that p .

(where ‘ p ’ is called *the embedded sentence*). In the simplest cases, the content of the belief that is attributed is exactly what the embedded sentence means (in context). I shall call these *notional attributions*.²

The proposed semantics *combines* structured meanings with possible-worlds propositions. The possible-worlds propositions will serve as *contents* of beliefs, and structured meanings as *modes of presentation* of propositions. The first part allows making use of basic ideas of possible-worlds framework for belief, as first developed by Hintikka (1962). The central idea in this framework is this:

- (H) x believes that p in w iff p is true in all worlds w' that are x -belief alternatives to w .³

In the present combined system, the left-to-right part of this biconditional still holds:

- (H⁺) If x believes that p in w , then p is true in all worlds w' that are x -belief alternatives to w .

However, the role of structured meanings in the semantics is to sometimes block the right-to-left part. That is,

¹I also, on occasion, use the expression “belief attributions” for *assertions* made by means of belief sentences. I trust that context will easily disambiguate.

²There are also more complicated attributions, where elements in the belief context contribute only by their reference, or by being anaphoric on/bound by antecedents outside the belief context:

- (i) Concerning Mary, John believes that she will come.
- (ii) Some girl is such that John believes that she will come.
- (iii) John believes that Mary will come. (Of course, John does not yet know her name.)

(i) and (iii) exemplify what is called *de re* belief attributions, and I will refer to (i) type attributions as *explicit de re* and (iii) type attributions as *implicit de re*. (ii) exemplifies what is called *quantifying into* belief contexts. I shall call these kinds *relational attributions* (the notional/relational distinction is drawn by Quine 1956, pp. 177–78, as a distinction between two kinds of belief, not, as here between two kinds of belief attributions).

The semantics for relational attributions extends the semantics for notional attributions and will be presented in a separate context.

³This is an informal and somewhat inaccurate way of presenting the principle. Strictly speaking, the belief sentence on the left-hand-side belongs to the object language, and the quantification over possible worlds on the right-hand-side to the meta-language.

(H⁻) If p is true in all worlds w' that are x -belief alternatives to w , then x believes that p in w .

does not hold. The structured meaning, which is derived from the embedded sentence of the belief attribution, filters out some ways of presenting propositions, while perhaps letting others through.

Before moving on to the formal semantics itself, I shall provide some background. The first piece of background material concerns the very need of combining structured and unstructured meanings in a single semantic framework. I'll discuss this in the present section.

We need both structured and unstructured meanings/contents. Why do we need unstructured contents? In our beliefs, we *represent* the world as being *thus and so*. In specifying a belief content one specifies what the *world is like* if the belief is accurate (and hence the belief content true). There is no need to *decode* the belief content in order to arrive at the conditions for it to be true.

This requirement is not met by any variety of structured meanings. The most common model of structured meanings in the literature has been that of nested ordered tuples (a version of which will be presented here, in section 4). As has been emphasized by several authors in recent years, ordered tuples don't *say* anything; they do not have intrinsic truth-conditions, and there is no non-arbitrary mapping from ordered tuples to propositions/truth-conditions. A tuple $\langle R, a, b, c \rangle$ of a relation R and objects a, b, c could be mapped equally well on six distinct propositions involving just these elements, so there is no privileged proposition associated with the tuple.⁴ The question how an ordered tuple can have truth-conditions has been emphasized as one aspect of the so-called problem of the "unity of the proposition" (cf. Russell 1903).

In Pagin 2016a, I generalized the argument that ordered tuples don't have a non-arbitrary mapping on propositions (truth-conditions) to a general idea of structured objects (characterized by means of Graph Theory). If that argument is correct, no type of structured meanings can serve as belief contents. "Endorsing" a structured meaning cannot be a kind of believing. At best, a belief can be associated with endorsing a structured meaning *with respect to* a particular mapping from structured meanings to propositions.

Hence, we need unstructured belief contents. Of course, this does not mean that we are justified in *identifying* propositions with sets of possible worlds (which are unstruc-

⁴See Scott Soames (2014, p. 97) and Peter Hanks (2015, pp. 42 ff.). For additional criticism of the ordered tuples model, see Bealer (1993, p. 22), Soames (2010, pp. 29–32), King (2014), Jespersen (2003).

tured). We are not. Scott Soames puts it as follows:⁵

Consider the set containing just world-states 1, 2, and 3. Is it true or false? Since the set doesn't represent anything as being one way or another, it can't be either. We could, if we wanted, interpret it as representing the actual world-state as being in the set, and so as being true iff no world-state outside the set were the unique world-state that is instantiated. But we could equally well interpret it as representing the actual world-state as not being in the set, and so as being true iff no world-state inside it is instantiated. Without interpretation by us, the set doesn't represent anything, and doesn't have truth conditions. Since propositions aren't things we interpret, but are themselves the interpretations we give to sentences, they aren't sets of world-states (Soames 2015, p. 10).

The only thing we can reasonably do, and which I think we should do, is to *model* propositions as sets of possible worlds. A set of worlds X models or represents a proposition p *with respect to membership* just in case p is true at a world w iff $w \in X$.

You might think that the need to appeal to the modeling relation rather than to claim identity would make it equally acceptable to model propositions by structured meanings. That is not correct, however. Sets of worlds provide a good model of propositions just because there is a plausible 1 – 1 correlation between sets of worlds and propositions with respect to truth. For any set of worlds there is a proposition such that they agree in truth value (with respect to the stipulation) at any world. This condition is not met by structured meanings, which are more fine-grained. This speaks strongly in favor of the sets-of-worlds model.

Does it also speak for the sets-of-worlds model against competing unstructured models, such as sets of world-time pairs, sets of centered worlds (world-time-subject triples), or sets of indices with more ingredients (standards of taste, standards of precision ...)? Intuitions about the typing of contents of beliefs and utterance may not be decisive here (cf. Stojanovic 2012). However, what does count for the classical choice of the sets-of-worlds model is that with a world as the index of evaluation, the index for an *unembedded* content is always *the same*: the actual world. As subjects we need not keep track of shifts of index.

This requirement is not met in the centered-worlds model. Since one element in the center is the *subject*, and what corresponds to belief is a self-ascription of a property (Lewis 1979a), there will be problems of communication: the speaker ascribes the property to herself while the hearer accepts the *same* content by ascribing it to himself, not to

⁵I argued essentially the same way in Pagin 2016a. Identifying propositions with characteristic functions is equally flawed.

the speaker.⁶

Analogously, if contents are *temporal* propositions, true or false with respect to world-time pairs, then we cannot *represent* times in our belief contents. We can only apply beliefs to the time of having them. This has a crippling effect on the possibility of preserving knowledge across time.⁷

Finally, there is the question whether there are *unstructured* contents that are *more fine-grained* than truth-conditions. Fregean senses have been taken to be such contents. It is often said that Fregean senses are intrinsically *representational* but nevertheless more fine-grained than truth-conditions (especially possible-worlds truth-conditions). But it is unclear to me what a *representational* content can be that is not captured in terms of truth-conditions. Standard coreference cases, of the ‘Hesperus’/‘Phosphorus’ kind, are resolved by appeal to difference in sense, but the kind of content that is added to reference in such examples, is just truth-conditional in nature.⁸

These considerations point to the conclusion that a) we need unstructured contents, b) these contents are best modeled by sets of possible worlds.

The properties of belief mentioned above can be summarized in three theses:

(CONTENT) Beliefs have content.

(TRUTHCOND) Belief contents have truth conditions.⁹

(DETERMINE) Belief contents are determined by their truth-conditions.

Let’s turn to the need for structured meanings. It is a consequence of the three principles above that if two sentences ‘ p ’ and ‘ p' ’ have the same truth-conditions, they determine the same content. Since what we believe are propositions, and since a belief *sentence*, of the form (1), seems to specify exactly what proposition a subject believes, sentences that express the same proposition should be intersubstitutable in belief *contexts*. That is, we should have the result that the two belief attributions

⁶This was first pointed out in Stalnaker 1988, revived in Egan 2007 and discussed by others after that. I discuss it in Pagin 2016b.

⁷The view called *Temporalism*, defended by David Kaplan (1989) and Brit Brogaard (2012) is the view that our semantics needs both temporal and classical propositions. The argument above only concerns the situation where we have access to temporal propositions only. For the mixed view, other arguments are needed, and have been provided and discussed, from Richard 1981 onwards.

⁸Mathematical examples are more difficult, since intuitive differences in sense are not captured by differences in possible-worlds truth-conditions. They are, however, captured by differences in structure. But see the discussion of synonymy and atomic concepts in section 6.

⁹You may want to say instead that *beliefs* are true or false, and that belief contents are the truth-conditions of belief, rather than entities that have truth-conditions. I’ll take this to be a terminological issue. Here, I’ll take a belief to be mental state of a subject and to be the *vehicle* of a content rather than individuated by the content itself. I sometimes say that a belief is “true”, but that should be seen as a shorthand.

- (2) a. x believes that p .
- b. x believes that p' .

have the same truth value. But, as has been a theme in analytic philosophy from the very start (Frege 1892), this need not be the case. A run-of-the-mill example is

- (3) a. John believes that pelicans fly.
- b. John believes that: pelicans fly and $\sqrt{529} = \pm 23$

Intuitively, John might well believe that pelicans fly without believing that $\sqrt{529} = \pm 23$. But, intuitively again, since the truth-conditions of the latter is satisfied at all indices, *that pelicans fly* is the same proposition as the proposition *that pelicans fly and $\sqrt{529} = \pm 23$* .

If we accept this truth value distribution ((3a) true, (3b) false), and also accept that a belief sentence is *true* iff the attributee has a belief with the content specified in the embedded sentence, then we have to conclude that the contents specified in (3a) and (3b) are different.¹⁰ However, if the embedded sentences have the same truth-conditions, it then follows that the contents they specify are *not* determined by their truth-conditions. Then, what we believe are *not* truth-conditionally determined contents.

We have a contradiction between our default, perhaps naive, ideas about the contents of beliefs (that they are truth-conditional) and the result of substitution problems. There are basically four alternative ways of resolving this contradiction:¹¹

- 1) Give up the naive idea about belief contents.
- 2) Give up the idea that the semantics of belief sentences has a close connection with belief contents.
- 3) Provide a semantics for belief sentences that allows the substitutions to go through.
- 4) Provide a semantics for belief sentences where truth and falsity are sensitive to how the proposition is expressed by the belief context sentence.

¹⁰John Perry (1977, pp. 475–7) proposes that Frege accepted as a *criterion of difference of sense* that

(Dif) If A understands S and S' , and accepts S as true while not accepting S' , then S and S' have different senses.

¹¹One further option, not discussed in the main text, that does not easily fit into either of these slots, is that of invoking *impossible worlds* (cf. Nolan 2013). By adding mathematically impossible worlds you can distinguish between the propositions attributed in (3a) and (3b), in that there is a world where it is true that pelicans fly but false that $\sqrt{529} = \pm 23$. This is a desirable effect. However, to do all the work desired within a (general) compositional semantics, we will also need logically impossible worlds, and in the end we will either have to make the semantic clauses relative to worlds (against the very idea of a possible-worlds semantics) or change the logic of the meta-language (which can't really be what the semantics for belief-sentences would justify).

An appeal to Fregean *senses* would fall under alternative 1). Another idea that falls under this option is making structured meanings the objects of belief. In his discussion of the corresponding problems of ‘says’, Max Cresswell (1985, p. 116) suggests that one type of possible objects of *says* are structured meanings. Cresswell uses the ordered-tuple model of structured meanings.¹² I have argued above, however, that structured meanings do not have intrinsic truth-conditions and cannot be the contents of belief.

I take alternative (2) to be a last resort. Our basic grasp of what belief is derives from accepted belief attributions, together with first-person experiences. Choosing option (2) would at least run a high risk of changing the topic.¹³

Alternative (3) has been defended in the literature, notably by Nathan Salmon (1986). Salmon appeals to *modes of presentation* of propositions. These are more fine-grained than propositions. What exactly they are, over and above playing that role, is not relevant to the analysis, since Salmon simply quantifies over them. On Salmon’s analysis, the belief sentence (3a) is true iff John stands in the belief relation to the proposition *that pelicans fly* under *some* mode of presentation. Then, since exactly the same proposition is presented in the belief context of (3b), if (3a) is true, so is (3b). The intuition that they differ in meaning is explained by appeal to pragmatics.¹⁴

The final alternative, (4), has been chosen by a number of authors through the years (e.g. Carnap (1947), Kaplan (1968), Burdick (1982), Forbes (1996)). The shared feature of such theories is that the embedded sentence *itself* is what the truth value of the attribution is sensitive to. The sentence itself plays the role of mode of presentation of the proposition it expresses. This makes sense, since typical differences between embedded sentences, such as between those in (3a) and (3b), seem to be crucial for the perceived difference in truth value.

The flip side of the coin is that some differences between sentences do not seem to matter. Consider

- (4) a. John believes that pelicans rarely fly.
- b. John believes that pelicans seldom fly.

¹²David Lewis (1970) provided a theory with meanings structured as syntactic trees. Larson and Ludlow (1993) provided an account where so-called *interpreted logical forms*, tree-structured entities, are the objects of belief. They combine structure and linguistic items as elements of belief content.

¹³An extreme version of this alternative would be to deny any connection between the meaning of the embedded sentence and the content of the belief attributed. I don’t know of anyone taking this option. A less extreme version is a theory according to which *pragmatics* adds ingredients to content attributions that go beyond what the meaning of the embedded sentence provides, even given contextual values to indexicals. A theory of this kind is proposed by Mark Crimmins (1992). A discussion of Crimmins’s view must be deferred to another occasion.

¹⁴Further discussion of Salmon’s views must be deferred to another occasion.

(4a) differs from (4b) only by containing the adverb ‘rarely’ where the latter contains ‘seldom’. These two words are as synonymous as adverbs get. The difference between the two belief sentences cannot plausibly be one that would have an impact on truth value: it is the very same belief that is attributed.¹⁵ Theories that let sentences play the role of modes of presentation then face a dilemma. Either they incorrectly predict that any difference between the embedded sentences of two belief attributions *can* induce a difference in truth value, or they simply leave it open which can and which cannot. In the latter case, the theory is simply incomplete. We want an answer to the question which sentence properties can and which cannot induce a truth value change.

Another question that arises for such theories is that the embedded sentence may not even belong to a language spoken by the attributee, and it is then unclear how it influences the truth value of the belief sentence. Clearly, it is not by directly corresponding to a mode of presentation of the content entertained by the *attributee*. Only the mental state of the attributee should matter for truth and falsity, and so the mere fact that two embedded sentences differ for the *speaker* indicates nothing about whether the choice between them has any bearing on truth and falsity.¹⁶

Carnap’s account (1947/1956, pp. 61–2) offers an answer to both questions. Here (b) is the analysis of (a):

- (Carnap) a. John believes that D
b. There is a sentence $\mathfrak{S}_i$ in a semantical system S' such that (a) $\mathfrak{S}_i$ is intensionally isomorphic to ‘D’ and (b) John is disposed to an affirmative response to $\mathfrak{S}_i$.

¹⁵This is a controversial issue. In particular, Tyler Burge (1978) has taken the opposite view. I’ll return to the question of belief and synonymy. If you want to go with Burge here, change the example so that the difference between the two sentences does not even depend on substitutions of words. It may e.g. derive from the difference between ‘rarely’ italicized and ‘rarely’ underlined, or even from just a change in typeface.

The account in Forbes 1996 suffers from this problem (that intuitively insignificant differences between sentences can affect truth value), in our view. The sentence (4a) is analyzed into (4b) (Forbes 1996, p. 351, informally rendered)

- (i) a. Lois believes that Superman is an extraterrestrial.
b. The situation type of Superman’s being an extraterrestrial is such that Lois believes her *so-labeled* way of thinking of it.

Here ‘so-labeled’ is a logofor, referring back to the *description* of the situation type. Forbes develops this idea technically. The problem with the appeal to descriptions is that the grain becomes too fine. Choosing between two simple synonyms should not make a difference for the truth of an attribution.

This effect can be reduced by requiring only a *relation* between the expression used and the expression relevant to the believer (as in Forbes 1990), but such accounts suffer from lack of information (cf. discussion below of Carnap’s analysis) or lack of predictive power (cf. the discussion of Richard’s analysis in note 31).

¹⁶An answer to this question is worked out in the account of Mark Richard (1990). For discussion of Richard’s account, see section 5.

The relation of intensional isomorphism between linguistic expressions is roughly as follows: two atomic expressions are intensionally isomorphic iff they are co-intensional, and two complex expressions are intensionally isomorphic iff they have the same structure and corresponding parts are intensionally isomorphic. The formal definition is given in Carnap 1947/1956, p. 59, where Carnap uses his technical notion of L-equivalence rather than co-intensionality.

Carnap's account is importantly similar to the account to be proposed here, since with respect to substitution conditions, the appeal to structured meanings is equivalent to the appeal to intensional isomorphism. This does not, however, make the two accounts equivalent, in particular since Carnap's analysis does not *say* what the belief content is, but only *mentions* the embedded sentence. The loss of information is stressed by Alonzo Church (1950), who brings out the point by appealing to the Langford translation test. Basically, the objection is the following: 'D' is an English sentence in (a) and (b). In a translation to German, it should be translated in (a), since it is *used* there, but not in (b), since it is only *mentioned* there by being quoted, in order to preserve truth value. But that means that a German speaker who does not understand English will not, despite knowing that the German translation of (b) is true, know what John believes, in contrast to knowing the truth of the translation of (a). Church finishes up with an extra twist: if we relax the translation relation by way of translating quoted expressions as well as used expressions, this problem is eliminated, but then a German translation will not be intensionally isomorphic to the English original, or to a translation into French, since the translations in the different languages contain reference to *different* sentences, the respective translations of 'D'.¹⁷

In order to get around the quotation problems, the type of entity that plays the role of mode of presentation must itself be *non*-linguistic but still sensitive to relevant linguistic differences. Structured meanings do make up such a type of entity. They reflect the relevant linguistic structure but do not inherit irrelevant properties from the sentences.

The problem with earlier appeals to structured meanings is that, for the most part, they have taken the structured meanings themselves as contents of belief (which I argued against above) and that they haven't been adequately or even systematically related

¹⁷ Compare 'The word 'dog' has three letters.', translated into German as 'Das Wort 'Hund' hat drei Buchstaben.', where a truth is translated into a falsehood. Putnam (1954, pp. 115–6) objects to Church that two analyses can be regarded as correct even if they are not intensionally isomorphic. This is certainly true in some cases, but not here. Let *s* be an English sentence and *s'* its translation into German, with meaning preserved. Let *r* be an English sentence that gives the meaning of *s*, and *r'* its translation into German, with meaning preserved. Then *r'* should also give the meaning of *s'*, since *s* and *s'* have the same meaning. But this condition is not met under the relaxed translation relation. This answers Putnam. Davidson (1963, p. 344) also finds Church's requirement of intensional isomorphism "unreasonable", but it is not clear why (cf. Pagin 2017a).

to *un*-structured meanings. Typically, providing structured meanings in the semantics leaves unstructured meanings out, and we are left with the shortcomings of the former.¹⁸

The current system, by contrast, which is an application of *Switcher Semantics*¹⁹ allows several semantic functions within the same semantic system, and controls the switching between them. Unstructured meanings will be taken as basic, and structured meanings will be called upon when needed: for what occurs in the belief context. The structured meanings will play a double role, on the one hand that of being related to the mental state of the attributee, and on the other that of determining the corresponding unstructured meaning. The latter does its job in the Hintikka style possible-worlds framework. Accordingly, I'll here take propositions, i.e. belief contents, to be individuated like classical possible-worlds propositions. That is, p and q are *the same* proposition iff they have the same truth value at every possible world.

The rest of the paper is organized as follows. In section 2, I'll give a brief overview of issues of compositionality in relation to belief sentences. In section 3, I present the idea of switcher semantics and *general compositionality*, a generalization of (standard) compositionality. In section 4, I present the current version of structured meanings. In section 5, I provide the formal semantics for notional attributions, in which possible-worlds semantics, structured meanings, and general compositionality are combined, and in section 6, I comment on some of its properties. Section 7 offers some concluding methodological remarks. In the appendix, David Chalmers's account is discussed and compared with the current proposal.

2. Belief sentences and compositionality

The principle of compositionality is typically stated in roughly the following way:

(CFI) The meaning of a complex expression is a function of the meanings of its parts and its mode of composition.

(CFI) is an *informal* statement of the *functional* version of (standard) compositionality.²⁰ The word 'meaning' that occurs in it does not refer to any particular type of se-

¹⁸Cresswell (1985, p. 116) actually allows 'says' to take *both* structured and unstructured meanings as complement denotations (claiming that this does not lead to ambiguity). Cresswell clearly acknowledges that both types are relevant as complement denotations, but not that they both are needed all the time. The interpreter freely chooses one complement at the expense of the other, and the speaker is given no syntactic means of indicating which is intended. For a detailed criticism of Cresswell's views, see Richard 1990, pp. 21–9.

¹⁹In Switcher Semantics, an operator may switch semantic function for one or more of its arguments, which is exemplified in Frege's idea of indirect reference, triggered in the second argument of attitude and indirect discourse operators. Cf. section 3.

²⁰We shall consider formal versions in the next section.

mantic value, but is placeholder for terms that do. For any particular type of value, we can ask whether that type, with respect to some particular interpreted language, satisfies the principle, or in short, is compositional. For example, is *extension* as a semantic value compositional for English? That is, is the following special principle true:

(CE_E) The extension of a complex expression in English is a function of the extensions of its parts and its mode of composition.

In this section, we shall assume that we start out with a fragment of English that is extensional, i.e. for which extension satisfies (CFI). To this fragment we add ‘believes that’. The only source of failure of compositionality for extension will therefore be the *believes-that* constructions. We shall then look at the character of compositionality failures for extension, intension, and intuitive synonymy. In order to carry this out, we need to consider the *substitution* version of compositionality:

(CSI) If in a complex expression A a constituent e is replaced by a constituent e' that has *the same* meaning as e , then the meaning of the new expression $A' = A[e'/e]$ has the same meaning as A .

(here ‘ $A[e'/e]$ ’ denotes the result of substituting e' for e , in at least one occurrence, in A). Under certain assumptions (e.g. that every grammatical expression is meaningful), the function and substitution versions of compositionality are equivalent.²¹ It is the substitution version that we will apply below.

With *extension* as the semantic value, consider the following pair of sentences:

- (5) a. John believes that Mark Twain is a novelist.
- b. John believes that Samuel Clemens is a novelist.

The embedded sentences in (5) (‘Mark Twain is a novelist’, ‘Samuel Clemens is a novelist’) contain only extensional constructions. So, any failure of compositionality in (5) derives from the belief sentence construction.

Moreover, the proper names in (5), ‘Mark Twain’ and ‘Samuel Clemens’, are corefering. Since everything else is the same, the two embedded sentences have the same truth value, hence the same extension. However, in accordance with an almost universally accepted intuition, since Frege, (5a) and (5b) themselves may differ in truth value. A typical scenario would be one where John is the neighbor of Clemens and knows him as such, without knowing that this same neighbor is also the author Twain. John knows by description that Twain is the author of *Huckleberry Finn*. It is then natural to say that (5a)

²¹Cf. Hodges 2001.

is true and (5b) is false.

With (5a) true and (5b) false, *extension* is not compositional for the belief-fragment of English.

Now, consider again the pair of co-intensional belief attributions:

(3a) John believes that pelicans fly.

(3b) John believes that: pelicans fly and $\sqrt{529} = \pm 23$

Because of miscalculation, John might believe that $\sqrt{529} = \pm 22$, and hence not believe the conjunction attributed to him in (3b). Then, it is natural to say that (3a) is true and (3b) false. Since the embedded sentences are co-intensional, it follows that *intension* is not compositional with respect to the belief fragment of English: this fragment is *hyper-intensional*.

Finally, consider again pairs of intuitively synonymous sentences:

(4a) John believes that pelicans rarely fly.

(4b) John believes that pelicans seldom fly.

If, in accordance with what Burge has claimed, it may be that (4a) is true and (4b) false, then *intuitive synonymy* is not compositional with respect to the belief-fragment of English. Burge's 1978 example uses the synonymy pair 'fortnight' and 'period of fourteen days', and claims that someone can truly say

(6) For years I believed that a fortnight was ten days, not fourteen, though of course I never believed that fourteen days were ten days.

(Burge 1978, p. 126). Hence, according to Burge, among

(7) a. John believes that a fortnight is a period of ten days

b. John believes that a period of fourteen days is a period of ten days.

(7a) may be true, while (7b) will remain false.²²

Where does this leave us for the prospects of a semantics for the belief fragment of English? Clearly, the problems for the compositionality of *extension* are not problems for

²²Considerations of synonymy-pairs in belief contexts go back to Benson Mates (1950), who used problems with substitutivity to argue against synonymy. Jeff Pelletier (1994b) accepts synonymy and uses the problems to argue against compositionality. A similar line was taken much earlier by Hilary Putnam (1954), but neither of Putnam's principles (g) and (h) (Putnam 1954, p. 118) is equivalent to standard compositionality, the former less and the latter more restrictive. Saul Kripke (1979, pp. 160–1) basically takes the stance that the disquotation principle is not applicable in Mates cases, much in line with the present view. Thanks to an anonymous reviewer for drawing my attention to Kripke's remarks.

intensional semantics, such as a standard possible-worlds semantics.

The problems for the compositionality of *intension* are the problems I propose to deal with in this paper. They cannot indeed be strictly speaking *solved*, since we don't get around the truth value distributions for examples like (3).²³ But the combination of structured meanings and semantic switching will turn out to satisfy the conditions of *general compositionality*. We turn to this topic in the next section.

I will come back to belief and synonymy in section 6. I shall argue that Burge's proposed truth value distribution should not be accepted strictly speaking, but that the Burgean intuition can be accommodated under a non-standard syntactic derivation.

3. General compositionality and switcher semantics

In order to introduce semantic switching and general compositionality, we must first provide a formal statement of *standard* compositionality. The formal framework used here is basically the algebraic one that was introduced by Wilfrid Hodges (2001) and is presented in Pagin and Westerståhl 2010a.²⁴

We assume given a *language* L and a domain of *meanings* M . The language is primarily, but only partially, a set of *expressions* E'_L together with a grammar, Σ'_L . Σ'_L is understood as a set of *syntactic operations*. There is a subset At'_L of E'_L that is the set of *atomic* expressions that are given independently of the operations. Formally, the language L is, partially, an *expression algebra* (E'_L, At'_L, Σ'_L) . A member $\sigma' \in \Sigma'_L$ is, for some n , a partial map $E'^n_L \rightarrow E'_L$, from n -tuples of expressions to expressions. E'_L is the *closure* of At'_L under Σ'_L . That is, the expressions of L are *generated* from the atomic expressions by the grammar.

Typically, a set of complex expressions contains many that are syntactically ambiguous, i.e. can be generated in more than one way by the grammar from the atomic expressions, and typically the semantic interpretation depends on which way is selected. This means that we cannot define a semantic function (interpretation function) directly on expressions, but only on *disambiguated expressions*. What plays the role of disambiguated expression can be different things, such as analysis trees (in Montague). Hodges uses *grammatical terms*, or *terms* for short.²⁵ Grammatical terms are built from names of expressions and syntactic operations, in a way that reflects the application of the latter to the former. To take an example, consider (8a), with the phrase structure of (8b).

²³Montague's semantics for belief contexts in (1973) is inadequate in that they are treated as intensional, not hyperintensional.

²⁴It has been applied in Pagin 2003 as well as in Pagin and Westerståhl 2010b.

²⁵Hodges distinguishes between terms, which may contain *parameters*, and grammatical terms, which don't. I don't make this distinction here, although I include terms with free but bindable variables.

- (8) a. John saw the dog
 b. [_S John_{NP} [_{VP} saw_{VT} [_{NP} the_{Det} dog_N]]]

(8b) indicates three syntactic operations: a) the Determiner + Noun (Det+N) combination, b) the Verb (transitive) + Noun Phrase combination (VT+NP), and c) the Noun Phrase + Verb Phrase combination (NP+VP).

Let A' be the set {'John', 'saw', 'the', 'dog'} and Σ' the set $\{\alpha, \beta, \gamma\}$, corresponding to the rules for forming (Det+N), (VT+NP), and (NP+VP), respectively, of example (8). The *names* of these operations are exactly ' $\bar{\alpha}$ ', ' $\bar{\beta}$ ', and ' $\bar{\gamma}$ '. Then the grammatical term that corresponds to the phrase structure analysis (8b) is

$$(9) \quad \bar{\gamma}(\text{'John'}, \bar{\beta}(\text{'saw'}, \bar{\alpha}(\text{'the'}, \text{'dog'}))).^{26}$$

The grammatical terms form their *own* algebra the *grammatical term algebra* (term algebra) G_L . This is a triple (T_L, A_L, Σ_L) , where T_L is the set of grammatical terms, A_L the set of *atomic* grammatical terms, and Σ_L the set of operations mapping n -tuples of grammatical terms on grammatical terms: a member $\sigma \in \Sigma$ is, for some n , a partial map $T_L^n \rightarrow T_L$. And T_L is the closure of A_L under Σ_L .

Here, we shall take G_L as the *basic* part of the grammar for a language L . From G_L will be derived *both* the meanings of grammatical terms and the linguistic expressions that they represent. There will be a *semantic function* μ that maps terms on meanings and an *expression value* function V that maps terms on expressions. That is, the algebraic syntax and semantics has the simple structure shown in Figure 1. The *full* specification of the language syntax is provided only by the term algebra, the expression algebra, and the V function.

Typically, atomic expressions are also atomic terms. In case of lexical ambiguity, however, we have distinct terms corresponding to the same expression, such as the terms 'bank₁' and 'bank₂' which correspond to the same atomic expression 'bank'. The V function maps simple terms on simple expressions, and it maps names of operations in the expression algebra on those operations. In the recursive step, where $t_1, \dots, t_n$ are grammatical terms, σ' an expression algebra operation, and ' $\bar{\sigma}'$ ' a name of σ' , and where ' $\bar{\sigma}'(t_1, \dots, t_n)$ ' is a complex grammatical term, we have

$$(V) \quad V(\bar{\sigma}'(t_1, \dots, t_n)) = V(\bar{\sigma}')(V(t_1), \dots, V(t_n)) = \sigma'(V(t_1), \dots, V(t_n))$$

In the simplest cases, a syntactic operation is mapped on surface structure simply as

²⁶This is an illustration. It might be preferable e.g. to have as top operation one that combines the term that corresponds to 'the dog' _{i} with the term that corresponds to 'John saw i ', analogous to a Montagovian quantification rule. Of course, I have left out the tense operation.

Figure 1: Algebraic syntax and semantics

concatenation with word spaces:

$$(10) \quad V(\overline{\sigma'}(t_1, t_2)) = V(t_1) \frown \sqcup \frown V(t_2)$$

where ' $\sqcup$ ' marks a word space. We get, after three applications of (10),

$$(11) \quad V(\overline{\gamma}(\text{'John'}, \overline{\beta}(\text{'saw'}, \overline{\alpha}(\text{'the'}, \text{'dog'})))) = \text{'John'} \frown \sqcup \frown \text{'saw'} \frown \sqcup \frown \text{'the'} \frown \sqcup \frown \text{'dog'}$$

The right-hand side of (11) is a structural-descriptive name, in Tarski's sense, of 'John saw the dog'.

In the grammatical term algebra, corresponding to an operation σ' in the expression algebra, there is an operation σ that is defined for the terms $t_1, \dots, t_n$ just in case σ' is defined for $V(t_1), \dots, V(t_n)$. If it is, it maps $t_1, \dots, t_n$ on exactly the complex term ' $\overline{\sigma'}(t_1, \dots, t_n)$ '. This means that V is a *homomorphism* from the term algebra to the expression algebra:

$$(V^*) \quad V(\sigma(t_1, \dots, t_m)) = \sigma'(V(t_1), \dots, V(t_m))$$

No abstract limitations are imposed on V (such as that the value of the first argument should appear in the expression to the left of the value of the second argument), although a number of restrictions will be empirically motivated, in order that the hearer be able to parse the string. The algebraic framework as such is very general; it can be adapted to phrase structure rules as well as to transformations or movements.

The *semantic* function μ is defined on the grammatical terms, mapping terms on entities in a domain M of semantic values. Given the algebraic framework we can now state a more formal counterpart to (CFI):

$$(CFF) \quad \text{For every } n\text{-ary syntactic operation } \sigma \in \Sigma \text{ there is a meaning operation } r_\sigma : M^n \longrightarrow$$

M , such that for all grammatical terms $t_1, \dots, t_n$ where $\sigma(t_1, \dots, t_n)$ is defined and μ -meaningful, it holds that $\mu(\sigma(t_1, \dots, t_n)) = r_\sigma(\mu(t_1), \dots, \mu(t_n))$.

If (CFF) holds for a semantic function μ , we say that μ is (functionally) compositional. Here μ is of course left unspecified. It is clear from (CFF) that standard compositionality is concerned with the *immediate* constituents, i.e. here the terms that are the arguments to the syntactic operation.

The restriction to immediate constituents does not hold for the *substitution* version. Rather, it is concerned with substitutions at any level of the grammatical term, provided only that the substitutions do not overlap. We use the following notation: the expression ' $s[\xi_1, \dots, \xi_n]$ ' denotes the term s together with n non-overlapping subterm *positions* $\xi_1, \dots, \xi_n$ in s . These can be positions anywhere in s . The expression ' $s[t_1/\xi_1, \dots, t_n/\xi_n]$ ' denotes the term s' which is like s except by possibly differing in containing t_i in position ξ_i , $1 \leq i \leq n$. One and the same term can occur several times in s , and hence also in the list. Substituting a subterm u for a subterm t , in a position ξ , is exactly the move from $s[t/\xi]$ to $s[u/\xi]$. That is, $(s[t/\xi])[u/\xi] = s[u/\xi]$. More on positions below.

The formal substitutional version can then be stated as follows:

(CSF) If $s[t_1/\xi_1, \dots, t_n/\xi_n]$ and $s[u_1/\xi_1, \dots, u_n/\xi_n]$ are both μ meaningful terms, and if $\mu(t_i) = \mu(u_i)$, for $1 \leq i \leq n$, then $\mu(s[t_1/\xi_1, \dots, t_n/\xi_n]) = \mu(s[u_1/\xi_1, \dots, u_n/\xi_n])$.

For a total semantic function μ (defined for all grammatical terms), (CFF) and (CSF) are equivalent (Hodges 2001). For a partial semantics, (CFF) can fail even though (CSF) holds. The reason is that not all immediate subterms of a μ meaningful term need be themselves μ meaningful. The substitution version still goes through, since we only consider subterms for substitution that *are* μ meaningful. The two principles are still equivalent even for a partial semantics if the *Domain principle* holds for μ : every immediate subterm of a μ meaningful term is itself μ meaningful (Hodges 2001).

Thereby we have characterized standard compositionality. *Switcher semantics* generalizes standard forms of semantics by allowing semantic systems with a finite number of semantic functions. Such a system $\mathcal{S}$ has a set of semantic functions $S = \{\mu_0, \mu_1, \dots, \mu_n\}$. In any system $\mathcal{S}$, one member μ_0 in S is the *designated* semantic function of $\mathcal{S}$. To be the designated member means that the value $\mathcal{S}(t)$ that the system assigns to the *unembedded* occurrence of t is exactly $\mu_0(t)$. That is to say, the interpretation of an unembedded term always *starts* with applying the designated member.

What characterizes switcher semantics in particular is that the value of a semantic function μ_i for a complex term $\sigma(t_1, \dots, t_n)$ may depend on values $\mu_{j_1}(t_1), \dots, \mu_{j_n}(t_n)$ that *other* members μ_{j_k} of S give to the subterms $t_1, \dots, t_n$. None of μ_{j_k} , $1 \leq k \leq n$, need be

μ_i itself. In case such a switch takes place, say at argument k of σ , we say that the pair (σ, k) is a *switcher*. Switcher semantics is the framework for handling so-called *ingredient sense*, and in fact *multiple ingredient sense*.

Now, it may be that in a set S of semantic functions, no member of S is standard compositional, and yet the functions in S *interlock* to the effect that every value of any function μ for a complex term $\sigma(t_1, \dots, t_n)$ is determined by the values of members of S for $t_1, \dots, t_n$, where at least some of those members are distinct from μ itself.

More precisely, in such a case there is a *selection function* Ψ that represents the interlocking of the members. Given a term algebra $G_L = (T, A, \Sigma)$, and a set of semantic functions $S = \{\mu_0, \mu_1, \dots, \mu_n\}$, a selection function is a function $\Psi_S : S \times \Sigma \times \mathbb{N} \rightarrow S$. Ψ_S takes as arguments triples (μ_i, σ_j, k) , where $\mu_i \in S$, $\sigma_j \in \Sigma$, and $1 \leq k \leq n$, where n is the arity of σ_j , and gives as value a member of S .

Given a term algebra G_L , a semantic system $\mathcal{S}$ defined over G_L is a structure $\mathcal{S} = (S, \mu_0, \Psi)$, where $S = \{\mu_0, \mu_1, \dots, \mu_n\}$ is a set of semantic functions, μ_0 is a designated member of S , and Ψ is a selection function defined on G_L and S .

We can now state what it is for a semantic system $\mathcal{S}$ to be *general compositional*:

(GCF) Given a term algebra $G_L = (T, A, \Sigma)$, a semantic system $\mathcal{S} = (S, \mu_0, \Psi)$ is *general compositional* iff it holds for every pair (μ_i, σ_j) , where $\mu_i \in S$ and $\sigma_j \in \Sigma$, that there is a meaning operation $r_{i,j}$ that for any terms $t_1, \dots, t_n$, provided μ_i is defined for $\sigma_j(t_1, \dots, t_n)$, satisfies

$$\mu_i(\sigma_j(t_1, \dots, t_n)) = r_{i,j}(\mu_{h_1}(t_1), \dots, \mu_{h_n}(t_n))$$

where $\Psi(\mu_i, \sigma_j, k) = \mu_{h_k}$, for $1 \leq k \leq n$.

This is the *functional* version of general compositionality. For it to be true of a semantic system $\mathcal{S} = (S, \mu_0, \Psi)$, a *Generalized Domain principle* must hold:

(GDP) For any $\mu_i \in S$, any $\sigma_j \in \Sigma$, and any $k \leq n$, where n is the arity of σ_j , if $\sigma_j(t_1, \dots, t_n)$ is μ_i meaningful and $\Psi(\mu_i, \sigma_j, k) = \mu_h$, then t_k is μ_h meaningful.

Note that (GDP) may be true of a system $\mathcal{S}$ even if the (non-generalized) domain principle isn't true of all or even any member of S .

If $S = \{\mu_0, \mu_1, \dots, \mu_n\}$ is part of a general compositional system, then the functions in S can be defined together by *mutual recursion*. The recursion clauses are exactly those we get from compositional equations in (GCF), once we substitute the actual values of Ψ for each instance. Each member in S is directly defined for simple terms, and for a complex term $\sigma(t_1, \dots, t_m)$ it is defined, if at all, in terms of the values of (possibly) other members

of S for the immediate subterms t_i . Note that once such a system of recursion clauses has been provided, Ψ is completely defined; it is nothing that comes *over and above* a semantics.

For the substitution version of general compositionality, we need to take account of which member of S is relevant for a position ξ in a term s . A position is here treated as a finite sequence ξ of zero or more pairs (σ, k) of an operation σ and a argument place k (within the arity of σ). $\xi \frown (\sigma, k)$ is the extension of the sequence ξ by the addition of (σ, k) . Given a semantic system $\mathcal{S}$ and an inductively defined set of positions X , we define the *extended selection function* $\Psi^* : X \longrightarrow S$ as follows:

- (Ψ^*) i) $\Psi^*(\langle \rangle) = \mu_0$
 ii) $\Psi^*(\xi \frown ((\sigma_j, k))) = \Psi(\Psi^*(\xi), \sigma_j, k)$.

Any term t is a subterm of itself, and occurs at its own $\langle \rangle$ position. We introduce a function T from terms and positions to terms:

- (T) i) $T(t, \langle \rangle) = t$
 ii) If $t = \sigma(t_1, \dots, t_n)$, and $T(s, \xi) = t$, then $T(s, \xi \frown \langle (\sigma, i) \rangle) = t_i$, $1 \leq i \leq n$.

It follows by straightforward induction over the length of sequences that if $T(s, \xi) = t$, then $T(s, \xi \frown \xi') = T(t, \xi')$, provided $T(t, \xi')$ is defined.

The formal substitution version can now be stated as follows:

- (GCS) Given a term algebra $G_L = (T, A, \Sigma)$, a semantic system $\mathcal{S} = (S, \mu_0, \Psi)$ has the *general compositional substitution property* iff for any instantiated arity n , any term v such that $T(v, \xi) = s$, and $\Psi^*(\xi)$ meaningful terms $s[t_1/\xi_1, \dots, t_n/\xi_n]$ and $s[u_1/\xi_1, \dots, u_n/\xi_n]$, if

$$\Psi^*(\xi \frown \xi_i)(t_i) = \Psi^*(\xi \frown \xi_i)(u_i), \quad 1 \leq i \leq n,$$

$$\text{then } \Psi^*(\xi)(s[t_1/\xi_1, \dots, t_n/\xi_n]) = \Psi^*(\xi)(s[u_1/\xi_1, \dots, u_n/\xi_n]).$$

As a special case, this holds if $\xi = \langle \rangle$, and hence $s = v$. The greater generality is needed for proving the following fact:

Fact 1. If (GDP) is true, then (GCF) and (GCS) are equivalent.

The proof from (GCS) to (GCF) is immediate, since (GCF) is the special case of immediate constituents. The proof from (GCF) to (GCS) proceeds by induction over the complexity of s .

Standard compositionality is the special case of general compositionality where S has only one member. It is of course automatically the designated member of the system. No switching occurs, and so Ψ is a constant function.

Let's say that in a semantic system $\mathcal{S} = (S, \mu_0, \Psi)$, a function μ is a *non-vacuous* member of S iff there is at least one term $s[t/\xi]$ such that $\Psi^*(\xi) = \mu$. That is, μ actually applies. We can now define switcher semantics. A *switcher semantics* is a general compositional semantic system $\mathcal{S} = (S, \mu_0, \Psi)$, where S has at least one non-vacuous non-designated member. Hence, in a switcher semantics, switching occurs.

Historically, Frege is probably the first semanticist to apply switching. The idea of *indirect reference*, as specified in Frege 1892, is the idea that in certain contexts, *indirect contexts*, an expression takes a different semantic value, the indirect reference, which is distinct from its default, ordinary reference. It also happens to be its ordinary *sense*. The switchers here are the belief and indirect discourse operators, with respect to their content argument (second position, if the first is the attributee argument position). Possibly, Frege's semantics is still not a switcher semantics, because possibly, the set of semantic functions is not finite. On many interpretations of Frege, the hierarchy of indirect references (and senses) contains infinitely many members.²⁷

Frege (1892) also applied the switching technique to the case of quotation. The idea was that an expression in a quotation context refers to itself. In this case, it is unclear whether Frege thought of the quotation marks as expressions that could themselves be quoted.

Montague credited Frege with the insight that the reference of a complex expression may depend on the senses of some its parts (Montague 1968, p. 101), and employed the idea himself, for instance in Montague 1973, with respect to what he called 'denotation' and 'sense', respectively. This comes out in the translation step from the analysis trees to IL, the formal language. Let the translation of α be α' and the translation of β be β' . Then the translation of $F_4(\alpha, \beta)$ is $\alpha'(\hat{\beta}')$.²⁸ This means that the denotation of α' applies to the *sense* of β' , not to the denotation of β' . This holds in general for the rules that correspond to function application in IL (Montague 1973, pp. 261–62).

It may look like ordinary intensional semantics employs switching, in that the extension of a sentence $\Box p$ at a world depends on the intension of p . However, the semantics

²⁷Church (1951) formalizes Frege's theory as an infinite type-theoretic hierarchy of senses and denotations. In Church's system, the type of an expression is explicitly marked by an index. For such a system, there is no semantic switching, since the same semantic function applies to an expression in all linguistic contexts where it can occur. However, switching *will* occur in the *translation* from the informal to the formal language. A sentence s will be mapped on some sentence p_0 in an extensional context, but on a corresponding sentence p_1 in a one-step indirect context.

²⁸ F_4 is the operation that forms a subject-predicate sentence from a term and an intransitive verb, and chooses the present tense, 3rd person singular form of the verb.

is standard compositional with respect to intension, and extension is derived by mere function application. No more than one semantic function is needed. Likewise, in Montague, *sense* itself is standard compositional, and denotation is derived by function application to the designated world-time pair. Denotation is not needed as a separate semantic function.

Michael Dummett (1973, pp. 446–7 and 1991, pp. 47–9) introduced the notion of *ingredient sense*, in part to characterize Frege’s idea of indirect reference, in part also to present his own idea that the *truth value* of a sentence, as opposed to the *correctness of an assertion* of the sentence, is what matters in the evaluation of antecedents of conditionals. Dummett thus distinguished between the *assertoric content* of a sentence, which is its default semantics, when used unembedded, and its ingredient sense, which is what it contributes in various embedded positions. This applies in particular in a many-valued semantics: an assertion of a sentence *s* in a many-valued semantics is *correct* iff *s* has any of the designated values, while in an *embedded* position in a sentence *s'*, the contribution to the correctness conditions of *s'* by *s* will depend on exactly which semantic value, designated or undesignated, *s* has where it occurs in *s'*.

A related distinction is drawn by David Lewis (1980) between *what is said* (or *content*) and *semantic value*. What is said is the meaning of an unembedded sentence *s* in a context of use, while the semantic value of *s* as *embedded* may depend on an index that is distinct from the context of use, for instance if *s* occurs in the scope of temporal adverb of quantification, such as ‘always’.

Switcher semantics as a semantic framework generalizes and systematizes Dummett’s and Lewis’s distinctions. It has been employed in a truth definition format to names in modal contexts, making the modal operator a switcher to the effect that the interpretation of a simple singular term is switched from the world of evaluation to the actual world, thus giving a rigidity-like effect (cf. Glüer and Pagin 2006; Glüer and Pagin 2008; Glüer and Pagin 2014). It has been employed, also in a truth definition format, to general terms in modal contexts, characterizing ordinary language natural kind terms (Glüer and Pagin 2012). And it has been applied, in the algebraic format, to quotation, developing Frege’s idea (Pagin and Westerståhl 2010b).

The latter paper also develops the theory of general compositionality, in particular the *generalization* procedure, which is the process of moving from a single non-compositional semantic function to a general compositional system. There may be a suspicion that general compositionality trivializes compositionality in the sense that any semantics can be rendered general compositional. This question can be rendered precise as follows: is (TRIV) true?

(TRIV) For any semantic function μ there is a switcher semantic system $\mathcal{S} = (S, \mu, \Psi)$.

Since a switcher semantic system is general compositional, this means that any function μ can be combined with finitely many other functions in a semantic system that is general compositional. This issue was investigated in Pagin and Westerståhl 2010b, section 6.

The answer to the question is yes. This might seem to make switcher semantics trivial. However, every application of switcher semantics to date, including the one proposed in the present paper, satisfies two further restrictions. The first, (MAX), says that if two terms t and u make the same semantic contribution in a particular context type c , with respect to μ , then they have the same relevant meaning in c . That is, no difference in meaning if no difference in original semantic contribution:

(MAX) Given a system $\mathcal{S} = (S, \mu, \Psi)$, if the context ξ is such that for any term s that contains ξ it holds that $\mu(s(t/\xi)) = \mu(s(u/\xi))$, then $\Psi^*(\xi)(t) = \Psi^*(\xi)(u)$.

(For both equations, the identity statement is true in case either side is well-defined).

The second restriction concerns Ψ . For any operation-argument place pair $\delta = (\sigma, k)$ we define the function $\Psi_\delta : S \rightarrow S$ such that $\Psi_\delta(\mu) = \Psi(\mu, \sigma, k)$. Then, in all applications so far of switcher semantics, including the present one, it holds that for all i , Ψ_i is either a *constant* function or else the *identity* function. In case i is a switcher, it is a constant function, switching to the same semantic function whichever the function applies to the immediate superterm (containing term), or else it is a *keeper*, i.e. applies the same function to the subterm as was applied to the superterm. In both cases, it depends only on the semantic function applying to the superterm which function applies to the subterm. We call this simplest kind of switcher semantics *First-Grade*.

(F-G) Given a system $\mathcal{S} = (S, \mu, \Psi)$, for all δ , there is $\mu_j \in S$ such that for all $\mu_i \in S$, $\Psi_\delta(\mu_i) = \mu_j$, or else for all $\mu_i \in S$, $\Psi_\delta(\mu_i) = \mu_i$.

We can now revisit the triviality question:

(TRIV*) For any semantic function μ there is a switcher semantic system $\mathcal{S} = (S, \mu, \Psi)$ satisfying (MAX) and (F-G).

It turns out that (TRIV*) is false. A semantic function μ that for a one-place operation σ has the following properties is a counterexample:

$$(12) \quad \mu(s) = \mu(t) = \mu(u), \mu(\sigma(s)) = \mu(\sigma(t)) \neq \mu(\sigma(u)), \mu(\sigma(\sigma(s))) \neq \mu(\sigma(\sigma(t))).^{29}$$

²⁹The proof, which is straightforward, can be found as Proof of Fact 7, in the Appendix of Pagin and West-

The present application (like all earlier applications) satisfies (MAX) and (F-G).³⁰ Here, switching will introduce *structured meanings*. We now turn to that topic.

4. Structured meanings

The most common version of structured meanings is that of *nested tuples*, i.e. tuples based on simple meanings that can themselves contain nested tuples. The most elementary form is that of an ordered pair. Russell (1903) used this model, with objects, properties and relations as the simple meanings. Here we will use possible-worlds intensions as simple meanings.

The basic idea is that we start from a compositional semantic function μ that assigns intensions to atomic terms, and meaning operations to syntactic operations, i.e. functions from n -tuples of intension to intensions. μ is assumed to be defined for an extensional language L that does not contain the belief operator or any other non-extensional operator.

On the basis of μ , a second semantic function μ' is defined that agrees with μ on the atomic terms, and is defined for the same the syntactic operations, but instead of *applying* the meaning operation to the arguments when applied to a complex term, it forms the $n + 1$ -tuple of the meaning operation and the argument meanings.

erståhl 2010b. For second-grade etc. versions of switcher semantics, we would need more iterations to produce counterexamples.

³⁰This triviality objection was pressed by an anonymous referee, who also worried about triviality in practice. Some semantic puzzles, as in an example from Higginbotham 1986, pp. 33–5, might be too easy:

- (i) a. Every student will succeed unless he goofs off.
- b. No student will succeed unless he goofs off.

Higginbotham claimed that ‘unless’ means *or* in (ia) and *and not* in (ib). He rejected standard compositionality (“The Indifference Principle”) in favor of a semantics where the meaning of an expression may depend on its linguistic context. He took the logical form of (ia) and (ib) to be

- (ii) $[Qx : \text{student } x] [(x \text{ succeeds}) \text{ unless}(x \text{ goofs off})]$

On Higginbotham’s proposed contextual semantics, the meaning of ‘unless’ depends on the meaning of the determiner Q . This is in accordance with switcher semantics provided the determiner is treated syncategorematically: the determiner has its own syntactic operation which triggers a switch that has an effect for the interpretation of ‘unless’, which occurs below. However, since according to Higginbotham the effect depends on the upward [downward] monotonicity of ‘every’[‘no’], this requires treating every monotone determiner, simple or complex, as syncategorematic: clearly an undesired multiplication of constructions.

The alternative for Higginbotham’s idea is to have a general binary quantifier construction and let the determiner belong to the left branch of the syntactic tree but still effect the interpretation of ‘unless’ on the *right* branch. This kind of dependence, however, *across sister nodes*, violates general compositionality. And clearly, on this syntactic analysis, no *legitimate* switching solves the puzzle of (i), as the context for ‘unless’ on the present account, is the same in (ia) and (ib). From the point of view of switcher semantics, the solution must then be found elsewhere. For further discussion, see Pelletier 1994a, Higginbotham 2003, Leslie 2009.

To spell this out, we need to define an extended domain of meanings. We start out with a domain M of intensions, and a set R of meaning operations. We then extend M to a domain O that also includes the structured meanings:

- (O) i) $M \subseteq O$
 ii) For $r \in R$, of arity n , and $o_1, \dots, o_n \in O$, if r is *quasi-defined* for $o_1, \dots, o_n$, then $\langle r, o_1, \dots, o_n \rangle \in O$.

The tuple-forming operation is a standard set-theoretic operation. The condition of being *quasi-defined* requires the definition of an *evaluation* function E from structured meanings to intensions: To define E :

- (E) i) $E(o) = o$, if $o \in M$
 ii) $E(\langle r, o_1, \dots, o_n \rangle) = r(E(o_1), \dots, E(o_n))$

For r to be *quasi-defined* for $o_1, \dots, o_n$ consists in r being defined (in the ordinary sense) for $E(o_1), \dots, E(o_n)$. O and E are thereby defined together by simultaneous induction/recursion. Once a new structured meaning o has been added to O , E is defined for it, and o can be an element in a more complex structured meaning.

We can now define μ' in terms of μ and the domain of structured meanings:

- (μ') i) $\mu'(t) = \mu(t)$, if t is atomic and μ meaningful.
 ii) $\mu'(\sigma_i(t_1, \dots, t_n)) = \langle r_i, \mu'(t_1), \dots, \mu'(t_n) \rangle$ iff $\sigma_i(t_1, \dots, t_n)$ is μ meaningful.

Here r_i is the meaning operation that corresponds to σ_i in the μ semantics.

μ' is also a compositional semantic function, where $\lambda x_1, \dots, \lambda x_n (\langle r_i, x_1, \dots, x_n \rangle)$ is the meaning operation corresponding to σ_i .

We can now show by induction that:

Fact 2. for all t , $E(\mu'(t)) = \mu(t)$.

Proof. This is immediate for atomic terms. For a complex term $\sigma(t_1, \dots, t_n)$ we have

$$\begin{aligned} E(\mu'(\sigma_i(t_1, \dots, t_n))) &= E(\langle r_i, \mu'(t_1), \dots, \mu'(t_n) \rangle) \\ &= r_i(E(\mu'(t_1)), \dots, E(\mu'(t_n))) \\ &= r_i(\mu(t_1), \dots, \mu(t_n)) \\ &= \mu(\sigma_i(t_1, \dots, t_n)) \end{aligned}$$

where the induction hypothesis is used in the step to the third line. So, by means of E we can return our original μ meanings from the structured μ' meanings. $\square$

What we will need for the belief sentence semantics, however, will be a slightly more complicated evaluation function E' . The reason is that in case we want to handle iterated belief sentences, like

- (13) a. John believes that Mary believes that pelicans fly.
 b. Cindy believes that John believes that Mary believes that pelicans fly.

etc. In interpreting (13b), μ' will produce a structured meaning from the entire embedded sentence (= (13a)). When reducing that structured meaning to an intension that, provided (13b) is true at a world w , is true at the Cindy belief alternatives w' in w , we don't want to lose the structure that still represents John's belief. We want to reduce that structure only in the next recursive step of evaluating the embedded sentence itself (this will become more clear in the next section). What we need, therefore, is an evaluation function E' that treats belief operators as special cases, that are not to be reduced.

We will use r_b as the meaning operation that corresponds to the syntactic belief term operation σ_b . Implementing this idea gives us

$$(E' \text{ def.}) \quad E'(o) = \begin{cases} o & \text{if } o \in M \\ r_b(E'(o_1), o_2) & \text{if } o = \langle r_b, o_1, o_2 \rangle \\ r(E'(o_1), \dots, E'(o_n)) & \text{if } o = \langle r, o_1, \dots, o_n \rangle, r \neq r_b \end{cases}$$

As can be seen from this definition, in the special case of r_b , E' leaves the second argument intact, which is the belief content position, and reduces only the first argument, which is the attributee position. Because of this, it does not hold automatically that $E'(\mu'(t)) = \mu(t)$. Rather, it will depend on the semantic clauses whether we in the end reduce all structured meanings to unstructured ones.

5. Semantics for notional attributions

The idea for the semantics of a notional belief sentence

- (1) x believes that p .

is the following, informally stated:

- (Bel) The belief sentence (1) is true at a world w iff the structured meaning o of $\ulcorner p \urcorner$, relative to E' , models a belief state of x in w , and $E'(o)$ is true in all x belief alternatives w' to w .

That the structured meaning models a belief state of the attributee ensures that the em-

bedded sentence presents the belief content in a way that corresponds to one way the content is presented to the attributee. That the state is modeled relative to E' is required to ensure that the structured meaning models the belief in the appropriate respect: if we were to only require the existence of an isomorphism f between the structured meaning and a mental representation α , this would not ensure that α is related in the right way to the appropriate intension. The modelling with respect to E' is the *psychological* aspect of the belief attribution.³¹

The requirement that $E'(o)$ be true in all belief alternatives provides the *semantic* aspect of the attribution. It identifies the content as the E' reduction of o , and requires this content to stand in the semantically modeled belief relation, i.e. to be true in the relevant belief-alternative worlds.

Thus, the structured meaning provides an abstract model of a mode of presentation that works as a psychological filter on belief attributions, blocking those attributions that do not correspond to the relevant mode of presentation.

In this section, I shall present the formal implementation of this idea, but without specifying the basic intensional semantics μ for other linguistic constructions.

³¹There are some striking similarities but yet more substantial differences between the present proposal and the account offered by Mark Richard (1990). Richard uses a version of structured meanings but does not coordinate them with possible-worlds contents. Still, what is called “modelling” here corresponds to a relation of *representation* in Richard’s account. The representation relation holds between so-called RAMs, *Russellian Annotated Matrixes* (1990, p. 137). A RAM is a set of ordered pairs of expressions and Russellian interpretations, such as

- (i) $\langle\langle\text{‘is dead’}, \text{being dead}\rangle, \langle\text{‘Twain’}, \text{Twain}\rangle\rangle$.

A set of such RAMs encodes a believer’s *representational system*, or RS, where each RAM derives from a sentence she accepts together with its Russellian (structured proposition) interpretation. RAMs are also determined by that-clauses, for instance (i) by ‘that Twain is dead’. On Richard’s view, the belief sentence

- (ii) Odile believes that Twain is dead

is true iff the RAM determined by its that-clause represents, under *some correlation*, a RAM in Odile’s RS (1990, p. 139).

In Richard’s account, correlations do the job of capturing the *how* of a believer’s acceptance of a Russellian proposition. Under a ‘Twain’-to-‘Twain’ (and ‘is dead’-to-‘is dead’) correlation, (ii) is true iff Odile’s RS contains (i). Under the alternative ‘Clemens’-to-‘Twain’ correlation (ii) is true iff her RS contains (iii)

- (iii) $\langle\langle\text{‘is dead’}, \text{being dead}\rangle, \langle\text{‘Clemens’}, \text{Twain}\rangle\rangle$.

The only *general* restriction on correlations is that they preserve reference (‘Clemens’ and ‘Twain’ are co-referring). This is a problem for Richard’s account, since co-referring terms are not always intersubstitutable in belief contexts. According to Richard, non-linguistic contexts can provide further restrictions on acceptable correlations, but Richard provides no condition for such restrictions to be *adequate*, only *ad hoc* examples (in some cases involving translation). Since Richard rejects appeals to more fine-grained notions of content, such notions are not available as a basis for adequacy. The truth value of a belief sentence is *relative to correlation*, and without further principles for selecting correlations in contexts, the theory has little predictive power.

We first need to add a syntactic operation σ_b that will be mapped on belief sentences:

($V(\sigma_b)$) Where $V(t_i)$ is a singular term, and $V(t_j)$ is a sentence,

$$V(\sigma_b(t_i, t_j)) = V(t_i) \hat{\ } \sqcup \text{'believes that'} \hat{\ } \sqcup V(t_j).$$

Next, we introduce the relation informally specified in (Bel):

(π) The relation

$$\pi(x, o, w)$$

holds between a subject x , a structured meaning o , and a world w , just in case o , relative to E' , models a belief state of x in w .

We add a notation for the belief-alternative, or accessibility relation: $B(x, w, w')$ holds just in case w' is a belief-alternative to w with respect to x .

The meaning operation r_b that corresponds to σ_b is given as follows:

$$(r_b) \quad r_b(x, o) = \{w : \pi(x, o, w) \wedge \{w' : B(x, w, w')\} \subseteq E'(o)\}$$

That is, the meaning operation applies to a subject x and a structured meaning o and returns the set of worlds w such that x , o , and w stand in the π relation and such that $E'(o)$ is true in all the x -belief-alternatives w' to w .

We can now specify the semantic system $\mathcal{S}_b = (S, \mu, \Psi)$, where μ , the designated member of S , is a default intensional semantics, $S = \{\mu, \mu'\}$, and μ' delivers structured meanings as specified in the preceding section. The selection function Ψ is as follows:

- (Ψ) i) For all $\mu_i \in S$, $\Psi(\mu_i, \sigma_b, 2) = \mu'$
- ii) For $(\sigma, k) \neq (\sigma_b, 2)$, for all $\mu_i \in S$, $\Psi(\mu_i, \sigma, k) = \mu_i$

According to (Ψ), $\mathcal{S}$ switches to μ' when encountering the belief operation σ_b , at the second argument position, and makes no other switches. Hence, everything within the scope of the belief operator gets a structured interpretation.

As can be read off from (Ψ), we have the following clauses for belief sentence terms:

- (BS) i) $\mu(\sigma_b(t_i, t_j)) = r_b(\mu(t_i), \mu'(t_j))$
- ii) $\mu'(\sigma_b(t_i, t_j)) = \langle r_b, \mu'(t_i), \mu'(t_j) \rangle$

$\mathcal{S}_b$ is a (general compositional) switcher semantics for belief sentences. As it is presented, it is defined for iterated belief attributions. To give an example, consider again

(13a) John believes that Mary believes that pelicans fly.

and let the term t such that $V(t) = (13a)$ be

$$(14) \quad \sigma_b(j, \sigma_b(m, p))$$

Assume that $\mu(p) = \{w : \text{pelicans fly in } w\}$, and let $\bar{p}$ be the structured meaning resulting from ‘pelicans fly’. We get this result, informally rendered:

$$(15) \quad \mu(t) = \text{the set of worlds } w \text{ such that } \langle r_b, m, \bar{p} \rangle \text{ models a belief state of John's in } w \text{ and it holds in every belief alternative } w' \text{ of John's in } w \text{ that } \bar{p} \text{ models a belief state of Mary's in } w' \text{ and it holds in every belief alternative } w'' \text{ of Mary's in } w' \text{ that pelicans fly in } w''.$$

The full derivation is as follows:

$$(16) \quad \begin{aligned} (i) \quad & \mu(\sigma_b(j, \sigma_b(m, p))) = r_b(\mu(j), \mu'(\sigma_b(m, p))) = \\ (ii) \quad & r_b(J, \langle r_b, \mu'(m), \mu'(p) \rangle) = \\ (iii) \quad & \{w : \pi(J, \langle r_b, \mu'(m), \mu'(p) \rangle, w) \wedge \{w' : B(J, w, w')\} \subseteq E'(\langle r_b, \mu'(m), \mu'(p) \rangle)\} = \\ (iv) \quad & \{w : \pi(J, \langle r_b, \mu'(m), \mu'(p) \rangle, w) \wedge \{w' : B(J, w, w')\} \subseteq r_b(E'(\mu'(m)), \mu'(p))\} = \\ (v) \quad & \{w : \pi(J, \langle r_b, \mu'(m), \mu'(p) \rangle, w) \wedge \\ & \{w' : B(J, w, w')\} \subseteq \{w_2 : \pi(\mu'(m), \bar{p}, w_2) \wedge \{w_3 : B(\mu'(m), w_2, w_3)\} \subseteq E'(\bar{p})\}\} = \\ (vi) \quad & \{w : \pi(J, \langle r_b, M, \mu(p) \rangle, w) \wedge \\ & \{w' : B(J, w, w')\} \subseteq \{w_2 : \pi(M, \mu'(p), w_2) \wedge \{w_3 : B(M, w_2, w_3)\} \subseteq \mu(p)\}\} = \\ (vii) \quad & \{w : \pi(J, \langle r_b, M, \mu(p) \rangle, w) \wedge \forall w_4 (B(J, w, w_4) \rightarrow \pi(M, \mu'(p), w_4)) \wedge \\ & \forall w_5 (B(M, w_4, w_5) \rightarrow w_5 \in \{w_6 : \text{pelicans fly in } w_6\})\} = \\ (viii) \quad & \{w : \pi(J, \langle r_b, M, \mu(p) \rangle, w) \wedge \forall w_4 (B(J, w, w_4) \rightarrow \pi(M, \mu'(p), w_4)) \wedge \\ & \forall w_5 (B(M, w_4, w_5) \rightarrow \text{pelicans fly in } w_5)\} \end{aligned}$$

where set inclusion is reduced to universal quantification in steps (vii) and (viii). The fact that $E'(\mu'(p)) = \mu(p)$ is used in the step to (vi).

The relation between μ' and μ , that $E'(\mu'(t)) = \mu(t)$, holds also in the extended fragment:

Fact 3. In the fragment that just adds σ_b to the atomic fragment, $E(\mu'(t)) = \mu(t)$.

Proof. Since the definition of E' agrees with the definition of E in the σ_b -free fragment, it follows by Fact 2 that the claim is true for σ_b -free terms t . For a term $s = \sigma_b(t, u)$, we see that $\mu(s)$ differs from $E'(\mu'(s))$ only in first argument to r_b : $\mu(s) = r_b(\mu(t), \mu'(u))$, while $E'(\mu'(s)) = r_b(E'(\mu'(t)), \mu'(u))$, by the definitions of r_b and E' . But since t is σ_b -free, $E'(\mu'(t)) = \mu(t)$. Hence, $E'(\mu'(s)) = \mu(s)$. $\square$

6. Properties of and extensions to the semantics

Under the appropriate understanding of “same structure”, which is the one intended by Carnap (1947/1956, pp. 61–2), two sentence terms t and t' of the same language, for which μ' is defined, count as intensionally isomorphic iff $\mu'(t) = \mu'(t')$.³² Because of this, the conditions of truth-preserving substitution in belief contexts are the same in Carnap’s analysis and the present account. But since the current account does not appeal to quotation of the embedded sentence, the Langford translation test poses no problem.³³

Let us return to the substitution failures discussed in section 2. We exemplified the failure of truth value preservation through co-extensional substitution with

- (5a) John believes that Mark Twain is a novelist.
- (5b) John believes that Samuel Clemens is a novelist.

This failure is accounted for here already by the fact that the designated semantic function μ is intensional, together with the assumption that the proper names ‘Mark Twain’ and ‘Samuel Clemens’ have different intensions.

This may appear question-begging from the perspective of Kripkean orthodoxy regarding proper names. However, the main point in Glüer and Pagin 2006 is that the modal intuitions that have been adduced in support of the rigidity thesis about proper names can be accounted for by a semantics that allows non-rigid names but uses semantic switching. In that paper, the idea is that the (metaphysical) modal operator acts as a switcher with the following effect (for a name N : it switches to a semantic function μ_a such that, for any world w , $\mu_a(N)(w) = \mu_0(N)(@)$, provided $\mu_0(N)$ exists in w . Hence, a proper name in a modal context semantically contributes its actual world referent, even if the name is not a rigid designator in the default semantics μ_0 . This accommodates the intuition about e.g. what Aristotle might have done (still being *Aristotle*), while allowing names to be non-rigid for the benefit of accounting for other linguistic contexts, such as attitude contexts. This is taken advantage of presently.

Next, we exemplified the failure of truth preservation through co-intensional substitution with

- (3a) John believes that pelicans fly.
- (3b) John believes that: pelicans fly and $\sqrt{529} = \pm 23$

³²This fact is non-trivial. We need to show that sameness of μ' value is both necessary and sufficient for intensional isomorphism. For a proof of a slightly more general fact, see Pagin 2003.

³³In addition, extending Carnap’s account, or any quotation account, to an account of relational attributions would involve quantifying into quotation contexts.

The failure of truth preservation in the substitution step from (3a) to (3b) is accounted for in the current semantics by the fact that $\mu'(t)$ may model a belief state of John's (in a world w), while $\mu'(u)$ still does not, where $V(t) = \text{'pelicans fly'}$ and $V(u) = \text{'pelicans fly and } \sqrt{529} = \pm 23$ '. Then, by (BS), $\mathcal{S}_b$ lets (3b) come out as false.

This also illustrates the fact that the well-known problem of *logical omniscience* that plagues standard possible-worlds models of belief is not a problem for the current account. In general, if a belief sentence

(1) x believes that p .

is true at a world w , and if $\ulcorner q \urcorner$ is a logical consequence of $\ulcorner p \urcorner$, then $\ulcorner q \urcorner$ is true at all x -belief-alternatives w' to w . However, it does not follow that

(17) x believes that q .

is true in w , since the structured meaning of $\ulcorner q \urcorner$ may fail to model a belief state of x . The informally stated principle

(H⁻) If p is true in all worlds w' that are x -belief alternatives to w , then x believes that p in w .

fails. There is no *logic* of belief in the sense that $Bp \models_B Bq$ provided $p \models_B q$. Here $\models_B$ is schematic for some restriction of the consequence relation that would support exactly inferences from belief attributions to belief attributions. Such inferences are psychological in nature, not logical.³⁴

This fact raises the question of belief in logical or mathematical falsehoods. It is true in no world that $\sqrt{529} = \pm 22$. Still, (BS) allows that

(18) John believes that $\sqrt{529} = \pm 22$.

can be true. It is true, however, only in case the set of John's belief alternatives is the empty set.³⁵ If it is not, any sentence that attributes to John a belief in a mathematical falsehood comes out false. In order to make a relevant distinction nevertheless, we

³⁴This holds even for the inference from *John believes that p and q* to *John believes that q and p* , an inference that is psychologically justified but not logically true. I am largely in agreement with Richard's (1990, pp. 30–4) discussion of this issue. However, according to Richard's view, the two attributions to John can be true because of *the same state*, despite relating John to two different (structured) propositions. This does not agree with the individuation of belief states as effected by structured meanings relative to an evaluation function. Of course, we can countenance a more coarse-grained *type* of state that is closed under conjunction commutation, but this does not render the inference above, relying on the more fine-grained type of states, any less psychological.

³⁵If the set of belief alternatives is the set of worlds compatible with everything John believes, then the set automatically becomes empty if he believes a necessary falsehood.

can introduce a notion of *weak belief* that leaves out the requirement of truth in belief alternatives but keeps the psychological part:

(Weak Belief) A sentence “*x* weakly believes that *p*” is true in *w* iff the structured meaning *o* of $\ulcorner p \urcorner$, relative to *E'*, models a belief state of *x* in *w*.

Then, interpreting ‘believes’ in (18) as *weakly believes* allows the sentence to come out true. Of course, if the set of belief alternatives is empty, anything weakly believed is also strictly believed. Weak belief is clearly inadequate as a general notion of belief, since it leaves the relation between believer and propositions believed completely undetermined.³⁶

Let’s move on to synonymy. We exemplified the possible failure of truth value preservation through synonymous substitution with

- (4a) John believes that pelicans rarely fly.
- (4b) John believes that pelicans seldom fly.

If indeed there can be such a substitution failure, the present semantics has no means of accounting for it. However, I deny that truth can fail to be preserved through such substitutions. The reason to think that it could is that John (a fairly normal speaker of English) might assent to (19a) but dissent from (19b):

- (19) a. Pelicans rarely fly.
- b. Pelicans seldom fly.

The reason he dissents from (19b) is that John misunderstands the word ‘seldom’, or attaches a different meaning to it. But it is a misunderstanding to think that this is a reason for the failure of truth preservation. In a belief attribution, the *speaker* uses an embedded sentence to characterize a belief state of the attributee. The speaker uses the words in the embedded sentence as *she* understands them. If ‘rarely’ and ‘seldom’ are synony-

³⁶The account of Scott Soames (1987) does, like the present one, combine structured contents with a kind of possible-worlds intension: sets of circumstances. Like the present account, the (would-be) intension of a belief sentence depends on the structured content of the embedded sentence. However, Soames’s approach is very different, especially in two main respects. First, the semantics itself only delivers structured contents, while sets of circumstances are assigned to the structured contents, in a stipulated fashion, similar to an ordinary truth definition (Soames 1987, p. 63; Salmon (1986, pp. 149–50) proceeds this way, too). Hence, sentences do not strictly speaking have intensions as meanings. For instance, the structured semantic value of a belief sentence, $\langle\langle o, \text{Prop } S \rangle, B\rangle$, where *o* is the believer, Prop *S*, the structured content value of *S*, and *B* is the belief relation, is said to be true at a circumstance *E* iff the pair $\langle o, \text{Prop } S \rangle$ is in the extension of *B* in *E*. This holds, by Soames’s account, *independently of* what the intension of Prop *S* is. This is so, since the belief sentence does not entail, for any intension, that the believer is belief-related to that intension. As rehearsed in section 1, structured meanings do not in themselves determine intensions. The intension part is therefore not integrated with the structure part.

mous for the *speaker*, then it is exactly the same belief state that is attributed by means of (19b) as by means of (19a). What meanings the *attributee* attaches to them is simply irrelevant. What matters to the truth or falsity of the belief sentences is only what belief states the attributee is in, irrespective of how he is disposed to express these states.

Of course, linguistic assent and dissent is an important source of knowledge about other persons' beliefs. If Mary, a competent speaker of English, believes that John is one too, and she observes John dissenting from (19b), she will be by default justified in inferring the falsity of (4b). But in exceptional cases, as in the example, the inference does lead from a true premise to a false conclusion.³⁷

Nonetheless, among

(7a) John believes that a fortnight is a period of ten days

(7b) John believes that a period of fourteen days is a period of ten days.

(7a) can be true and (7b) false, according to Burge, despite the fact that the mistake on the part of the attributee is linguistic, a misunderstanding of 'fortnight'. For the reasons just adduced, I think that in such a case, (7a) is literally false. It looks acceptable only because the occurrence of 'a fortnight' under such a reading is tacitly embedded in a quotation context. This comes out in the unacceptability of

(20) [The concept of] a fortnight is such that John believes that it is [the concept of] a period of ten days.

which indicates that 'a fortnight' cannot be straightforwardly extracted, and the embedding in a tacit quotation context explains this.³⁸

³⁷This is the so-called *disquotational principle* discussed in Kripke 1979, pp. 248–9. The problem there is that the attributee, Pierre, is as competent as his fellow speakers, although in two different languages, and not fully competent in translating between them. For a suggestion to handle such cases in terms of *entrenchment*, see Pagin 1992.

³⁸The appeal to a quotation analysis was made early on in Church 1954 and in Sellars 1955 against Mates 1950. Burge objected to the quotation analysis (1978, pp. 123–31) in a number of ways, which cannot all be discussed here. Oddly enough, Burge on the one hand says that his arguments show that "the fact that a belief involves linguistic considerations does not show that the belief content must be reformulated so as to contain quotation marks or otherwise denote linguistic expressions" (1978, p. 127), and on the other himself suggests that "expressions occurring in *that*-clauses function autonomously—to denote themselves". It should also be noted, however, that Burge's view is at least partly motivated by his *linguistic externalism* (later more known through Burge 1979): "When the presupposition is in force, communal conventions about the meaning of a speaker's words tend to override what a speaker mistakenly associates with his words in determining what he says and even, sometimes, believes." (1978, pp. 134–5).

As I understand it, Burge's view can be broken down into the following elements. 1) John's attitude to the embedded sentence 'A fortnight is a period of ten days', depends on his mistaken *personal* interpretation of 'a fortnight'. 2) Because John *holds* this sentence true, the belief sentence (7a) *is* true. 3) The *content* of John's belief, truly attributed by (7a), is determined by the *communal* interpretation of 'a fortnight'. 4) When we substitute 'a period of fourteen days' for 'a fortnight', we get the sentence 'A period of fourteen days is

Since there is a tendency to accept attributions such as (7a) to report misunderstandings that are admittedly linguistic, we can, and perhaps should,³⁹ treat this as a linguistic phenomenon in its own right. That is, we can interpret (7a) as reporting what is more explicitly reported by

(21) John believes that what is called ‘a fortnight’ is a period of ten days.

In the imagined circumstances, (21) is, of course, true. In order to get a reading of (7a) with the explicit meaning of (21), we will need a non-standard way of producing it. What we need is a syntactic operation that allows a change in meaning without altering the surface syntax. Let’s call it *quasi-quotation*.

To implement the idea, we add a syntactic operation σ_{qq} with the following properties:

- (σ_{qq}) i) $V(\sigma_{qq}(t)) = V(t)$
- ii) $\mu(\sigma_{qq}(t)) = r_{qq}(V(t))$
- iii) $\mu'(\sigma_{qq}(t)) = \langle r_{qq}, V(t) \rangle$
- iv) $r_{qq}(V(t))(w) = \text{that which is called } V(t) \text{ in } L \text{ in } w$

As can be seen from clauses (ii) and (iii), σ_{qq} is itself a switcher, inducing a switch from μ/μ' to V as applied to the immediate subterm. This works exactly as in ordinary quotation, where V is also added as a semantic function as applying to quoted material (cf. Pagin and Westerståhl 2010b). Adding σ_{qq} thus produces a system with three semantic functions.

Let’s set $V(f) = \text{‘a fortnight’}$, $V(p) = \text{‘a period of ten days’}$, and $V(\sigma_=(t, u)) = \lceil V(t) \wedge \text{is} \wedge V(u) \rceil$. We also need additions to the semantics: $\mu(\sigma_=(u, t)) = \{w : \mu(t)(w) = \mu(u)(w)\}$; $\mu(p)(w) = \text{a period of ten days in } w$.

Then result of applying (σ_{qq}) delivers (leaving out explicit mention of word spaces):

- (22) a. $V(\sigma_=(\sigma_{qq}(f), p)) =$
 $V(\sigma_{qq}(f)) \wedge \text{‘is’} \wedge V(p) =$
 $V(f) \wedge \text{‘is’} \wedge V(p) =$
 $\text{‘a fortnight’} \wedge \text{‘is’} \wedge \text{‘a period of ten days’} =$

a period of ten days’, which John does *not* hold true (because his personal understanding agrees with the communal interpretation), and hence the corresponding belief sentence (7b) is false. This accounts for the substitution failure.

If we accept Burge’s proposed truth value distribution between (7a) and (7b), his externalism can be invoked to explain it. Allowing a meta-linguistic reinterpretation, explicit or, as in the present account, implicit, eliminates the need for an explanation, and thus weakens the case for externalism.

³⁹This was insisted upon by Robert Stalnaker in a Q&A in Umeå.

‘a fortnight is a period of ten days’.

- b. $\mu(\sigma_{=}(\sigma_{qq}(f), p)) =$
 $\{w : \mu(\sigma_{qq}(f))(w) = \mu(p)(w)\} =$
 $\{w : r_{qq}(V(f))(w) = \mu(p)(w)\}$
 which contains a world w' iff that which is called ‘a fortnight’ in L in $w' =$ a
 period of ten days in w' .

Thus, this extension of $\mathcal{S}_b$ provides the resources to accommodate the Burge example insofar as accepting surface forms such as (7a), without compromising the basic view that under their default syntactic derivation, they are literally false.

Further, the account involves the assumption that atomic co-intensional expressions are also synonymous, as with the pair ‘rarely’ and ‘seldom’. But the assumption is open to question. Taken in a trivial sense, the assumption is clearly false. The adjectives ‘triangular’ and ‘trilateral’ are both (syntactically) atomic, co-intensional, and intuitively non-synonymous. Other examples, including morphologically simple ones, can be constructed *ad libitum*.

But the real assumption is less trivial and weaker: if two expressions are *semantically* atomic, i.e. express a simple *concept*, and are co-intensional, they are also synonymous. Again, ‘rarely’ and ‘seldom’ provide a plausible example, but ‘triangular’ and ‘trilateral’ do not, since they intuitively mean *has three angles* and *has three sides*, respectively. They would then not be represented as having atomic *structured* meanings. The question then is whether there are plausible examples of pairs of co-intensional expressions that are semantically simple and yet intuitively non-synonymous. If there are, the appeal to structured meanings for achieving the relevant semantic fineness of grain, fails.

A possible candidate is provided by the nouns ‘knowledge’ and ‘evidence’ according to Tim Williamson’s (2000) account of knowledge.⁴⁰ Williamson says

Suppose that knowledge, and only knowledge, justifies belief. That is, in any possible situation in which one believes a proposition p , that belief is justified, if at all, by propositions $q_1, \dots, q_n$ (usually other than p) which one knows. On that supposition, if justified belief is central to epistemologico-sceptical inquiry and the philosophy of science, then so too is knowledge. Now assume further that what justifies belief is *evidence* (this assumption is briefly discussed in section 9.8). Then the supposition just made is equivalent to the principle that knowledge, and only knowledge, constitutes evidence. This chapter defends that principle; it equates S’s evidence with S’s knowledge, for every individual or community S in any possible

⁴⁰This possible counterexample was proposed by Dave Chalmers during the Q&A in Umeå. Chalmers himself considers the issue in Chalmers 2011, p. 601 (although not with the Williamson example) and takes a stance similar to the present one.

situation. Call this equation $E = K$. (Williamson 2000, p. 185)

The suggestion is that both ‘knowledge’ and ‘evidence’ are semantically atomic, co-intensional ($E = K$), and intuitively non-synonymous. Williamson clearly takes knowledge not to be analysable in terms of belief, and probably not analysable at all (Williamson 2000, pp. 2–5). However, he clearly seems not to take *evidence* as a simple concept. In the quoted passage he assumes that “what justifies belief is evidence”, and later he adds

The present case for $E = K$ is now complete. If evidence is what justifies belief, then knowledge is what justifies belief. (Williamson 2000, p. 208)

As far as I have understood, the assumption part is meant to leave the possibility that there is, over and above truth-directed justification of belief, also *pragmatic* justification. If this is right, then Williamson at least takes evidence to be exactly what gives truth-directed justification of belief. Hence, on Williamson’s view, evidence is not a simple concept. This means that there is scope for Williamson’s argument for the $E = K$ equation by way of arguing that it is exactly knowledge which provides truth-directed justification of belief.

Therefore, the $E = K$ equation does not, even by Williamson’s standards, provide a counterexample to equating co-intensionality of semantically simple expressions with synonymy. Nevertheless, it is hard to disprove the possibility of such counterexamples. I believe that there aren’t any, but the semantic account proposed relies on this assumption.⁴¹

⁴¹It is instructive at this point to compare the present approach with that of Cresswell and von Stechow (1982). Cresswell and von Stechow (CvS) also use a version of structured meanings to model belief contents, and combine structured meanings with a possible-worlds framework. There are big differences, however. Perhaps the main difference is that CvS attain the hyperintensional fineness of grain not by the structured meanings themselves, but through various *de re* readings. Thus, the sentence

- (i) Poirot believes that $2 + 2 = 4$.

can be interpreted as relating Poirot to the structure

- (ii) $\langle \textit{that}, \langle \langle 2, +, 2 \rangle, =, 4 \rangle \rangle$

but this is relating him only to the intension of the structure, the necessary proposition. The reason is that $\langle \langle 2, +, 2 \rangle, =, 4 \rangle$ is the structured necessary proposition, and *that* a function that maps it on an object. To get finer grain, (i) must be interpreted to let Poirot’s belief concern entities *de re*, which amounts to combining *that* directly with these elements, for instance as in

- (iii) $\langle \textit{that}, \langle \lambda, x, y \langle \langle x, +, y \rangle, =, 4 \rangle \rangle, 2, 2 \rangle$

(p. 516), which is the structure containing 2, 2, and the property of summing to 4. *that* in this case maps these *three* intensions on a structured object (‘that’ is infinitely ambiguous). It is up to pragmatics to select an appropriate *de re* reading. The *de re* structures are the *objects* of belief. By contrast, here we use structured meanings as directly supplying fineness of grain, while unstructured propositions still are the objects of belief.

Finally, the account of notional attributions provided in section 5 involves a substantial and often unrealistic idealization. It consists in using, in the relation $\pi(x, o, w)$, the structured meaning o that is constructed from the meanings of terms in the language of the *speaker*. This would seem to often get things wrong, since different members of the same speech community often have slightly different conceptions of objects and properties. In particular, the meanings associated by the speaker with terms in a belief context of an attribution might be different from the meanings associated with the same terms by the attributee. For instance, suppose Bill says

(23) John believes that Elsa has acquired a dog.

Suppose further that John and Bill are ordinary speakers of English and both use the word ‘dog’, but that there is a small difference between Bill’s notion of a dog, dog_b , and John’s notion of a dog, dog_j .⁴² As a consequence, the possible-worlds proposition p_b that Bill associates a time t with the sentence

(24) Elsa has acquired a dog.

is slightly different from the possible-worlds proposition p_j that John associates with the same sentence: there is a world w where p_b is true but p_j false. Should we say that because of this, Bill’s notional attribution to John is *false*?

My view is that we should not. I think that the truth-conditions of belief attributions should be coupled with the success-conditions of communication. In the present case, this means that if Bill would successfully communicate to John that Elsa has acquired a dog, by means of the sentence, (24), despite their conceptual differences, then in most ordinary cases, Bill’s attribution to John should be counted as true as well. I also think that the communication *should* be considered successful if the conceptual difference is small enough (as in the example extended in the footnote). The difference between p_b and p_j is simply too marginal to induce failure of communication about such mundane affairs as acquisitions of dogs.

The idea that truth-conditions of belief attributions should be coupled with success conditions of communication has also been defended by David Chalmers (2011, pp. 619–

⁴²For example, suppose that according to John, an animal is a dog only if it belongs to a sub-species of *Canis lupus* that was domesticated no later than 35.000 years ago, while according to Bill, an animal is a dog only if it belongs to a subspecies of *Canis lupus* that was domesticated no later than 15.000 years ago. Since these are differences with respect to how they interpret ‘dog’, the difference is conceptual, not just empirical. Suppose that they agree on all other conditions. To complete the story, assume that in fact, any sub-species of *Canis lupus* that was domesticated more than 15.000 years ago was domesticated more than 35.000 years ago as well. Hence, despite their conceptual difference, John and Bill will agree on the dog-hood of any existing animal.

21). Chalmers proposes a relation of *coordination* between propositions (of a certain kind), both for the truth-conditions of belief sentences and for communication, and we shall look more closely at the former in the appendix.

The main point of coordination in belief attributions is that the attributee is ascribed a direct relation to a proposition that need not be the same as that expressed/denoted by the speaker, but it is required that they stand in an appropriate relation. In the present account, when allowing for conceptual differences between speaker and attributee, we need a relation corresponding to Chalmers's coordination relation. I shall here just sketch the main idea. I shall say that two concepts α and β are *@-similar* if, and only if, for any *close* possible world w , the extension of α is the same as the extension of β . That is, within the set of *close* worlds, $C_@$, their extensions coincide. Unless α and β are identical, their extensions will diverge outside $C_@$ but they are similar in the sense that their *restrictions* to $C_@$ are the same. This relation of @-similarity is clearly an equivalence relation.⁴³

A mini-gloss on what it is to be a close world, would run something like this. A possible world that is *close* to the actual is a world where things are normal, from the perspective of the actual world. The laws of nature are the same. The geography is the same. The history of the world up to the relevant time is the same except possibly for some marginal late deviations. In short, pretty much all the typical landmarks by which ordinary objects and properties are identified still identify the same objects and properties. Then, for instance, if in the actual world the celestial body seen in position X in the morning is the same as the celestial body seen in position Y in the evening, then in any close possible world it is still the same possible celestial body. But if in the actual world the man I rode the bus with yesterday is the man who robbed the bank today, this need not hold in all close worlds. In a sense, things that could *easily* have been different are different in some world in $C_@$, while other things are the same.

We extend the similarity relation to a relation between structured meanings in the obvious way: two isomorphic structured meanings o and o' are @-similar iff all their corresponding simple constituents are @-similar. The question now is how to implement this similarity relation in the semantics. The idea is to introduce an *approximate* belief attribution, where the structured meaning used in the attribution is @-similar to one that strictly models the belief state of the attributee. This is implemented by replacing the meaning operation r_b with an operation r'_b , as follows:

$$(r'_b) \quad r'_b(x, o) = \{w : \exists o' (@\text{-SIM}(o, o') \wedge \pi(x, o', w) \wedge \{w' : B(x, w, w')\} \subseteq E'(o'))\}$$

where @-SIM is the @-similarity relation. On this way of implementing similarity, we

⁴³This account is developed in Pagin 2017b.

deem an approximate belief attribution true just in case the structured meaning that results from the embedded sentence is @-similar to a structured meaning that *does* model a belief state of the attributee, and which is mapped on a proposition true in all relevant doxastic alternatives. To trigger the use (r'_b), we need an approximate belief operator σ'_b in the syntax, providing approximate attributions over and above strict attributions.

Approximate attributions allow for the attributions of approximately inconsistent beliefs. To exemplify, suppose that John has two distinct but very similar *dog* concepts, *dog'* and *dog''*, and that *both* are @-similar to Bill's concept of a dog. Now, suppose the following are strictly true:

- (25) a. John believes that Elsa has acquired a *dog'*.
b. John believes that Elsa has not acquired a *dog''*.

Since both *dog'* and *dog''* are @-similar to Bill's concept *dog*, by a semantics that uses (r'_b), both (26a) and (26b) count as true:

- (26) a. John believes that Elsa has acquired a dog. (= (23))
b. John believes that Elsa has not acquired a dog.

Hence, by the resulting semantics, John is attributed two approximately inconsistent beliefs. Clearly, John does not have strictly inconsistent beliefs. The example has the same basic character as standard Fregean coreference cases where reference is the same but senses differ: here coarse-grained (restricted) intension is the same but the fine-grained (unrestricted) intensions differ.

This outcome is acceptable: just as Ralph can believe of Ortcutt both that he is a spy and also believe of Ortcutt that he is not a spy, under different modes of presentation, so John can believe for some doggyish kind that Elsa has acquired a specimen of that kind and also believe for some (other) doggyish kind that Elsa has not acquired a specimen of that kind. The surface syntax is the same, but the semantics makes it clear that in (26) there is no strict attribution of inconsistent beliefs.

(r'_b) helps implement the intuitive idea that small conceptual differences shouldn't matter, in the following sense. We restrict attention to the set of normal worlds, $C_@$. Suppose that John's beliefs are *normal* in the sense that there is some normal world where all John's beliefs are true: there is an overlap between John's doxastic alternatives and the set of normal worlds: $\{w : B(j, @, w)\} \cap C_@ \neq \emptyset$. This intersection is the set of John's *normal* doxastic alternatives. Now, suppose that propositions p and q are formed from only extensional operations and that they result from @-similar structured meanings o_p and

o_q . Then, for any world w in $C_@$, and if p is true in w , then q is true in w .⁴⁴ Therefore, an *approximate* belief attribution is true (in @) if, and only if, the corresponding *strict* belief attribution is true (in @), as long as we consider only normal doxastic alternatives. Of course, if two approximately inconsistent attributions are true, like (26a) and (26b), then the set of normal doxastic alternatives (for John, in @) is empty.

This has the further desirable consequence that communication by means of belief sentences can be successful across small conceptual differences. For instance, if Bill and Susan have @-similar concepts of John, Elsa, acquiring and dogs, and if John has normal doxastic alternatives, then Bill can use (26a) to communicate successfully that fact of John's belief to Susan. We would think that successful communication in this case would also depend on the similarity of their respective concepts of *belief*. In the proposed framework, however, the concepts of belief are represented by the (r_b) and (r'_b) meaning operations, which do not vary from subject to subject. At most, Bill might intend a strict and Susan an approximate attribution, or vice versa. But in ordinary circumstances, where the attributee has normal doxastic beliefs, strict and approximate attributions coincide in truth value in normal worlds, and this suffices for communicative success in such cases.

Finally, it appears to be an empirical fact that many concepts differ somewhat between speakers (cf. Majid 2010, Pagin 2017b). Therefore, the idea of approximation is important. Without it, too many belief attributions are false. On top, we lose the distinction between attributions that are largely correct (false only because of small conceptual differences) and attribution that are clearly incorrect (false because the attributee lacks the ascribed attitude even to an approximate content). In addition, the device of approximation can be exploited for some otherwise problematic belief attributions.⁴⁵

⁴⁴Since the operations are extensional, the extensions of the values of the operations in a particular world depend only on the extensions or their arguments in that world. And the extensions of the elements of @-similar structured meanings are the same in any world in $C_@$.

⁴⁵For instance, to take John Perry's (1977, p. 487) example of Heimson who believes that he is Hume. I can truly say

- (i) Heimson believes that he is Hume.

because Heimson's conception of himself does pick out Heimson in all normal worlds; this is the only metaphysically possible alternative. Heimson's belief alternatives, however, are all non-normal (in fact metaphysically impossible). In those worlds, his self-conception picks out Hume. Since my conception of Heimson and his conception of himself are @-similar, I can still truly but *approximately* attribute the belief (i), even though that belief itself is true in no metaphysically possible world, nor in any of my belief-alternatives.

This account has an interesting affinity with Stalnaker's idea (2014, pp. 215–6, 221–2) of an I-concept that is rigid *within* the context set of worlds, but may be *non-rigid* outside the context set. Such a non-rigid I-concept in fact transfers to belief *content* what is otherwise relegated by Stalnaker to the belief *state* in chapter five of the same book (Stalnaker 2014, pp. 120–3), where he uses centered worlds to track who the subject "takes himself to be" in a particular possible world. In that apparatus, whom a subject takes himself

7. Concluding remarks

In this paper I have argued that we need both unstructured meanings, in particular standard possible-worlds propositions, and structured meanings. The former are needed for the intuitive subject matter of beliefs, i.e. ordinary belief contents, while the latter are needed for the intuitive individuation of belief states, as ascribed by means of belief sentences.

I have also tried to show that by allowing *general* instead of standard compositionality, we can devise a switcher semantics that combines a structured meanings account with a possible-worlds framework, allowing the structure meanings to model belief states and thereby act as a filter on admissible attributions, allowing some and ruling out other attributions of the very same content.

I would like to end with a brief comment on what Stephen Schiffer (1992, pp. 512–9) calls “the meaning-intention problem”. The meaning-intention problem is the problem that a semantic theory requires a speaker of a sentence to have as the content of the belief expressed exactly the meaning ascribed to the sentence by the theory. Schiffer exemplifies it with what he calls the *hidden-indexical* theory of belief attributions, which is an indexical variant of the account by Salmon mentioned above. Schiffer (1992, p. 512) considers an utterance by Flora of

(27) Harold believes that TWA is offering a New York-Paris return ticket for \$318.

He goes on to point out that the hidden-indexical theory has the consequence that what Flora *meant* by her utterance is

(28) that there is something [a mode of presentation] that both has Φ and is such that Harold believes the proposition *that TWA is offering a New York-Paris return ticket for \$318* under it.

for some particular property Φ of modes of presentation. He adds that neither Flora nor her audience is aware of her meaning any such thing. The problem is that a semantic theory in order to be true should ascribe contents to sentences that agree with the intentions of speakers in using these sentences. Schiffer considers and rejects the idea that speakers and audiences have *tacit* beliefs with these contents. To what extent should a belief sentence semantic theory, and in particular the one on offer here, live up to the requirement that speakers must *intend* what the theory ascribes? The current theory is technically far more complex than the hidden-indexical theory, and so the problem might appear even more acute in this case.

to be in a possible world does not make a difference to belief content.

The answer has two parts. The first part is that the present idea of structured meanings as filters simply *blocks* the attribution of the theory outputs to speakers. For instance, if Flora utters

(13a) John believes that Mary believes that pelicans fly.

the theory assigns as truth-conditions of her utterance that

(29) The actual world is a member of $\mu(t)$ = the set of worlds w such that $\langle r_b, m, \bar{p} \rangle$ models a belief state of John's in w and it holds in every belief alternative w' of John's in w that $\bar{p}$ models a belief state of Mary's in w' and it holds in every belief alternative w'' of Mary's in w' that pelicans fly in w'' .

But Flora clearly is not in any belief state that would be modeled by the structured meaning of (29) itself. So, it wouldn't follow from the theory that Flora has a belief properly attributed to her by means of using (29).

The second part is the more important. We want an answer to the question what the empirical basis for a semantics for belief sentences is, given that speakers do not in any ordinary sense intend the contents attributed to them by the semantic theory. The empirical basis is the following: if the semantic theory is adequate for a speaker S, then S's intuitions about the circumstances under which an ordinary belief attribution that she makes is true are exactly those where the semantic theory entails that they *are* true. When that condition is met, the theory provides a model of the contents of S's belief attributions, even if S is not able to state the theory herself or even to understand a statement of it. No stronger requirement than that on the speaker's intentions is justified. The mere fact that a theory is technically complicated does not undermine it.

Appendix: David Chalmers's account

The idea of an existential quantification over contents together with an appropriate relation between the content of the speaker and the content of the attributee is central to David Chalmers's (2011) approach. Chalmers states his general principle for notional attributions as follows (Chalmers 2011, p. 608):

(CEP) ' x believes that S ' is true of i iff $\exists q(E(i, q) \ \& \ C(q, p))$

Here ' p ' is short for the *enriched proposition* expressed by the sentence ' S ', and ' q ' is a bound variable ranging over enriched propositions. ' E ' stands for the *endorsement* relation, and ' C ' stands for the *coordination* relation. I'll briefly explain these notions.

Chalmers's account is stated within his *two-dimensionalist* framework. In this paper

and in other more recent papers by Chalmers (e.g. Chalmers 2006), two-dimensionalism is not the classical system with two-dimensional intensions that are evaluated at pairs of worlds. In the classical framework, one-dimensional intensions were defined from the two-dimensional intension, I^2 in two ways. The *primary* intension, I^p , was the intension whose extension at a world w is the extension of I^2 at $\langle w, w \rangle$, i.e. extension of I^2 along the diagonal of the two-dimensional matrix. The *secondary* intension, I_w^s , at a world w , has as its extension at a world w' the extension of I^2 at $\langle w, w' \rangle$.

In the more recent two-dimensionalist framework, the ideas of a primary and a secondary intensions are retained, but they are partially redefined and are independent of the original I^2 . Rather, the system is two-dimensional simply in having two *or more* intensions (Chalmers 2011, p. 599). The secondary intension is a function from (metaphysically) possible worlds to extensions (2011, pp. 597–8). Since ‘Hesperus’ and ‘Phosphorus’ are co-referring rigid designators, in Kripke’s (1980) sense, ‘Hesperus is Phosphorus’ has a necessary secondary intension, true in all worlds. The secondary intension of Chalmers 2011 is thus the old $I_{@}^s$, as fixed in the actual world @.

The primary intension is a function from *scenarios* to extensions. A scenario is (“on a standard understanding”) a *centered world*, which in turn is a triple $\langle w, a, t \rangle$ of a world w , an individual a (in the domain of w) and a time t (2011, pp. 597–8). The centered-world construction is used for characterizing indexical beliefs, as suggested in Lewis 1979b. The primary intension of a sentence is true at a scenario “when, if the subject were to accept that they were inhabiting the scenario in question, they would endorse the sentence in question” (2011, p. 598). The primary intension of ‘Hesperus is Phosphorus’ can be false in a scenario; it is not epistemically necessary, i.e. a priori.⁴⁶

The *enriched* intension of a simple expression is an ordered pair of a the primary intension of an expression and its *extension*. In Chalmers’s notation, the enriched intension h of ‘Hesperus’, is an ordered pair, written ‘ h'/v ’, where h' is the primary intension of ‘Hesperus’, and v , i.e. Venus, is the extension (Chalmers 2011, p. 600). The extension of a one-place predicate is here not a set of individuals, but a *property*. Since a predicate is taken to rigidly denote the same property in all worlds, the extension, in this sense, of a predicate, determines its secondary intension.

The enriched intension of a *complex* expression is a structure that reflects the logical form of the expression. The simple parts of the structure are simple enriched intensions. The enriched intension of a sentence is an *enriched proposition*. The enriched proposi-

⁴⁶Chalmers has in other places, e.g. in Chalmers 2006, emphasized that the world component in a scenario is an *epistemically possible world*.

tion of (an utterance of) the sentence ‘Hesperus is Phosphorus’ is the structure

$$[=' / =, \langle h' / v, p' / v \rangle]$$

where $='$ is the primary intension of $'='$ and p' the primary intension of ‘Phosphorus’. Enriched propositions are evaluated both at scenarios and at worlds. An enriched proposition is true at a scenario iff the associated primary proposition (extracted by composing the primary intensions of the components) is true at that scenario, and it is true at a world iff the associated secondary proposition (extracted by composing the extensions of the components) is true at that world (2011, p. 601).

The *endorsement* relation is introduced simply as a technical counterpart to the belief relation: when someone has a belief with primary intension i_1 they *endorse* i_1 , and analogously for secondary intensions i_2 and enriched propositions p (2011, p. 604). Later on, however, a distinction is made in that one can believe a proposition indirectly that one does not endorse, by way of endorsing a proposition coordinate with it. This is formally stated in the (B) principle (2011, p. 608):

$$(B) \quad B(i, p) \text{ iff } \exists q(E(i, q)) \ \& \ C(q, p)$$

where C is the coordination relation.

Concerning the *coordination* relation, Chalmers says

To a first approximation, we can say that p is coordinate with q iff (i) p and q have the same Russellian component and (ii) p determines an S -appropriate primary intension, where S is the sentence used to express q (Chalmers 2011, p. 609).

So, coordination is to be understood in terms of S -appropriateness, regarding which Chalmers says:

“ S -appropriate” is used to accommodate the fact that the truth of the ascription is sensitive to the primary intensions of the subject’s beliefs (as example (2) suggests) but does not require a belief with a specific primary intension (as examples (5) and (6) suggest). Rather, it seems to require a belief with a primary intension that falls into a certain class. For a subject to believe that Clark Kent is handsome, the relevant belief must pick out Clark under a ‘Clark Kent’-appropriate primary intension, which may allow various sorts of intension associated with Clark’s role as reporter and ordinary citizen, for example, but not a primary intension associated with his role as Superman.

Except for application examples, we don’t get further elucidation, and Chalmers remarks that no account of what it is for a primary intension to count as S -appropriate is “yet

close to hand” (2011, p. 609). We can only conclude that to serve its intended function, it must be some kind of similarity relation on primary intensions.

The main motivations for the main features are, firstly, that structured meanings are required for fineness of grain (2011, p. 601), and, secondly, that in some cases, the primary intension is not the relevant parameter for the intuitively correct truth-conditions. This is exemplified with

(30) Twin-Oscar believes that water is wet.

(31) Fred believes that I am hungry.

(Chalmers 2011, p. 605, examples (3b) and (5)). On the centred-worlds approach, if Fred believes the centred-worlds proposition expressed by ‘I am hungry’, as uttered by Chalmers, he believes that he, Fred, is hungry (even if he believes the utterance is made by Chalmers).⁴⁷ It is not altogether clear why the coordination account on primary intension should not take care of this, but here Chalmers clearly thinks that it doesn’t.

The example (30) has the usual Twin-Earth background. According to Chalmers, Twin-Oscar cannot believe that water is wet even if he has a belief with the same primary intension as his duplicate Oscar on Earth (Chalmers takes primary intensions to play the role of so-called narrow contents - Chalmers 2011, p. 605), since he would need an appropriate environment (with H₂O rather than XYZ), which according to the example he doesn’t. Here I don’t understand why the simple understanding of “water” as equivalent to “local watery stuff”, which seems in agreement with the two-dimensionalist interpretation, does not yield the correct attribution:

(32) Twin-Oscar believes that the local watery stuff is wet.

is ambiguous between reading ‘local’ as *local-to-the-speaker*, which yields a false attribution, and *local-to-the-attributee*, which yields a true attribution. Anyway, Chalmers thinks that the secondary intension is needed to get the right result for (30). Therefore, sometimes the primary intension is relevant, and sometimes the secondary, and that is why we need the enriched versions.

These are the basics of Chalmers’s account. There are some obvious similarities with the present account, since Chalmers appeals both to a kind of structured meanings and to kinds of possible-worlds intensions. Below, I’ll discuss three main aspects of the relation between the two: belief as a relation, compositionality, and structures as objects of endorsement.

⁴⁷This is related to problem of *de se* communication on the centred-worlds approach. For recent discussions, see Pagin 2016b and other contributions in García-Carpintero and Torre 2016 .

Belief as a relation

The present account combines structured meanings with a possible-worlds account of belief, i.e. a Hintikka-style treatment of belief attribution as a modal operation. By contrast, Chalmers understands belief attribution as a relation between subjects and enriched propositions. By the (B) principle, this reduces to two other relations, those of endorsement, between subjects and enriched propositions, and the coordination relation between propositions.

What happens then if the enriched intension of 'believes', or 'endorses', itself is already a *constituent* of the enriched proposition? This situation arises, of course, in iterated belief attributions, such as

(33) John believes that Mary believes that Hesperus is Phosphorus.

(example in Chalmers 2011, p. 623). If the enriched intension of 'believes' is defined in terms of the relation between subjects and enriched propositions, as in Chalmers's account, then this relation will be non-well-founded, since it will have to be defined partly in terms of itself.

This problem arises in general for theories that treat belief as a relation between subjects and propositions. It is noted by Mark Richard (1990, pp. 145–6). For his own relational theory of belief, Richard proposes as a solution a type hierarchy of belief relations, where at each level in the hierarchy the belief relation at that level is defined only for propositions that contain lower-level belief relations. The verb 'believes' itself is then systematically ambiguous; each occurrence expresses a relation at some level in the hierarchy. This proposal can clearly run into trouble in cases of simultaneous belief attributions ('Everyone believed what his mother had told him'), where there need be no unique level that is applicable to each instance of believing, even though such an attribution need not be paradoxical in any other way.

Chalmers does not take note of the problem. He does, however, provide a treatment of iterated belief attributions, such as (33), in terms of an intensional hierarchy, in parallel to Frege's hierarchy of indirect senses. For this purpose, Chalmers introduces an *ascension* operator, which he explains as follows:

The ascension operator A maps an enriched proposition p to an enriched intension $A(p)$ whose extension is p . If in ' N believes that S ', S expresses the enriched proposition p , then 'that S ' expresses the enriched intension $A(p)$, or simply $\langle p \rangle$. Then ' N believes that S ' expresses the enriched proposition $[B(n, \langle p \rangle)]$, where B and n are the enriched intensions of 'believes' and N .

What is the ascension operator, in the two-dimensional framework? Given an enriched proposition p , $A(p)$ will be an ordered pair $[A'(p)/p]$. The primary intension

$A'(p)$ will be a function from scenarios to enriched propositions, mapping a scenario s to the enriched proposition $[p'/p'(s)]$, where p' is the primary intension associated with p . So $A(p)$ will be wholly determined by p . Although it is not needed above, one can also define an ascension operator mapping arbitrary enriched intensions i to enriched intensions $A(i)$ whose extension is i in an analogous manner. (The primary intension associated with $A(i)$ will be a function mapping s to the enriched intension $[i'/i'(s)]$, where i' is the primary intension associated with i (Chalmers 2011, p. 622).)

Chalmers points out that that the ascension operator forms canonical primary intensions that do offer a road “back from reference to sense”, thus avoiding a common criticism of Frege’s hierarchy.⁴⁸ But precisely because $A(p)$ depends on and is fully determined by p , we cannot define the intensions for ‘believes’ as applied to $A(p)$, if the enriched proposition p already has the primary intension of ‘believes’ as a constituent, for the primary intension of ‘believes’ is then needed for determining $A(p)$. Hence, the belief relation problem remains.

Note that this problem does not arise within the Hintikka type of belief account. The problem with belief relation theories is that the interpretation of ‘believes’ depends on the interpretations of embedded sentences which themselves in many cases depend on the interpretation of ‘believes’. On the doxastic logic approach, the meaning of the belief operator is given by the recursive clauses of the semantics and does not depend on the interpretation function. What we need to know for the truth-conditions of (33) is the accessibility relations, that is the set of John-belief-alternatives to the actual world, and for each member of that set, the set of Mary-belief-alternatives to it. The accessibility relations themselves do not depend on the interpretation function either.

Compositionality and switching

Chalmers’s principle (CEP) is stated for truth simpliciter, which most probably amounts to truth at the actual world (not a scenario, since no center is mentioned), which in turn is a special case of truth at a world. Clearly, since the truth value according to (CEP) depends on both the primary and secondary intension (or extension) of the embedded sentence, the semantics is not compositional for the semantic notion of *truth at a world*.

⁴⁸In fact, Chalmers’s ascension operator is a two-dimensional analogue to Montague’s hat operator ‘ $\hat{\cdot}$ ’ (Montague 1973, p. 256). In Montague’s PTQ, if expression α is of type a , and therefore denotes a meaning of type a , then expression ‘ $\hat{\alpha}$ ’ is of type $\langle s, a \rangle$, and denotes the *sense* of α , a function from world-time pairs to denotations of α at those pairs (1973, p. 259). The indirect sense is therefore determined by the sense itself. More precisely, since ‘ $\hat{\alpha}$ ’^{2*l*,*i*,*j*,*g*} is that function h that applied to a pair $\langle i', j' \rangle$ gives ‘ α ’^{2*l*,*i'*,*j'*,*g*} as value, and since this value does not depend on the pair $\langle i', j' \rangle$, the sense of the sense of α is a constant function from indices that gives the sense of α as value at each index. Although the sense of the sense of α is distinct from the sense of α , it is determined by the sense of α . Hence, in PTQ as well, there is a road from reference to sense, when the reference is itself a sense.

This feature is emphasized by Chalmers himself, who says that on a “broadly Fregean account” one must deny *extensional compositionality*, “the thesis that the extension of a complex expression is always determined by the extensions of its parts” (Chalmers 2011, p. 621). Of course, this situation does not arise because of belief sentence or hyperintensionality more generally; the same holds whenever an *intensional* operator is involved: the truth value of $\Box p$ at a world depends on the intension of p .

In a purely intensional framework, we still have *intensional compositionality*, since the intension of a sentence is determined by the intensions of its parts (and the mode of composition). But the two-dimensional framework cannot be intensionally compositional either. Neither the primary nor the secondary intension of an expression is in general determined by the other. Hence, since the truth of a belief sentence at a world depends on both the primary and the secondary intension of the embedded sentence, secondary intension is not a compositional semantic value. Analogously, since the truth at a scenario of a belief sentence depends on the secondary intension, primary intension cannot be a compositional semantic value either.

Prima facie, it does look as if a kind of semantic switching, in the sense of section 3, does take place, precisely because the the primary intension of a belief sentence depends on the secondary intension of the embedded sentence, and the secondary sentence of a belief sentence depends on the primary intension of the embedded sentence. The semantics would then turn out to be general compositional (in the sense of chapter 3).

Chalmers does indeed say that his account is “compatible with compositionality in a broader sense, in that the meaning of a complex expression is determined by the meanings of its parts” (Chalmers 2011, p. 621). He immediately adds that “the enriched intensions of a complex expression is determined by the meanings of its parts”. This is of course just the standard notion of compositionality as applied to semantic value of *enriched intension*. Chalmers’s claim is indeed correct: the semantics is standardly compositional for enriched intension. This does not, however, entail that switching does not take place.

Any general compositional system $\mathcal{S} = (S, \mu_0, \Psi)$ can be trivially reformulated by *adding* a standard compositional semantic property that consists of an *ordered tuple* $\mu^* = \langle \mu_0, \mu_1, \dots, \mu_n \rangle$ (where $S = \{\mu_0, \mu_1, \dots, \mu_n\}$). The property is defined such that

$$\mu^*(t) = \langle \mu_0(t), \mu_1(t), \dots, \mu_n(t) \rangle.$$

To state the semantics we then need *projection functions* $h_0, h_1, \dots, h_n$ that pick the relevant component of μ^* to apply to a term: $h_i(\mu^*(t)) = \mu_i(t)$. Now, in a semantic system where switching does *not* take place, we would have for a complex expression $\alpha(t)$ the

result that for for each $0 \leq i \leq n$,

$$(h_i(\mu^*))(\alpha(t)) = r_{\alpha,i}((h_i(\mu^*))(t))$$

for some meaning operation $r_{\alpha,i}$, while in a system *with* switching, we will have

$$(h_i(\mu^*))(\alpha(t)) = r_{\alpha,i}((h_j(\mu^*))(t))$$

with $i \neq j$. Thus, that μ^* is formally *standard* compositional is fully compatible with the fact that switching is going on in the computing of μ^* by means of computing its component parts. The system as a whole is not standard compositional, since some of the components of μ^* are not.

Almost exactly this is going on with Chalmers's enriched meanings (cf. the quote above from p. 622), which have the primary and secondary intensions as the component parts, except that this semantics is even more general: the value of a component semantic function (e.g. the primary intension) for some complex expressions–belief sentences–depends both and its own value for a sub-expression *and* the value of another component function (the secondary intension). This goes even beyond ordinary switcher semantics.⁴⁹

Endorsing structures

Chalmers introduces 'endorsement' as a technical term for a relation between a subject and an enriched proposition. As we saw above, the explanation of the term is rather thin. Endorsement appears to differ from belief in that belief in an enriched proposition p can be indirect, via endorsement of a p -coordinate proposition p' . This relative sparseness of technical detail in the explanation of endorsement invites the interpretation that the enriched proposition is simply the *object of endorsement*, i.e. *what is endorsed*, just as in other theories ordinary possible-worlds propositions, or Fregean thoughts, are the objects of belief.

The idea is problematic in light of arguments that structured entities such as ordered

⁴⁹A different kind of two-dimensional switching takes place in Stalnaker's account of belief sentences (Stalnaker 1987 and 1988). Stalnaker's idea is that the belief context triggers a shift of the *propositional concept*. The propositional concept is a two-dimensional matrix, with alternative worlds of utterance on the vertical axis and alternative worlds of evaluation on the horizontal axis. The propositional concept for an utterance tells you which proposition the utterance would express in a particular world of utterance and what truth value that proposition has in a particular world of evaluation. The belief context triggers a shift from the propositional concept of the *speaker's* utterance to the proposition concept of a would-be utterance by the *attributee* of the embedded sentence (or a suitable grammatical variant). This is not strictly speaking a case of semantic switching, since the propositional concept, on Stalnaker's account, is not itself a semantic property.

tuples cannot be bearers of truth, nor objects of belief, rehearsed in section 1.

If an ordered pair, for instance an enriched proposition, does not *say* anything, since it does not have any intrinsic truth-conditions, then it cannot be the object of belief either, as far as Chalmers has explained the notion, nor endorsed.⁵⁰

References

- Bealer, George (1993). 'A Solution to Frege's Puzzle'. In: *Philosophical Perspectives* 7, pp. 17–60.
- Brogaard, Brit (2012). *Transient Truths: An Essay in the Metaphysics of Propositions*. Oxford: Oxford University Press. Electronic resource at SUB.
- Burdick, Howard (1982). 'A Logical Form for the Propositional Attitudes'. In: *Synthèse* 52, pp. 185–230.
- Burge, Tyler (1978). 'Belief and Synonymy'. In: *Journal of Philosophy* 75, pp. 119–38.
- (1979). 'Individualism and the mental'. In: *Midwest Studies in Philosophy*, 4. Ed. by P A French, T E Uehling, and H K Wettstein. Minneapolis: University of Minnesota Press, pp. 73–121.
- Carnap, Rudolf (1947). *Meaning and Necessity*. 1st ed. Chicago, Ill.: The University of Chicago Press. Second edition published in 1956.
- (1956). *Meaning and Necessity*. 2nd ed. Chicago, Ill.: The University of Chicago Press. First edition published in 1947.
- Chalmers, David (2006). 'The Foundations of Two-Dimensional Semantics'. In: *Two-Dimensional Semantics*. Ed. by Manuel García-Carpintero and Josep Macià. Oxford: Oxford University Press, pp. 55–140.
- (2011). 'Propositional Attitude: A Fregean Account'. In: *Noûs* 45, pp. 595–639.
- Church, Alonzo (1950). 'On Carnap's Analysis of Statement of Assertion and Belief'. In: *Analysis* 10, pp. 97–9.
- (1951). 'A Formulation of the Logic of Sense and Denotation'. In: *Structure, Method, and Meaning: Essays in Honor of Henry M. Scheffer*. Ed. by Paul Henle, H. M. Kallen, and S. K. Langer. Liberal Arts Press, pp. 3–24.
- (1954). 'Intensional Isomorphism and Identity of Belief'. In: *Philosophical Studies* 5, pp. 65–73.

⁵⁰Eric Johannesson (2017, Chapter 6) has developed a technically more detailed account of belief sentences that combines two-dimensional semantics and structured meanings. Johannesson elaborates Chalmers's *coordination* relation into a relation called *alignment*. Johannesson still takes subjects to be belief-related to structured propositions, and hence inherits the endorsement problems noted above (in addition, noticing related problems). It may also be remarked that the semantics for unstructured content provided by Johannesson (2017, pp. 102–3) does contain a switcher clause for the belief operator, similar to the one proposed here.

- Cresswell Max, J (1985). *Structured Meanings. The Semantics of Propositional Attitudes*. Cambridge, Mass.: MIT Press.
- Cresswell Max, J and A von Stechow (1982). 'De re Belief Generalized'. In: *Linguistics & Philosophy* 5, pp. 503–35.
- Crimmins, Mark (1992). *Talk About Belief*. Cambridge, Mass.: MIT Press.
- Davidson, Donald (1963). 'The Method of Extension and Intension'. In: *The Philosophy of Rudolf Carnap*. Ed. by Paul Arthur Schilpp. The Library of Living Philosophers. La Salle, Ill.: Open Court, pp. 311–49.
- Dummett, Michael (1973). *Frege: Philosophy of Language*. London: Duckworth.
- (1991). *The Logical Basis of Metaphysics*. Cambridge, Mass.: Harvard University Press.
- Egan, Andy (2007). 'Epistemic Modals, Relativism, and Assertion'. In: *Philosophical Studies* 133, pp. 1–22.
- Forbes, Graeme (1990). 'The Indispensability of Sinn'. In: *The Philosophical Review* 99, pp. 535–63.
- (1996). 'Substitutivity and the Coherence of Quantifying In'. In: *The Philosophical Review* 105, pp. 337–72.
- Frege, Gottlob (1892). 'Über Sinn und Bedeutung'. In: *Translations from the Philosophical Writings of Gottlob Frege, 1952*. Ed. by Peter Geach and Max Black. Oxford: Blackwell, pp. 56–78.
- García-Carpintero, Manuel and Stephan Torre, eds. (2016). *About Oneself*. Oxford: Oxford University Press.
- Glüer, Kathrin and Peter Pagin (2006). 'Proper names and relational modality'. In: *Linguistics & Philosophy* 29, pp. 507–35.
- (2008). 'Relational Modality'. In: *Journal of Logic, Language and Information* 17, pp. 307–22.
- (2012). 'General Terms and Relational Modality'. In: *Nous* 46, pp. 159–99.
- (2014). 'Vulcan Might Have Existed, and Neptune Not. On the Semantics of Empty Names.' In: *Empty Representation. Reference and Non-Existence*. Ed. by Genoveva Martí and Manuel García-Carpintero. Oxford: Oxford University Press, pp. 117–41.
- Hanks, Peter (2015). *Propositional Content*. Oxford: Oxford University Press.
- Higginbotham, James (1986). 'Linguistic Theory and Davidson's Program in Semantics'. In: *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Ed. by Ernest LePore. Oxford: Blackwell, pp. 29–48.
- (2003). 'Conditionals and compositionality'. In: *Philosophical Perspectives* 17, pp. 181–94.
- Hintikka, Jaakko (1962). *Knowledge and Belief*. Ithaca, NY: Cornell University Press.

- Hodges, Wilfrid (2001). 'Formal Features of Compositionality'. In: *Journal of Logic, Language and Information* 10, pp. 7–28.
- Jespersen, Bjørn (2003). 'Why the Tuple Theory of Structured Propositions isn't a Theory of Structured Propositions'. In: *Philosophia* 31, pp. 171–83.
- Johannesson, Eric (2017). 'Analyticity, Necessity, and Belief. Aspects of Two-Dimensional Semantics'. PhD thesis. Department of Philosophy, Stockholm University.
- Kaplan, David (1968). 'Quantifying in'. In: *Synthese* 19, pp. 178–214.
- (1989). 'Demonstratives'. In: *Themes from Kaplan*. Ed. by J Almog, J Perry, and H Wettstein. Oxford: Oxford University Press, pp. 481–563. Electronic resource at SUB.
- King, Jeffrey C (2014). 'Structured Propositions'. In: *Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Stanford University. URL: <http://plato.stanford.edu/archives/spr2014/entries/propositions-structured/>.
- Kripke, Saul (1979). 'A Puzzle About Belief'. In: *Meaning and Use*. Ed. by Asa Kasher. Dordrecht: Reidel, pp. 239–83. Reprinted in Kripke 2011, 125–61. Page references to the reprint.
- (1980). *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.
- (2011). *Philosophical Troubles*. Vol. 1. Collected Papers. Oxford: Oxford University Press.
- Larson, Richard K and Peter Ludlow (1993). 'Interpreted Logical Forms'. In: *Synthese* 95, pp. 305–55.
- Leslie, Sarah-Jane (2009). '"If", "Unless", and Quantification'. In: *Compositionality, Context, and Semantic Value*. Ed. by Robert J. Stainton and C. Viger. Berlin: Springer, pp. 3–30.
- Lewis, David (1970). 'General Semantics'. In: *Synthese* 22, pp. 18–67. Reprinted in Lewis 1983a, 189–232.
- (1979a). 'Attitudes *de dicto* and *de se*'. In: *The Philosophical Review* 88, pp. 513–43. Reprinted in Lewis 1983a, 133–56.
- (1979b). 'Attitudes *de dicto* and *de se*'. In: *The Philosophical Review* 88, pp. 513–43. Reprinted in Lewis 1983b, 133–56.
- (1980). 'Index, Context, and Content'. In: *Philosophy and Grammar*. Ed. by Stig Kanger and Sven Öhman. Dordrecht: Reidel, pp. 79–100. URL: http://www.andrewmbailey.com/dkl/Index_Context_Content.pdf. Available at Andrew M Bailey's page.
- (1983a). *Philosophical Papers. Volume I*. Oxford: Oxford University Press.
- (1983b). *Philosophical Papers. Volume I*. Oxford: Oxford University Press. Electronic resource at SUB.

- Linsky, Leonard, ed. (1952). *Semantics and the Philosophy of Language*. Urbana, Ill: University of Illinois Press.
- Majid, Asifa (2010). 'Words for Body Parts'. In: *Words and The Mind. How Words Capture Human Experience*. Ed. by Barbara C. Malt and Philip Wolff. Oxford University Press. Chap. 3, pp. 58–71.
- Mates, Benson (1950). 'Synonymity'. In: *University of California Publications in Philosophy* 25. Reprinted in Linsky 1952, pp. 201–26.
- Montague, Richard (1968). 'Pragmatics'. In: *Formal Philosophy. Selected papers by Richard Montague*. Ed. by Richmond H Thomason. edited and with an introduction by Richmond H Thomason. New Haven: Yale University Press, pp. 95–118. Originally published in Raymond Klibansky (ed.), *Contemporary Philosophy: A Survey*, La Nuova Italia Editrice, Florence, 1986, pp. 102–22.
- (1973). 'The Proper Treatment of Quantification in Ordinary English'. In: *Approaches to Natural Language*. Ed. by Jaakko Hintikka, J. Moravcsik, and P. Suppes. Dordrecht: Reidel. Reprinted in Montague 1974, pp. 247–70. Page references to the reprint.
- (1974). *Formal Philosophy. Selected papers by Richard Montague*. edited and with an introduction by Richmond H Thomason. New Haven: Yale University Press.
- Nolan, Daniel P. (2013). 'Impossible worlds'. In: *Philosophy Compass* 8.4, pp. 360–72.
- Pagin, Peter (1992). 'Names in and out of thought'. In: *Philosophical Studies* 66, pp. 27–51.
- (2003). 'Communication and strong compositionality'. In: *Journal of Philosophical Logic* 32, pp. 287–322.
- (July 2016a). 'A General Argument Against Structured Propositions'. In: *Synthese On-line*. DOI: [10.1007/s11229-016-1244-4](https://doi.org/10.1007/s11229-016-1244-4).
- (2016b). 'De se Communication: Centered or Uncentered?' In: *About Oneself*. Ed. by Manuel García-Carpintero and Stephan Torre. Oxford: Oxford University Press, pp. 272–306.
- (2017a). 'Compositionality in Davidson's Early Philosophy'. Under review.
- (Jan. 2017b). 'When does communication succeed? The case of general terms.' Draft.
- Pagin, Peter and Dag Westerståhl (2010a). 'Compositionality I: Definitions and Variants'. In: *Philosophy Compass* 5, pp. 250–64.
- (2010b). 'Pure Quotation and General Compositionality'. In: *Linguistics and Philosophy* 33, pp. 381–415.
- Pelletier, Francis Jeffry (1994a). 'On An Argument Against Semantic Compositionality'. In: *Logic and Philosophy of Science in Uppsala*. Ed. by Dag Prawitz and Dag Westerståhl. Dordrecht: Kluwer Academic Publishers, pp. 599–610.

- Pelletier, Francis Jeffry (1994b). 'The Principle of Semantic Compositionality'. In: *Topoi* 13. Expanded version reprinted in S. Davis and B. Gillon (eds.) *Semantics: A Reader* (2004), pp. 133-157, pp. 11-24.
- Perry, John (1977). 'Frege on Demonstratives'. In: *The Philosophical Review* 86, pp. 474-97.
- Putnam, Hilary (1954). 'Synonymity, and the Analysis of Belief Sentences'. In: *Analysis* 14, pp. 114-22.
- Quine, Willard Van Orman (1956). 'Quantifiers and propositional attitudes'. In: *Journal of Philosophy* 53, pp. 177-87. Reprinted Quine 1976.
- (1976). *The Ways of Paradox and Other Essays*. 2nd ed. Cambridge, Mass.: Harvard University Press.
- Richard, Mark (1981). 'Temporalism and eternalism'. In: *Philosophical Studies* 39, pp. 1-13.
- (1990). *Propositional Attitudes. An Essay on Thoughts and How We Ascribe them*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press.
- Russell, Bertrand (1903). *Principles of Mathematics*. 2nd ed. Routledge. Available at the Internet Archive.
- Salmon, Nathan (1986). *Frege's Puzzle*. Cambridge, Mass.: MIT Press.
- Schiffer, Stephen (1992). 'Belief Ascription'. In: *Journal of Philosophy* 89, pp. 499-521.
- Sellars, Wilfrid (1955). 'Putnam on Synonymy and Belief'. In: *Analysis* 15, pp. 117-20.
- Soames, Scott (1987). 'Direct Reference, Propositional Attitudes, and Semantic Content'. In: *Philosophical Topics* 15, pp. 47-87.
- (2010). *What is Meaning?* Princeton, N.J.: Princeton University Press.
- (2014). 'Cognitive Propositions'. In: *New Thinking About Propositions*. Oxford: Oxford University Press. Chap. 6, pp. 91-124. Electronic resource at SUB.
- (2015). *Rethinking Language, Mind, and Meaning*. Carl G. Hempel Lecture Series. Princeton University Press.
- Stalnaker, Robert (1987). 'Semantics for Belief'. In: *Philosophical Topics* 15, pp. 177-90. Reprinted in Stalnaker 1999, pp. 117-29. Page references to the reprint.
- (1988). 'Belief Attribution and Context'. In: *Contents of Thought*. Ed. by Robert Grimm and Daniel Merrill. University of Arizona Press, pp. 140-55. Reprinted in Stalnaker 1999, pp. 150-66. Page references to the reprint.
- (1999). *Context and Content*. Oxford: Oxford University Press.
- (2014). *Context*. Oxford: Oxford University Press.
- Stojanovic, Isidora (2012). 'The Problem of De Se Assertion'. In: *Erkenntnis* 76, pp. 49-58.
- Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.