Subject obviation and self-locating knowledge

Julie Goncharov

(accepted for a special volume of *Lingvisticae Investigationes*)

Abstract

Subject obviation is a restriction on having coreferential subjects in sentences like *#Je veux que je parte* 'I want that I leave'. Since early 1980s, linguists tried to explain this restriction, assigning it to different domains of formal grammar (syntax, semantics, semantic-pragmatic interface). In this paper, I defend the view that subject obviation is due to a property of "general intelligence" rather than linguistic competence (see Ruwet 1984). In my proposal, the role of "general intelligence" is played by the principle of non-triviality that disallows ascription of a proposition whose content does not exclude any possibility from a belief state. I formulate my proposal in the framework developed by Robert Stalnaker where belief states include information about how the believer locates herself in the world (Stalnaker 2008, 2014). Since embedded propositions in sentences with subject obviation involve self-locating information, I argue, their use in attitude ascriptions violates the principle of non-triviality. That is to say, sentences with subject obviation are deviant for the same reason as sentences like *#I believe I am sane* when the speaker has no doubts about her sanity. I conclude by discussing how my proposal allows us to see in a fresh light some phenomena that are problematic for grammar-based accounts.

1 Introduction

Subject obviation is a restriction on having coreferential subjects in some attitude ascriptions with finite complement clauses.[1] For example, *vouloir* 'want' in French can be used with a subjunctive complement, but not when the subject of the main clause and the embedded subject refer to the same individual, compare (1)a with (1)b. This restriction is not absolute. Coreferential constructions can be used when the interpretation of the subject in the main clause does not fully align (in the sense to be discussed) with the interpretation of the embedded subject, see (1)c.

(1) a. #Je veux que je parte. (French)
'I want that I leave-SBJV'
b. Je veux que tu partes.
'I want that you leave-SBJV'

[1]The initial name for the phenomena was 'subjunctive obviation'. 'Subject obviation' is a recent label that is used for a broader range of constructions (e.g., Kaufmann 2019; Stegovec 2019). Although my focus in this paper is on what was previously called 'subjunctive obviation', I will use the name 'subject obviation' for two reasons. First, I believe, a unified account with other similar phenomena is justified. And, second, as we will see, 'subjunctive obviation' is a misnomer since the restriction is not limited to subjunctives.

c. Je veux absolument que j'amuse ces enfants.
'I want absolutely that I amuse-SBJV these children' (Ruwet 1991: 2, 30)

Subject obviation is found in different languages and with a variety of attitude predicates.[2] The examples in (2) and (3) show subject obviation in Italian and Hungarian. The sentences in (2)a and (3)a have the same structure as the French example in (1)a and are judged unacceptale. I will call such sentences *obviative*. I will call sentences with non-coreferential subjects, like in (1)b and (2)b, *non-obviative*, and sentences that have coreferential subjects, but are not judged unacceptable, like in (1)c, (2)c, and (3)c, *ameliorated*.

(2) a. #Penso che io parta domani. (Italian)
'I think that I leave-SBJV tomorrow'
b. Penso che parta domani.
'I think that he/she leave-SBJV tomorrow'
c. Penso che io abbia fatto molti errori.
'I think that I have-SBJV made many mistakes' (Costantini 2023)

(3) a. #Remélem, hogy (nem) szédülök (Hungarian)
'I hope that I (don't) have-IND vertigo'
b. Remélem, hogy nem untatlak.
'I hope that I not bore-IND you' (Szabolcsi 2021)

Generative linguists have studied subject obviation since early 1980s and there are many things that we now know about this restriction. We know that the restriction is not due to some syntactic condition, because ameliorated sentences come with varying degrees of acceptability, whereas syntactic well-formedness judgements are assumed to be categorical (Ruwet 1984). We also know that subject obviation is not due to a competition between infinitival and finite complements, because in some languages and constructions, subject obviation obtains when there is no infinitival alternative to compete with (Szabolcsi 2021; Costantini 2023). Nor can subject obviation be explained as a competition between the subjunctive and the indicative mood, since subject obviation is attested with both (Szabolcsi 2021). Finally, we know that subject obviation is not directly related to decisive modality (desires, hopes, commands), since it is also found with epistemic attitudes (Costantini 2016, 2023).

Most analyses of subject obviation have assigned the restriction to some domain of formal grammar (syntax, semantics, semantic-pragmatic interface). In this paper, I defend the view that subject obviation is due to a property of "general intelligence" rather than linguistic competence that was initially articulated in Ruwet (1984). In my proposal, the role of "general intelligence" will be played by the principle of non-triviality that disallows ascription of a proposition whose content does not exclude any possibility from a belief state. I formulate my proposal in the framework developed by Robert Stalnaker (Stalnaker 2008, 2014).

The structure of the paper is as follows: In the next section, I will discuss the explanation of subject obviation given by Ruwet (1984). This discussion is important, because my own explanation of subject obviation will share two key features with Ruwet's explanation. First, like Ruwet, I will attribute subject obviation to a property of "general intelligence". This feature will be the topic of

[2](E.g., Bouchard 1983; Picallo 1985; Ruwet 1984; Kempchinsky 1986, 2009; Farkas 1992; Costantini 2006, 2016, 2023; Schlenker 2005; Szabolcsi 2010, 2021; Stegovec 2019)

section 3. Second, also like Ruwet (and a number of scholars after him), I will say that obviative sentences are unacceptable, because there is an illegitimately tight "self-to-self" relation between the two coreferential subjects. This second feature is discussed in section 4. In section 5, I will show how subject obviation is derived using the apparatus introduced in sections 3 and 4. Then, in section 6, I compare my proposal with current semantic-pragmatic accounts of subject obviation and discuss what I believe to be the advantages of my proposal in section 7. Section 8 summarizes the paper.

2 Ruwet's explanation

I begin with the explanation of subject obviation in Ruwet (1984).[3] Although Ruwet's study is one of the first on the topic and since then, we have learned more about subject obviation and grammar, in general, his explanation has a feature that recent accounts lack. It attributes subject obviation to a property of "general intelligence", rather than linguistic competence. I divide Ruwet's explanation into four points. With the first and the last points, I will agree without reservation. The second point was shown to be wrong by the studies that came after Ruwet's; so, I will reject it. The third point is what distinguishes Ruwet's explanation from other accounts of subject obviation. It holds that subject obviation is due to a property of "general intelligence" and it will be central to my explanation as well. I will agree with Ruwet's claim broadly construed, but disagree about the specific mechanism of "general intelligence" that is the source of subject obviation.

Ruwet's explanation of subject obviation goes as follows: First, he maintains that obviative sentences, as in (4)a repeated from above, are syntactically well-formed. The main empirical contribution of Ruwet's work is that he identifies a range of strategies that can improve acceptability of obviative sentences. These strategies include (among others) the use of passives or modals and changing the Aktionsart or the viewpoint aspect, see (4)b-e. If (4)a were syntactically ill-formed, Ruwet argues, it would have been difficult to explain different degrees of acceptability of ameliorated sentences in (4)b-e, because syntactic well-formedness is, presumably, a categorical judgment.

(4) a. #Je veux que je parte. (French)
'I want that I leave-SBJV'
b. ?Je veux que je sois enterré dans mon village natal.
'I want that I be-SBJV buried in the village of my birth'
c. ?Je veux que je puisse attaquer à l'aube.
'I want that I can-SBJV attack at dawn'
d. ?Je veux que je réussisse.
'I want that I succeed-SBJV'
e. Je veux (absolument) que je sois parti dans dix minutes.
'I want (absolutely) that I be-SBJV gone in ten minutes' (Ruwet 1991)

Second, Ruwet argues that the unnaturalness of (4)a comes from the fact that it competes with the infinitival construction *Je veux partir* 'I want to leave' that expresses the same thought. As was mentioned in the introduction, the competition conjecture cannot be maintained. Since Ruwet's work, it has been shown that subject obviation obtains in languages and constructions that do not

[3]It was first published as Ruwet (1984) in French and was later translated into English and published as Ruwet (1991). I will refer to this work by its first date of publication, but use the English version for citations.

have a suitable infinitival alternative for an obviative sentence. We will see evidence against the competition conjecture at the end of the section.

Third, the competition between a subjunctive and an infinitive construction, Ruwet continues, is due to a property of "general intelligence or central cognitive processes".[4] The property that he has in mind is "an **iconic** link between the (superficial) form of the sentence – simple or complex [...] – and the content, experienced as relatively simple or relatively complex".[5] In other words, Ruwet hypothesizes that a formally simple infinitival construction like *Je veux partir* tends to be iconically associated with a relatively simple construal, whereas more complex bi-clausal constructions like *Je veux que je parte* tend to be used to convey relatively complex relations. Thus, if we want to express a relatively simple thought, we tend to choose a simple infinitival construction, which makes sentences with subjunctive clauses less natural for expressing that same thought.

Fourth, Ruwet's classification of a construal as "relatively simple" or "relatively complex" has to do with the interpretation of a coreferential relation between the subject of the main clause and the embedded subject. A "relatively simple" construal is when the two subjects "are viewed from fundamentally the same point of view" or "the internal distance between the two instances of the self [brought in by the two subjects – Author] tends to vanish".[6] A "relatively complex" construal is when "the relation of self-to-self [...] involves an internal differentiation and highlights two distinct facets of the self, and/or introduces a certain distance between self and self".[7] This intuition that a "self-to-self" relation can be either "viewed from fundamentally the same perspective" or involve a "shift in perspective"[8] and "distancing between self and self"[9] is also central to most recent accounts of subject obviation. It comes under different guises: as a distinction between a (simple) *de se* and an event *de se* interpretation (Schlenker 2005), as the difference between objective information and "direct experience" (Szabolcsi 2021), or the difference in how knowledge is obtained – through reasoning or introspection (Costantini 2016, 2023). I think the intuition is correct. In this paper, I will formalize it using self-locating knowledge.

As I said earlier, my explanation of subject obviation will share key features with Ruwet's explanation. I agree with Ruwet's reasoning about ameliorated sentences: their varying degrees of acceptability show that the source of subject obviation is not a syntactic restriction. Nor is it, I believe, a semantic or semantic-pragmatic restriction (for the reasons that I will discuss in sections 6 and 7). The source of subject obviation (I again agree with Ruwet) is a property of "general intelligence". In my explanation, the role of "general intelligence" will be played by the principle of non-triviality that disallows ascription of a proposition whose content cannot be used to exclude any possibility from a belief state. This principle will be formulated and discussed in the next section (section 3). I will propose that embedded propositions in obviative sentences are unacceptable, because they violate the principle of non-triviality. What makes embedded propositions in obviative sentences to violate the principle of non-triviality (and here again, I agree with Ruwet's initial insight and similar ideas in recent accounts) is the presence of a tight "self-to-self" relation between the interpretation of the subject of the main clause and that of the embedded clause. This relation,

[4](Ruwet 1991, 19)
[5](Ruwet 1991, 8)
[6](Ruwet 1991, 16)
[7](Ruwet 1991, 16)
[8](Ruwet 1991, 13)
[9](Ruwet 1991, 15)

which I will formulate in terms of self-locating knowledge, will be discussed in section 4. After introducing the principle of non-triviality and self-locating knowledge in the next two sections, in section 5, I will show how this apparatus can help us to explain subject obviation.

In my explanation of subject obviation, there will be no room for iconicity or "an iconic link" which is central for Ruwet's explanation. I put aside the objection that the concept of iconicity is hard to articulate and dismiss its usefulness on the ground that it requires the competition conjecture which has been shown to be wrong. As mentioned above, the competition conjecture says that attitude reports with subjunctive complements are unnatural because they compete with similar and simpler infinitival constructions. But subject obviation has been shown to obtain in languages and constructions that do not have a suitable infinitival alternative for an obviative sentence.

For example, Costantini (2023) shows that subject obviation in Italian holds with epistemic attitudes, see (5)a, and for this construction, it is possible to have an acceptable infinitival alternative, see (5)b. However, there are constructions for which an infinitival alternative is not available. This is the case with periphrastic progressive which in Italian is formed with the auxiliary *stare* 'stay' followed by a gerund. Obviative sentences with periphrastic progressive, see (5)c, do not have a suitable infinitival alternative, because (5)d is ungrammatical due to a language-specific morpho-syntactic restriction. Therefore, it is difficult to explain the unnaturalness of (5)c in terms of the competition conjecture.

(5) a. #Penso che io parta domani. (Italian)
'I think that I leave-SBJV tomorrow'
b. Penso di partire domani.
'I think to leave-INF tomorrow'
c. #Penso che io stia leggendo il giornale.
'I think that I stay-SBJV reading the newspaper'
d. *Penso di star leggendo il giornale.
'I think to stay-SBJV reading the newspaper' (Costantini 2023)

Also, Szabolcsi (2021) argues against the competition conjecture using data from Hungarian. In Hungarian, subject obviation is found with attitude predicates that take either a subjunctive complement, such as *akarni* 'want', see (6)a, or an indicative complement, such as *ramélni* 'hope' and *sajnálni* 'regret/be sorry', see (6)b,c.

(6) a. #Azt akarom, hogy meg-látogassam Marit. (Hungarian)
'I want that I visit-SBJV Mary'
b. #Remélem, hogy fél lábon állok.
'I hope that I stand-IND on one leg'
c. #Sajnálom, hogy ugrándozok.
'I regret that I jump-IND around' (Szabolcsi 2021)

Since only *akarni* 'want' has a corresponding infinitival alternative, but not *ramélni* 'hope' or *sajnálni* 'regret/be sorry', Szabolcsi reasons, the competition conjecture cannot be right. More precisely, I would add, the competition conjecture cannot be the source of subject obviation if we want to develop a unified explanation for subject obviation. But it remains to be seen whether the availability of an infinitival or any other alternative construction is an additional non-necessary contributor to the unacceptability of obviative sentences.

3 Non-triviality, belief ascription, and autonomy of pragmatics

This section and the next one are based on the ideas developed by Robert Stalnaker, the way I understand them. I will use more formalism than Stalnaker usually does; this is to make his ideas easier to use for a formal linguist. In some places, I will have to simplify the picture for reasons of space, but I hope to stay faithful to the spirit of the ideas. Because Stalnaker's ideas are rather on the philosophy side of the linguistics-philosophy divide and may not be all that familiar to the linguistic audience, I will structure my presentation in this section as follows: I will start with the idea that is most familiar to linguists, namely, that the role of an assertion is to exclude from the context in which the assertion is made those possibilities that are incompatible with the asserted proposition. I will call this *a principle of non-triviality*. After that, I will describe Stalnaker's views on belief ascription and how the principle of non-triviality is extended to beliefs. At this point, I will introduce a (partial) knowledge model that we will use later on to talk about subject obviation and I will define the principle of non-triviality in that model. Finally, I will discuss in what sense we can view the principle of non-triviality as a property of "general intelligence". As we will see, it has to do with the thesis of *an autonomy of pragmatics* that Stalnaker defends in his book *Context* (Stalnaker 2014).

In formal semantics, it is standard to assume that propositions are functions from possible worlds to truth values. Stalnaker also makes this assumption. But his conception of a possible world is different from the conception usually assumed in formal semantics, which comes from David Lewis's work. As we will see, this difference affects a number of theoretical choices, so it is worthwhile to keep it in mind throughout the paper. In his recent work, Stalnaker labels Lewis's position on possible worlds 'modal realism' contrasting it with his own position that he calls 'actualism'.[10] In Lewis's view, possible worlds are things of the same sort as our actual world; they differ from the actual world "not in kind but only in what goes on at them".[11] For Lewis, to believe in possible worlds is to believe in "the existence of entities that might be called 'ways things could have been' ".[12] In Stalnaker's view, possible worlds are (mere) possibilities, properties of the universe, relative to which truth is determined. For Stalnaker, to believe in possible worlds is to believe simply that they have a certain structure that "people distinguish [. . .] in their rational activities".[13]

Stalnaker's conception of possible worlds as possibilities distinguished by a rational agent (rather than real entities just like the actual world but not actualized) is fundamental for his contextualism. It allows him to assimilate the context, in which an utterance takes place, to the domain of possible worlds, which determines the meaning of the uttered sentence. Here, propositions are not functions from the domain of all possibilities, but only from the domain of those possibilities that are distinguished in the context of utterance.[14] This, in turn, provides the basis for the position that

[10] In earlier work, Stalnaker describes Lewis's position as 'extreme realism' and his own position as 'moderate realism' (e.g., Stalnaker 1984)

[11] (Lewis 1973, 85)

[12] (Lewis 1973, 84)

[13] (Stalnaker 1984, 57)

[14] In contextualism, the relation between semantics and pragmatics is more complex than in the relativist view, usually assumed in formal semantics where a sentence is evaluated with respect to a model and a context of utterance. This is a complicated issue that I will not take up in this paper. What is important to keep in mind for the purpose of this paper is that for Stalnaker, "context is not just information that mediates between utterance and proposition" either pre- or

Stalnaker is most known for, namely, that the role of an assertion is to exclude from the context in which the assertion is made those possibilities that are incompatible with the proposition expressed by the asserted sentence. The corollary of this statement is that a sentence cannot be asserted in a context which entails the proposition expressed by that sentence. I will call this corollary *the principle of non-triviality*.

In Stalnaker's framework, the principle of non-triviality can be also formulate at the level of belief ascription. This is the level that I will have in mind when talking about the principle of non-triviality in this paper. For Stalnaker, the sentence '*Alice believes that* p' expresses a proposition that is a function of the proposition expressed by the embedded sentence '*that* p'. That is to say, the whole proposition expressed by '*Alice believes that* p' fulfills its assertoric function (i.e., excludes non-compatible possibilities from the context of utterance) in virtue of the embedded proposition '*that* p' excluding non-compatible possibilities from Alice's belief states. Following Stalnaker, I will refer to the context relative to which the whole proposition is evaluated as a *basic* context and to the context for the embedded proposition as a *derived* context. A derived context for '*Alice believes that* p' is determined as follows: For each possible world in the basic context, Alice is in a particular belief state (which itself is a set of possible worlds compatible with Alice's beliefs in that world). The union of these belief states is a set of possible worlds that the speaker takes to be compatible with Alice's beliefs in the basic context. This union is the derived context in which the ascribed proposition expressed by '*that* p' is interpreted.[15] Having in mind the principle of non-triviality, I will say that if the derived context entails the proposition expressed by '*that* p', '*that* p' is not ascribable in that context. As I will argue in the next section, subject obviation obtains when there is an attempt to create an attitude ascription with a non-ascribable proposition. Thus, I propose that the source of subject obviation is a violation of the principle of non-triviality.

To make the principle of non-triviality more concrete, let me introduce a (partial) knowledge model. Our (partial) model here is a tuple $\langle W, S, R\rangle$ where W is a (non-empty) set of possible worlds assuming Stalnaker's conception of possible worlds. As a reminder of this conceptual change, I will use $x, y, z, ...$ as variables for possible worlds (instead of customary $w, w', w'', ...$). I will also use Greek letters $\alpha, \beta, \gamma, ...$ as names for possible worlds when I need to make an illustration (α will be always the actual world). S is a (non-empty) set of individuals or subjects. I will use capital letters $A, B, C, ...$ for members of this set. R is a regular epistemic relation on W (transitive, Euclidean, serial). In this model, let us say that R_A is a Hintikka-style indexed epistemic relation for Alice. Then, for any world x, $\{y : xR_Ay\}$ is a set of possible worlds compatible with what Alice believes to be the case in x. Let us say that the sentence '*Alice believes that* p' is uttered in the basic context C (where C is a set of possible worlds in each of which the sentences was uttered). Then, the derived context for interpreting '*that* p' will be the union of Alice's belief states in C, i.e., $\bigcup_{x\in C}\{y : xR_Ay\}$. Given this, the principle of non-triviality with respect to '*that* p' can be formulated as follows:

(7) *The principle of non-triviality*
a proposition p is ascribable to an individual A in a context set C iff
$\bigcup_{x\in C}\{y : xR_Ay\} \cap p \neq \emptyset$ and $\bigcup_{x\in C}\{y : xR_Ay\} \cap W - p \neq \emptyset$

post-semantically when the meaning of indexicals is determined or implicatures are computed, rather "it is the material out of which propositions are constructed" (Stalnaker 1999, 156)

[15](See Stalnaker 1999, 157)

To say that a proposition is 'ascribable to some individual in a context' is to say that the speaker of the utterance that takes place in that context can use a sentence that expresses this proposition as a complement of an attitude predicate.

Let me illustrate the principle of non-triviality using a simple example. Suppose Alice at some point truthfully says *I believe Mabel is having vertigo*. Suppose we have three possible worlds $\alpha, \beta,$ and γ, such that in α and β, Mabel indeed is feeling dizzy and lightheaded, but not in γ. Suppose further that the context set C consists of two possible worlds α and γ (that is, $C = \{\alpha, \gamma\}$) and that Alice's accessibility relation R_A relates every world to itself and α and β to each other (that is, $R_A = \{\langle\alpha, \alpha\rangle, \langle\alpha, \beta\rangle, \langle\beta, \beta\rangle, \langle\beta, \alpha\rangle, \langle\gamma, \gamma\rangle\}$). To distinguish α from β, let us say that a fair coin was tossed and the results came out as Heads in α and Tails in β and γ, but we are not interested in Alice's belief about the results of the coin flip. This setup is schematized in Figure 1 (where Vm = 'Mabel is having vertigo', H = Heads, T = Tails).

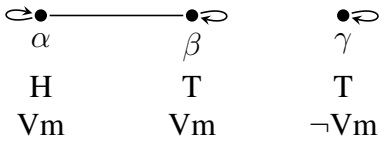

Figure 1: Alice *"I believe Mabel is having vertigo"*

In this setup, the principle of non-triviality is satisfied since the derived context (the union of possible worlds accessible from the basic context set C) is $\{\alpha, \beta, \gamma\}$ and it is compatible with both Mabel having vertigo and Mabel not having vertigo. Also, the belief ascription is true in the actual world α since $\{\alpha, \beta\} \subseteq \{x : \text{Mabel is having vertigo in } x\}$. We can say that the role of the belief ascription is to eliminate from the context set C the possible world γ in which Alice believes that Mabel is not having vertigo.

The principle of non-triviality as it is formulated in (7) bears close resemblance to a more familiar uncertainty presupposition that is argued for some attitude predicates. To see this, suppose that the context set C consists only of one world – the actual world α. Then, the principle of non-triviality reduces to:

(8) a proposition p is ascribable to an individual A in a context set C iff
$\{x : \alpha R_A x\} \cap p \neq \emptyset$ and $\{x : \alpha R_A x\} \cap W - p \neq \emptyset$

I will discuss the connection between the principle of non-triviality and the uncertainty presupposition in section 6. There, I will argue that using the principle of non-triviality puts us in a better position for explaining subject obviation than using the uncertainty presupposition.

The last point to discuss in this section is in what sense the principle of non-triviality is a property of "general intelligence". Recall that following Ruwet, I want to argue that the source of subject obviation is a property of "general intelligence" rather than linguistic competence. To begin with, let me state where the line between "general intelligence" and linguistic competence can be drawn, at least for the purpose of this paper.[16] By "general intelligence", I will mean features and

[16]The discussion of the partition between "general intelligence" and linguistic competence is complex and can lead as beyond what I want to do in this paper. It touches on the most fundamental debates in the philosophy of language having to do with the relation between natural language and the language of thought, the problem of intentionality and how our current theories of meaning address it.

mechanisms that allow us to understand (or theorize about) behaviour and reasoning of rational agents *independently* of the means they are using to express their behaviour or reasoning (including language). That is to say, these are features and mechanisms that we can use to understand (or theorize about) the behaviour and reasoning of non-linguistic creatures with cognitive capacities comparable to ours, as well as the behaviour and reasoning expressed using other means than language. Note that my formulation of the dividing line is at the theoretical level (which mechanism we use to understand or theorize about), and not at the phenomenal level (which phenomena are part of "general intelligence" rather than language). I believe that before we have much better understanding than we have now of the correlation(s) between physiological processes (including those in the brain) and manifested behaviour (including linguistic behaviour), it is premature to insist on any kind of phenomenal division. With this clarification about the partition between "general intelligence" and linguistic competence, it is easy to see why the principle of non-triviality is on the "general intelligence" side of the partition. The principle is formulated in Stalnaker's framework that is explicitly designed to represent speech acts and contents of propositional attitudes in a way that is independent of the "vehicle" (to use Stalnaker's own metaphor) that rational agents use to encode and manifest their thoughts. Stalnaker maintains that "it is possible and fruitful to theorize about the structure and function of discourse independently of specific theory about the mechanisms that languages use to serve those functions".[17] He calls this position a thesis of an *autonomy of pragmatics* and defends it in his 2014 book. One of the outcomes of the autonomy of pragmatics thesis is that there might be linguistically manifested phenomena that are better explained as consequences of general reasoning than of linguistic encodings. In this paper, I suggest that subject obviation is one of such phenomena.

4 Phenomenal information and self-locating knowledge

In the previous section, we discussed the principle of non-triviality which, I want to suggest, is the property of "general intelligence" that is the source of subject obviation. This principle disallows ascription of propositions whose content cannot be used to exclude any possibility from a belief state. In this section, I will discuss what makes embedded propositions in obviative sentences non-ascribable. As mentioned earlier, Ruwet (1984) and a number of scholars after him had an intuition (correct in my view) that obviative sentences are unnatural because there is an illegitimately tight "self-to-self" relation between the subject of the main clause and the subject of the embedded clause. This intuition will be the starting point of our discussion in this section. Then, I will suggest that various ways of capturing this intuition in the literature can be subsumed under the idea that embedded propositions in obviative sentences convey phenomenal information. The goal of the remainder of this section, then, will be to introduce phenomenal information, its representation as self-locating knowledge, and its analysis using the knowledge model from the previous section.

Describing the illegitimately tight "self-to-self" relation between the two coreferential subjects in obviative sentences, Ruwet, first of all, points out that the distancing that the two subjects lack does not amount to different interpretations of indexical expressions, as in the famous McCawley's example *I dreamed that I was Brigitte Bardot and that I was kissing myself*. Nor is it the same as the distinction sometimes made between "a person and his portrait [. . .], an actor and the character

[17](Stalnaker 2014, 1)

he plays" or "an author and his literary work".[18] What he has in mind is a finer distinction between "the soul, broadly construed, and the body and their conflation in the person considered globally" or the distinction drawn by Plato between the rational part of the soul and its desires.[19] These kind of distinctions, he says, can be present even in simple sentences like *I am torn between my love for my family and my love for my country* or *Hey, I am talking to you*. Later on, Anna Szabolcsi will call this discontinuity of the self "mind-boggling".[20]

Schlenker (2005) uses the concept of an event *de se* ('of self') interpretation to describe the tight "self-to-self" relation in obviative sentences, which he borrows from Higginbotham (2003). The examples in (9) illustrate the contrast between a simple *de se* and an event *de se* interpretation. Whereas the sentence in (9)a requires only that the speaker remembers the fact of his going to school in the 5th grade and thus, can be truthfully uttered by any adult, the sentence in (9)b can be asserted only by an adult with an exceptionally good memory who remembers a particular event of his going to school at this young age.

(9) a. I remember that I walked to school in the 5th grade. (simple *de se*)
b. I remember walking to school in the 5th grade. (event *de se*)

Schlenker's explanation of subject obviation is based on the competition conjecture which, as we saw earlier, is not viable. But his intuition that subjunctive clauses in obviative sentences in addition to a simple *de se* interpretation have a perspectival component that comes with an event *de se* interpretation – so that the event is seen from within – and this is what makes obviative sentences unacceptable is, I think, correct and points to the same idea that Ruwet tries to describe in his paper.

Szabolcsi (2021) makes another step towards clarifying the nature of the "self-to-self" relation in obviative sentences. She adds to the picture Hungarian data that directly point to the experiential nature of the information that is conveyed by embedded propositions in obviative sentences. Here are some of her examples:

(10) a. Remélem, hogy benne vagyok a csapatban. (Hungarian)
'I hope that I'm on the team'
b. Remélem, hogy nem untatlak.
'I hope that I'm not boring you'
c. Remélem, hogy biztonságban vagyok.
'I hope that I'm safe'
d. #Remélem, hogy fél lábon állok.
'I hope that I'm standing on one leg'
e. #Remélem, hogy (nem) szédülök.
'I hope that I (don't) have vertigo'
f. #Remélem, hogy (nem) fázom.
'I hope that I'm (not) cold'
g. #Remélem, hogy ugrándozok.
'I hope that I'm jumping around'
h. #Remélem, hogy simogatom a macskát.
'I hope that I'm stroking the cat' (Szabolcsi 2021, 10)

[18](Ruwet 1991, 12)
[19](Ruwet 1991, 12)
[20](Szabolcsi 2021)

All the unacceptable examples in (10) have what Szabolcsi calls "classical predicates of direct experience" which convey agent's own perceptual or cognitive state. This is what according to Szabolcsi makes them obviative.

Finally, discussing subject obviation with epistemic attitudes in Italian, Costantini (2016, 2023) postulates that a sentence is obviative when the embedded proposition "is accessible through introspection". Introspection here is described as something pertaining to self-knowledge that gives the believer "a direct access to mental states and is highly epistemically secure".[21]

I want to suggest that the various ways of describing the illegitimately tight "self-to-self" relation in obviative sentences that we saw above can be subsumed under the idea that embedded clauses in obviative sentences contain *phenomenal information*. 'Phenomenal' is a term that philosophers use to describe a type of knowledge of what it is like to be in a particular cognitive state.[22] For example, what it is like to be in a cognitive state corresponding to seeing the colour red, or feeling cold, or experiencing vertigo. There are several approaches to explaining phenomenal knowledge in philosophy of mind and epistemology; some of them attempt to reduce phenomenal knowledge to more familiar self-locating knowledge (e.g., Perry 1999; Stalnaker 2008). In this paper, I will adopt the reductionist approach developed by Robert Stalnaker. An advantage of his reductionist approach (for us) is that it is framed in a model designed to theorize about linguistic phenomena.

Self-locating information is information about where an agent locates herself in the world. Self-location can be with respect to place, time, or person's own identity. This kind of information is usually associated with the interpretation of indexical expressions (*today*, *here*, *I*) and is contrasted with objective, impersonal information. This contrast is usually illustrated using intricate scenarios involving learned amnesiacs, mountain-dwelling gods, spilled over goods, or cloth items on fire.[23] However, Stalnaker (among others) correctly points out that self-locating knowledge is more ubiquitous than those exotic scenarios make us believe. This kind of knowledge is usually taken for granted and we notice its presence only when we notice its absence. Here's Stalnaker's own mundane example of self-locating knowledge that shows the ubiquity of this kind of knowledge:

> "It is Monday afternoon. After shopping in the mall, I take the elevator down to level B of the parking garage. I had gone up a different elevator, one in the center of the garage. The one I came down is either at the east or the west end, I am not sure which – there is an elevator at each end. I know my car is about in the middle along the northern edge, but is that to the right or to the left? I have a clear mental map of level B, but it has no 'you are here' marker, so I don't know how to orient myself on it. The garage is pretty symmetrical, so it is hard to tell by looking around just where I am. I do know that there is a pale green Prius with Massachusetts license plate [...] to my right as I come out of the elevator, but knowing that does not help, since of course my mental map of level B does not tell me what cars are parked in what places." (Stalnaker 2014, 121)

In this situation, the agent lacks self-locating knowledge because no objective information available

[21] (Costantini 2023)

[22] The most famous illustration of phenomenal knowledge and its (alleged) difference from impersonal knowledge is a thought experiment in which Mary, a brilliant scientist, who knows everything there is to know about the physics and physiology of colour-seeing, but who was raised in a black-and-white world, sees the colour red for the first time (Jackson 1982).

[23] (E.g., Perry 1977; Lewis 1979, among many others)

from that position can allow the agent to choose between two epistemic possibilities: an actual world in which, say, the agent descends at the east end and has to go right to find his car at the northern edge and a counterfactual possibility in which the agent descends at the west end and has to turn left to find his car.

To analyze self-locating knowledge, Stalnaker (2008, 2014) uses the concept of centered possible worlds proposed by David Lewis (Lewis 1979). A centered possible world is a pair $\langle c, x\rangle$ where x is a possible world which represents an objective possible situation and c is a person with whom the knower or believer identifies herself in x. In Lewis's system, centered possible worlds are more fine-grained objects than uncentered possible worlds, in the sense that a single uncentered possible world can correspond to a number of centered possible worlds. This is what happens when an agent is confused about his self-location, as in the parking garage story above. The distinction between centered and uncentered possible worlds allows Lewis to capture the intuitive difference between objective and self-locating knowledge. The content of self-locating knowledge is a proposition centered on the individual whose knowledge is being represented. By contrast, the content of objective knowledge is a proposition in which centers are irrelevant.[24]

Although Lewis's system (commonly assumed in formal semantics) is successful in capturing the intuitive difference between objective and self-locating knowledge, the system, as it stands, has serious limitations. As Stalnaker (2008) points out, because in Lewis's system the content of objective and self-locating knowledge is represented by different kinds of propositions (i.e., a set of centered worlds where the center is irrelevant versus a set of centered worlds with the believer at the center), the system is solipsistic. It cannot be used to explain how self-locating information is communicated, or how objective and self-locating knowledge of different individuals is integrated, or even how self-locating beliefs of one individual are revised. As an illustration for self-locating knowledge being communicated, suppose that in the parking garage story above the unfortunate agent confused about whether he has to turn left or right to find his car sees a young couple and asks them whether he is at the east or west end of the garage. They reply: "*You are at the west end*". There is nothing self-locating about the couple's reply, but on hearing and accepting their reply, the agent acquires self-locating information and can now conclude that he has to turn left to find his car.

To overcome these limitations, Stalnaker proposes a number of modifications of Lewis's system. Here, I discuss two of his modifications immediately relevant to the topic of this paper. The first modification has to do with how an accessibility relation between possible worlds specifies whose beliefs are being represented. Lewis uses the classical Hintikka-style accessibility relation that specifies whose beliefs are being represented by an index on the relation, e.g., xR_Ay says that y is compatible with Alice's beliefs in x. However, the Hintikka-style accessibility relation is insufficient for capturing intentionality of beliefs, in general, and is especially problematic for self-locating beliefs.[25] For example, if Alice is talking to Bennett, her beliefs about the beliefs of the person she is talking to may be different from her beliefs about what Bennett believes. This may be the case when Alice does not realize that she is talking to Bennett. In this situation, we might assent to '*Alice believes the person she is talking to believes that* p' without assenting to '*Alice believes*

[24]More precisely, Lewis is talking about centered properties of an individual broadly construed, but since there is a one-one correspondence between these 'properties' and possible worlds, it is proper to talk about centered possible worlds (Lewis 1979).. It is also worth noticing that in Lewis's system objective information is a sub-case of self-locating information (a sub-case where the center is irrelevant).

[25](Stalnaker 2014, 120)

Bennett believes that p'. The classical indexed accessibility relation does not allow us to express this difference.

Stalnaker proposes that there is only one accessibility relation established between centered possible worlds. That is, $\langle A, x\rangle R\langle B, y\rangle$ says that A in x locates herself as B in y. Unlike for Hintikka and Lewis, for Stalnaker the believer's specification is in the relata and not on the accessibility relation, see (11).[26]

(11) a. Hintikka/Lewis: y is compatible with what A knows in x iff xR_Ay
b. Stalnaker: y is compatible with what A knows in x iff for some center c, $\langle A, x\rangle R\langle c, y\rangle$

With this modification, we can represent the difference between Alice's beliefs about the beliefs of the person she is talking to and Alice's beliefs about Bennett's beliefs (even if she is talking to Bennett). Suppose f is an individual concept (i.e., a function from a possible world to an individual) which for any possible world picks out the person with whom Alice is talking in that world. Then, beliefs of the person Alice is talking to will be centered on the individual (whoever it might be) who is the value of f, see (12)a. At the same time, Bennett's beliefs (according to Alice) will be beliefs that have Bennett as the center, see (12)b. And these beliefs do not have to be the same.

(12) a. '*Alice believes the person she is talking to believes that* p' is true in α iff
for all worlds x and y and all subjects C,
if $\langle A, \alpha\rangle R\langle A, x\rangle$ and $\langle f(x), x\rangle R\langle C, y\rangle$, then $y \in p$
b. '*Alice believes Bennett believes that* p' is true in α iff
for all worlds x and y,
if $\langle A, \alpha\rangle R\langle A, x\rangle$ and $\langle B, x\rangle R\langle B, y\rangle$, then $y \in p$

The second modification proposed by Stalnaker changes the relation between centered and uncentered possible worlds. As mentioned above, for Lewis, centered possible worlds are fine-grained objects, in the sense that multiple centered possible worlds may correspond to one uncentered possible world. For Stalnaker, however, there is one-one correspondence between centered and uncentered possible worlds, see (13).[27] In other words, in Stalnaker's system, if an individual locates herself differently a possible world (i.e., there are two distinct centered possible worlds), there are also two distinct uncentered possible worlds she locates herself in.[28]

(13) *One-one correspondence principle*
for all $c, c', c'' \in S$ and all $x, y \in W$, if $\langle c, x\rangle R\langle c', y\rangle$ and $\langle c, x\rangle R\langle c'', y\rangle$, then $c' = c''$

To sharpen this difference between Lewis's and Stalnaker's views, let me illustrate it with a vivid analogy using Stalnaker's own words:

> "A misleading picture sometimes accompanies the Lewis account of self-locating belief: belief about what possible world you are in is like belief about what country you are in, while beliefs about where in the worlds you are is like a more specific belief about where in the country you are (what village, street corner, or mountain top). But ordinary belief about where you are in the world is always also belief about what possible world you are in (what possible state of the world is actual). If I am not sure as I drive along

[26](Stalnaker 2014, 119)

[27](See Stalnaker 2014, 118)

[28]This modification reflects the disagreement between Stalnaker and Lewis about the metaphysics of possible worlds discussed earlier.

the highway toward New York, whether I am still in Massachusetts, then I am not sure whether I am in a possible world in which this stretch of highway is located in Massachusetts." (Stalnaker 2008, 51)

The 'misleading picture', Stalnaker argues, is due to the fact that we erroneously accept the *principle of phenomenal indistinguishability*. He formulates the principle as follows: "if a possibility is an epistemic alternative for a knower at a time (that is, it is compatible with his or her knowledge), then it is phenomenally indistinguishable from the actual world to the knower at that time".[29] This principle comes from the empiricist's view that we have direct epistemic access to phenomenal experience or to our "internal world". Stalnaker argues that our knowledge of the internal world is as indirect as our knowledge of the external world; therefore, we should be skeptical about the principle of phenomenal indistinguishability. Here is a version of one of the arguments that he makes against the principle of phenomenal indistinguishability. Suppose Alice is a brilliant medical student who knows everything there is to know about physiological processes of various illnesses. But she was born with a sensory deprivation disorder so severe that she never had any experiences, such as pain, sadness, dizziness, tranquility, uplifting, happiness, or euphoria. Alice agrees to participate in an experiment where she will be given a drug inducing one of two conditions: vertigo (in which case she will experience dizziness and lightheadedness) or hypermania (in which case she will experience upbeat and euphoria). Alice does not know which drug she will be given, but she knows that the choice will depend on a toss of a fair coin: if Heads, she will be given a vertigo-inducing drug, if Tails a hypermania-inducing drug. The argument goes as follows: because Alice has never experienced any sensations, both before and after the experiment she is in a state of ignorance, not able to distinguish between the epistemic possibility where the coin landed Heads and she was given a vertigo-inducing drug and the one where the coin landed Tails and she was given a mania-inducing drug. These two possibilities are depicted in Figure 2 (where Va = 'Alice was given a vertigo-inducing drug', Ma = 'Alice was given a mania-inducing drug', H = Heads, T = Tails). The worlds α and β in Figure 2 are both physically and phenomenally different.

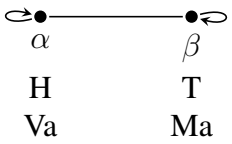

Figure 2: Alice's belief state before and after the experiment

Stalnaker argues that to accept the principle of phenomenal indistinguishability is to suppose that there is a possible world β* where the coin landed Tails and Alice was given a mania-inducing drug, but she had vertigo associated sensations. In other words, β* is physically like β, but phenomenally like α. Those who accept β* as an epistemic possibility, Stalnaker says, accept the existence of phenomenal experiences. I find Stalnaker's argument that β* is not a viable epistemic possibility in the situation above convincing (assuming that the drug worked as expected and Alice did not have an abnormal reaction to it). Thus, I agree that we should abandon the principle of phenomenal indistinguishability and endorse, instead, the position that if an agent is ignorant or confused with respect to her phenomenal (or self-locating) information, the agent is ignorant or confused about which possible world she is in. This position is reflected in the one-one correspondence principle.

[29](Stalnaker 2008, 88).

Let me show how the two modifications discussed above work using a simple example with self-ascription of phenomenal information. Suppose Alice at some point says or thinks to herself:

(14) Alice: *I believe I am having vertigo*.

Our modified (partial) knowledge model is a tuple $\langle W, S, E, R\rangle$ where W and S are, as above, a (non-empty) set of possible worlds and a (non-empty) set of individuals or subjects respectively. E is a set of centered possible worlds (that is, a set of pairs $\langle c, x\rangle$ where $c \in S$ and $x \in W$) that meets the following condition: the individual represented by the center c exists in the possible world x of that center. R is an epistemic accessibility relation (transitive, Euclidean, serial) which differs from the classical Hintikka-style accessibility relation in two respects. First, it is a relation on E (that is, it relates centered possible worlds rather than uncentered possible worlds) and the individual whose beliefs are being represented is specified in the relata rather than as an index (the first modification). Second, R relates centered possible worlds in a way that satisfies the one-one correspondence principle; that is, if a possible world has a different center, we are dealing with a different (uncentered) possible world (the second modification).

In our example in (14), Alice plays three roles: she is the speaker, the attitude holder, and the person who is experiencing vertigo. As a speaker, Alice assumes the basic context for the utterance. Let us say that the basic context C that Alice assumes consists of two worlds α and γ. In this basic context, Alice, the speaker, represents beliefs of the attitude holder (who happens to be Alice herself). This representation of Alice's belief states in C determines the derived context for the embedded proposition *that I am having vertigo*. Let us say that the accessibility relation R in our present case is the same as in Figure 1; that is, the derived context is the set $\{\alpha, \beta, \gamma\}$, see Figure 3. To satisfy the principle of non-triviality, the derived context must include an epistemic alternative in which Alice, the experiencer, is not having vertigo. Since having vertigo is phenomenal (or self-locating) information, the epistemic alternative in which Alice is not having vertigo cannot be a possible world in which Alice locates herself as Alice with the same "experiential profile"[30] as the one she has in the actual world. This is because we abandoned the principle of phenomenal indistinguishability (see above). The epistemic alternative in which Alice is not having vertigo must be a possible world in which Alice locates herself as a different individual, possibly with the same name, shared memories, etc., but with a different "experiential profile" at the relevant time. To simplify things, I will represent Alice with a different "experiential profile" than the one in the actual world as Mabel.

To express Alice's ignorance (or confusion) about her own identity, we can use what Stalnaker calls an I-concept. An I-concept with respect to a possible world x is an individual concept f (a function from a possible world to an individual) whose value for any possible world y accessible from x is the center of y. That is, f is an I-concept with respect to x iff for any world y and any individual B, if $\langle f(x), x\rangle R\langle B, y\rangle$, then $B = f(y)$. An I-concept may be rigid in case the believer is not ignorant or confused about her self-location or phenomenal experience. In this case, the value of f is the same individual in all worlds accessible from x. An I-concept may be non-rigid – this is the case when the believer is ignorant or confused about her self-location or phenomenal experience. In this case, the value of f is different for different worlds accessible from x. In Alice's case that we are discussing here, the I-concept that represents Alice with respect to the actual world α is

[30](E.g., Godfrey-Smith 2020)

non-rigid: it picks Alice in α and β, but picks Mabel in γ (that is, $f(\alpha)$ = A, $f(\beta)$ = A, $f(\gamma)$ = M).[31] The full setup is summarized in Figure 3 (where Va = 'Alice is having vertigo', ¬Vm = 'Mabel is not having vertigo', H = Heads, T = Tails).

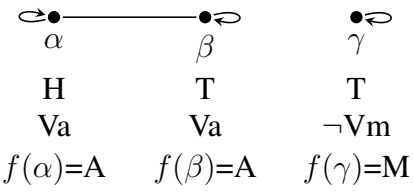

Figure 3: Alice *"I believe I am having vertigo"*

In this setup, the principle of non-triviality is satisfied since the derived context (the union of possible worlds accessible from the basic context set C) is $\{\alpha, \beta, \gamma\}$ and it is compatible with both propositions *that I am having vertigo* (true in α and β where Alice locates herself as Alice) and with *that I am not having vertigo* (true in γ where Alice locates herself as Mabel). That is, $\bigcup_{x \in C}\{y : xR_f y\} \cap \{z : f(z)$ is having vertigo in $z\} \neq \emptyset$ and $\bigcup_{x \in C}\{y : xR_f y\} \cap W - \{z : f(z)$ is having vertigo in $z\} \neq \emptyset$ (where $xR_f y$ iff $\langle f(x), x\rangle R\langle f(y), y\rangle$ for any x and y). Also, the belief ascription as a whole is true in the actual world α since $\{\alpha, \beta\} \subseteq \{x : f(x)$ is having vertigo in $x\}$. Finally, we can say that the role of the belief ascription is to eliminate from the context set C the possible world γ in which Alice believes of herself that she is not having vertigo.

Before concluding this section, let me briefly mention that in order to capture the familiar cases of self-locating knowledge in this framework (the cases with learned amnesiacs, mountain-dwelling gods, etc., including the case where one is lost in the underground parking garage), one will say that the alternative with the different center is accessible from the world of evaluation. For example, in Figure 3, this will amount to saying that γ is accessible from α (i.e., $\langle\langle A, \alpha\rangle, \langle M, \gamma\rangle\rangle \in R$). In this case, Alice cannot truthfully say *I believe that I am having vertigo*.

Let me summarize our discussion so far. I started in section 2 with the explanation of subject obviation proposed by Nicolas Ruwet. The exposition of his view was important because, as I said, my own explanation of subject obviation resembles his in key features. Like Ruwet, I attribute subject obviation to a property of "general intelligence". And also like Ruwet (and a number of scholars after him), I want to say that obviative sentences are unnatural because there is an illegitimately tight "self-to-self" relation between the two coreferential subjects. Then, I proceeded to elaborate on each of these two key features. In section 3, I discussed the principle of non-triviality which, I suggested, is the property of "general intelligence" to which subject obviation can be

[31]This is a simplified picture. Because γ is not accessible from α, we need two I-concepts in C: one, say f_1, for Alice's representation of herself with respect to α and the other one, say f_2, for Alice's representation of herself with respect to γ. Such a setup will be more expressive and will allow us to capture cases in which Alice 'chooses' between being in a knowledgeable and ignorant cognitive state. However, this refinement will complicate the representation of indexical expressions which I represent simply as the value of an I-concept f, for example, representing *I am having vertigo* as $\{x : f(x)$ is having vertigo in $x\}$. We independently need a more elaborate representation of indexical expressions and there is a lot of linguistic work on this topic. But I want to leave this discussion outside of the scope of this paper since this is an independent issue and once suitably settled, can be easily incorporated into the explanation of subject obviation that I propose in this paper.

attributed. The principle of non-triviality disallows ascription of a proposition whose content does not exclude any possibility from a belief state. I showed how this principle can be formulated in the framework developed by Robert Stalnaker. After that, in section 4, I discussed phenomenal (or self-locating) information with a view to saying that this kind of information makes embedded propositions in obviative sentences non-ascribable. Self-locating knowledge was also captured in Stalnaker's system. In the next section, I show how the apparatus presented in the last two sections enables us to explain subject obviation.

5 Obviative, non-obviative, and ameliorated sentences

Time has come to look at obviative, non-obviative, and ameliorated sentences. Above, I discussed self-ascriptions like *I believe I am having vertigo* as uttered by Alice. I said that in order to satisfy the principle of non-triviality, Alice's cognitive state in the context should include an epistemic possibility in which Alice is not having vertigo. Since having vertigo is a phenomenal (or self-locating) information, as was discussed in the previous section, the epistemic possibility in which Alice is not having vertigo is also a possible world in which Alice locates herself as an individual with a different "experiential profile".

But suppose Alice utters the same sentence in a context that forbids her to doubt or be confused about her self-location. In this case, the self-ascription becomes unnatural. Consider the sentences in (15) uttered in a non-pathological, non-dynamic situation that rules out Alice's slightest doubt about how she feel in the world.

(15) [Context: a non-pathological, non-dynamic situation, in which Alice is fully aware of her perceptual or cognitive state and has no doubts about it.]
   a. #I believe I am having vertigo.
   b. #I believe I am sane.
   c. #I believe I am standing on one leg.
   d. #I believe I am jumping around.

The examples in (15) sound unnatural because there is tension between the principle of non-triviality which asks us to suppose that there is a possibility in which Alice is not aware of her cognitive state and the context (or assumed common ground) that suppresses this supposition.

I want to suggest that this kind of tension is what makes sentences with coreferential subjects and self-locating information in the embedded clause obviative. Consider again subject obviation in Hungarian, repeated from (10). The embedded propositions in these examples convey self-locating information, as we discussed earlier.

(16) a. #Remélem, hogy (nem) szédülök. (Hungarian)
   'I hope that I (don't) have vertigo'
   b. #Remélem, hogy fél lábon állok.
   'I hope that I'm standing on one leg'
   c. #Remélem, hogy (nem) fázom.
   'I hope that I'm (not) cold'
   d. #Remélem, hogy ugrándozok.
   'I hope that I'm jumping around'

e. #Remélem, hogy simogatom a macskát.
'I hope that I'm stroking the cat' (Szabolcsi 2021, 10)

Now, suppose that 'hope' in Hungarian is a predicate that induces a "particularly intimate"[32] relation between the attitude holder and the subject in the embedded clause. This "particularly intimate" relation disallows the attitude holder to view herself as distinct individuals in the context in which an attitude ascription with 'hope' is made. Let me call such "particularly intimate" attitude predicates *rigid* attitudes and using the apparatus from the previous section, define them as predicates which admit only a rigid I-concept in the context in which they are used:[33]

(17) An attitude predicate is *rigid* iff
for all $x, y \in C,\ f(x) = f(y)$
where f is an I-concept

My proposal is that obviative sentences are sentences in which there is an attempt to ascribe phenomenal (or self-locating) information using a rigid attitude. Such an attempt cannot be successful because an ascription should satisfy the principle of non-triviality, to satisfy the principle of non-triviality, an ascription of self-locating information should involve an epistemic alternative in which the agent locates herself differently than she does in the actual world (or the world of evaluation), but rigid attitudes forbid varying self-location in the same context.

Let me make a quick illustration, using the English gloss of the Hungarian example in (16)a. Suppose Alice says *#I hope I am having vertigo.* We can describe Alice's epistemic situation needed to satisfy the principle of non-triviality using the setup that we used for (14), see Figure 4 (where Va = Alice is having vertigo, Vm = Mabel is having vertigo, H = Heads, T = Tails). The difference between our present (obviative) sentence and the sentence in (14) is the type of the attitude. 'Hope' is a rigid attitude which disallows a non-rigid I-concept. Therefore, the sentence is irreparably unnatural.

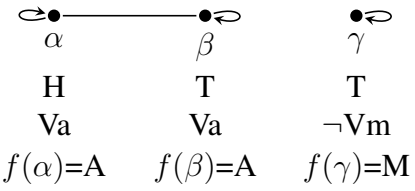

Figure 4: Alice *"#I hope I am having vertigo"*

If the proposal that subject obviation is an interplay between the type of an attitude and the type of ascribed information (within the frame set by "general intelligence") is maintainable, there are two questions that need to be resolved: (a) what makes a particular attitude (in a particular language) rigid and (b) how do we draw a line between phenomenal and impersonal information. In this paper, I will not give a satisfactory answer to either of these questions. But I will try to show that my way of formulating these questions puts us in a better position for answering them compared to recent semantic-pragmatic accounts of subject obviation (sections 6 and 7).

[32]This is an expression used by Ruwet in connection with French *vouloir* 'want' (Ruwet 1991, 18).

[33]In a more elaborate version of this proposal, this is a condition on having identical I-concepts in the same context. See fn. 31.

Let us now look at ameliorated sentences. In ameliorated sentences, as in (18) repeated from above, the embedded proposition does not involve phenomenal (or self-locating) information. So, the principle of non-triviality can be satisfied without introducing a non-rigid I-concept in the context. The absence of the tension between the need to satisfy the principle of non-triviality and the requirement on a rigid attitude makes these sentences (more) acceptable.

(18) a. Remélem, hogy benne vagyok a csapatban. (Hungarian)
'I hope that I'm on the team'
b. Remélem, hogy nem untatlak.
'I hope that I'm not boring you'
c. Remélem, hogy biztonságban vagyok.
'I hope that I'm safe' (Szabolcsi 2021, 10)

Again, as a quick illustration, let us consider the English gloss of the example in (18)a. Suppose Alice says *I hope I am on the team*. In this case, we can represent Alice's epistemic situation as in Figure 5 (where Ma = Alice is on the team, H = Heads, T = Tails). Since being on the team is not self-locating information, adding an epistemic possibility in which Alice is not on the team does not require Alice to locate herself differently than she does in the actual world. Therefore, there is no tension between satisfying the principle of non-triviality and adhering to the requirement of a rigid attitude.

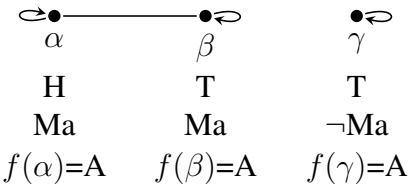

Figure 5: Alice *"I hope I am on the team"*

At this point, let me briefly comment on Ruwet's empirical observation that ameliorated sentences have varying degrees of acceptability and that different factors (and their combinations) may contribute to the amelioration effect. This observation is in line with my proposal in this paper. Amelioration occurs when the embedded clause is not interpreted as conveying phenomenal (or self-locating) information. But in the framework assumed in this paper, the divide between self-locating and impersonal information is not the divide in a type of information. In Stalnaker's (and Lewis's) view self-locating and impersonal information are of the same type. In addition to that, in Stalnaker's system, self-locating beliefs have the same kind of content as impersonal beliefs. What sets self-locating beliefs aside, according to Stalnaker, is the subject's relation to the content of this beliefs and not the content itself. Of course, there are clear cases of phenomenal and clear cases of impersonal information on which we will all agree. For example, being cold versus being on the team. But there are also borderline cases like being safe or being allowed whose classification as phenomenal versus impersonal depends on their interpretation. This dependence on interpretation can account for varying degrees of acceptability of ameliorated sentences.

To complete the picture, let us consider a non-obviative sentence corresponding to (16)a in which the subject of the main predicate and the embedded subject do not refer to the same individual. Suppose Alice says in Hungarian *I hope Mabel is having vertigo*. Then, Alice's epistemic situation

can be represented as in Figure 6 (where Vm = 'Mabel is having vertigo', H = Heads, T = Tails). Of course, Mabel's having vertigo is not self-locating information for Alice; so, the tension that we have in obviative sentences does not arise.

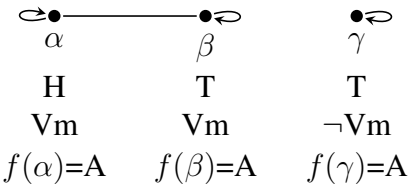

Figure 6: Alice *"I hope Mabel is having vertigo"*

6 Non-triviality, uncertainty, and introspection

As we saw in the previous section, a part of my explanation of subject obviation is an interplay between the meaning of the matrix predicate and the kind of information in the embedded clause. This idea is also present in the explanation in Ruwet (1984) and some recent semantic-pragmatic accounts of subject obviation. I already discussed the differences and similarities between my explanation of subject obviation and Ruwet's (section 2). In this section, I will compare my explanation with two recent semantic-pragmatic accounts.

The first account I look at is a Szabolcsi-Kaufmann account (Szabolcsi 2021; Kaufmann 2019). According to this account, the source of subject obviation is the same as the source of *directive obviation*, which is an inability of first person exclusive forms to be subjects of regular root imperatives or subjunctives used for directives. For example, in Greek, *na*-subjunctives can be used as directive speech acts, see (19)a, but not in the first person singular form, see (19)b. Directive obviation can be described as a ban on having the *director* and *instigator* refer to the same individual. In case of first person singular, to the speaker.

(19) a. Na aniksis to parathiro. (Greek)
SBJV open-2SG the window
'Open the window!'

b. #Avrio na stilo ena e-mail stin Ana.
tomorrow SBJV send-1SG an e-mail at.the Anna
'Tomorrow, I should send an email to Anna.' (Oikonomou 2016, 73, 168)

Kaufmann (2019) builds her account of directive obviation on her earlier work, where directive speech acts are carried out with modalized sentences (i.e., *Open the door!* ≈ *You must/should open the door*) that come with a set of pragmatic presuppositions restricting the context. These restrictions derive the non-assertive character of imperatives and non-canonical directives. In particular, there are two general conditions on the use of imperatives: (a) the director of an imperative presents herself as uncertain about the course of events (Epistemic Uncertainty Condition) and (b) the director believes that the commanded proposition becomes true if the instigator takes it to be necessary (a combination of Decisive Modality and Director's Anticipation). If the director and the instigator are one and the same individual, the two conditions come into conflict. As a result, there can be no context that satisfies both of these conditions and directives like (19)b are judged unacceptable.

Kaufmann (2019) makes a cursory remark that her account of directive obviation can extend to subject obviation in sentences with desire verbs like *vouloir* 'want'. Szabolcsi (2021) accepts this conjecture without giving a detailed analysis of how the account might work. I presume the idea is that the uncertainty presupposition of *want* and its kin[34] comes in conflict with the certainty of the subject of the embedded clause. The latter, I take it, is due to the type of information conveyed in the embedded clause.

The virtue of this account is that it proposes a general explanation of subject obviation in terms of semantic-pragmatic restrictions that can unify different obviative phenomena. The shortcomings of this account will, of course, depend on a concrete implementation of the general schema. However, already at the general level, it is possible to name a number of problematic points. First, if decisive modality remains an important ingredient of the analysis, it is unclear how the account can be extended to cases of subject obviation with epistemic attitudes. We saw such cases in Italian earlier in the paper (see also (20) below). Second, since the uncertainty condition is attributed to the presuppositional meaning of an attitude predicate, factive predicates that show subject obviation cannot be explained by this account. I discuss subject obviation with factives in the next section. Finally, to account for cross-linguistic variation of subject obviation, the account will have to employ lexical (presuppositional) differences between those predicates that show subject obviation and those that do not. If two predicates are lexically similar, it may turn out to be difficult to locate such a difference. This point will also be discussed in the next section.

Let me now look at another pragmatic-semantic account of subject obviation. Costantini (2016, 2023) studies subject obviation with epistemic attitudes in Italian, see (20) repeated from above. His explanation also attributes subject obviation to a conflict between the interpretation of the matrix verb and the information conveyed by the embedded clause. In his view, (20)a is obviative because the verb is "a predicate implying an indirect access to a proposition" and the embedded clause conveys "a proposition accessible through introspection".[35]

(20) a. #Penso che io parta domani. (Italian)
'I think that I leave-SBJV tomorrow'
b. Penso che io abbia fatto molti errori.
'I think that I have-SBJV made many mistakes'

The merit of this account is that it broadens the empirical landscape of the phenomenon. Previous studies of subject obviation were focussing on desire predicates which led to erroneous generalizations. The shortcomings of the analysis are similar to what we saw for the Szabolcsi-Kaufmann analysis. Bypassing the question of concrete implementation in this case as well, let me again list the problematic points. First, it is hard to see how the account can be extended to other cases of subjection obviation. "An indirect access to a proposition" is a plausible restriction on an epistemic attitude, but it is unclear what it means in case of desire predicates even in the same language. Second, factives are also problematic for this account (as noted by Costantini as well). Finally, cross-linguistic variation is also an issue for the same reason as above.

[34](See Heim 1992; von Fintel 1999)
[35](Costantini 2023)

7 Avoiding a wild goose chase

In this section, I will first discuss the problem that factives pose for the semantic-pragmatic accounts of subject obviation from the previous section. I show that my explanation of subject obviation avoids this problem. After that, I will make some remarks about cross-linguistic variation of subject obviation.

Factives

Subject obviation obtains with factive attitudes like 'regret' and 'know'. Consider the Italian examples in (21) and the Hungarian examples in (22). Notice that in Italian, *rammaricarsi* 'regret' selects the subjunctive mood, whereas in Hungarian, *sajnál* 'regret' selects the indicative. This, once again, illustrates the point discussed earlier that subject obviation is not limited to subjunctive clauses. The Hungarian example in (22)b also shows the familiar amelioration effect found in other subject obviation cases.

(21) a. #Mi rammarico che io parta domani. (Italian)
'I regret that I leave-SBJV tomorrow'
b. #Non so se io parta domani.
'I don't know if I leave-SBJV tomorrow' (Costantini 2016: 127)

(22) a. #Sajnálom, hogy ugrándozok. (Hungarian)
'I regret that I jump-IND around'
b. Sajnálom, hogy untatlak.
'I regret that I bore-IND you' (Szabolcsi 2021: 11-12)

Discussing subject obviation with *sajnál* 'regret' in Hungarian, Szabolcsi (2021) notes that it cannot be easily explained in terms of a conflict between a certainty and uncertainty conditions on common ground.[36] This is because the uncertainty condition is not plausible in case of factives, since common ground (and hence, attitude holder's beliefs if common ground is defined as common beliefs of the participants of a conversation) is assumed to entail the prejacent of a factive predicate. As an alternative, Szabolcsi (2021) outlines an explanation with reference to counterfactual reasoning which approximates the certainty-uncertainty conflict. Hungarian 'regret', she argues, can be paraphrased using 'wish', as in (23)a, which makes it similar to English *I find it regratable* or *I wish it weren't the case*. The paraphrase in (23)a entails the propositions in (23)b, which (according to Szabolcsi) together have "an appropriate whiff of contradiction to them".[37]

(23) a. I wish I weren't jumping around (paraphrase of (22)a)
b. If it were up to me whether I am jumping around, I would not be jumping around
AND I am jumping around (per factivity of 'regret'/'I wish I weren't')
AND It is up to me whether I am jumping around (per RESP[38])
(Szabolcsi 2021: 16)

The propositions in (23)b, indeed, seem to form an inconsistent set. But without specifying the interpretation of the counterfactual and the status of the factive and RESP-induced implications, it

[36] Costantini (2016) leaves factives for future research.

[37] (Szabolcsi 2021: 16)

[38] RESP is a responsibility operator proposed by Farkas to account for control (Farkas 1988, 1992) and often used in the literature on subject obviation to describe obviation with intentional or volitional actions.

is hard to say what exactly is wrong when (23)b is used to communicate (22)a. Szabolcsi (2021) does not provide these details. In this section, I will discuss one way of formalizing the reasoning in (23)b. This will help us to see more clearly why (23)b seems inconsistent and give us grounds for evaluating the counterfactual strategy proposed in Szabolcsi (2021). As we will see, counterfactual reasoning is neither necessary nor sufficient for explaining subject obviation with factives. It might be correct that some factives, like *sajnál* 'regret' in Hungarian, have a counterfactual interpretation, but this is not the reason why they show subject obviation. The reason why factives like *sajnál* 'regret' in Hungarian show subject obviation, as I will argue below, is the same as for subject obviation with non-factive attitudes.

Let me begin by pointing out that Szabolcsi's appeal to counterfactual reasoning in case of 'regret' in Hungarian cannot be extended to other factives, for example, 'know' and 'be surprised'. In (21)b, we saw that *sapere* 'know' in Italian shows subject obviation, but it is implausible to engage counterfactual reasoning in case of epistemic factives. One might object that in (21)b the factive implication is not guaranteed because of the negation. It is well documented that in the so-called 'projective' environments (under negation, in questions, with possibility adverbs, etc.), the presence of a factive implication depends on the context or question under discussion.[39] So, one could say that in (21)b no factive implication is projected and subject obviation is due to a certainty-uncertainty conflict as it is the case for non-factive examples. However, subject obviation with *sapere* obtains in positive sentences as well, see (24). Negation is used in (21)b to ensure that *sapere* selects the subjunctive mood in which case the unacceptability is robust. In positive sentences like (24), *sapere* takes the indicative and subject obviation holds only with a neutral intonation; a high pitch on *sapere* makes (24) acceptable (see Costantini 2023, fn. 5 for discussion).

(24) ?So che sto male. (Italian)
'I know I feel sick.' (Costantini 2023: fn. 5)

In addition to *sapere* 'know', other factives in Italian, such as *essere sorpresi* 'be surprised' and *avere saputo* 'come to know', show subject obviation, see (25). These factives are epistemic in nature and, as mentioned above, it is hard to see how the counterfactual strategy in (23) can be used to explain subject obviation with epistemic attitudes. It is true that the constructions in (25) have a dynamic flavour, in the sense that they convey that the attitude holder's beliefs or expectations have been revised to align with the actual situation described by the prejacent. One might try to capitalize on this dynamic component of 'be surprised' and 'come to know' and devise an explanation similar to the counterfactual strategy. But, first, these explanations will not be fully parallel and, second, non-dynamic epistemic factives like 'know' in (21)b and (24) will still be left behind.

(25) a. #Sono sorpreso che io parta domani.[40]
'I'm surprised that I leave-SBJV tomorrow'
b. #Ho saputo che sto male.
'I've come to know that I am-IND feeling sick.'
c. #Ho saputo che sto leggendo il giornale.
'I've come to know that I am-IND reading the newspaper.'
(Costantini 2023: (41), (42))

Cross-linguistic variation of subject obviation is a complex issue which we will discuss in the

[39](E.g., Tonhauser et al. 2013; Simons et al. 2016; Roberts and Simons 2022)
[40]Thanks to Francesco Costantini (p.c.) for this example.

next section. Some of the challenges it poses can be already seen in our discussion of factives. We saw that for some factives, like 'regret' in Hungarian, counterfactual reasoning may be used (at least according to Szabolcsi 2021); for other factives, like epistemic factives in Italian, for which an explanation in terms of counterfactual reasoning is implausible, a different strategy must be found. The point I am trying to make in the above discussion is that if the source of subject obviation is the same across languages, counterfactual reasoning cannot be a necessary part of the explanation for factives.

To show that counterfactual reasoning may not be sufficient for explaining subject obviation even with 'regret' in Hungarian, I will first discuss a way to formalize the counterfactual strategy. This will allow us to see exactly where the conflict lies. Then, I will introduce an example that avoids this type of a conflict but still shows subject obviation. Let me start by repeating the counterfactual strategy in (26).[41]

(26) a. If it were up to me whether I am jumping around, I would not be jumping around
b. I am jumping around (factive implication)
c. It is up to me whether I am jumping around (RESP-induced implication)

We can capture the conditional proposition in (26)a using the account in Stalnaker (1968) et seq.[42] According to this account, the sentence 'if ϕ, then ψ' is true in a world x iff ψ is true in the nearest world to x in which ϕ is true, see (27). The nearest world is determined by a selection function s that maps a proposition and a possible world into a possible world. The selection function (among other restrictions) has to choose the actual world α if the antecedent ϕ is true in α (that is, if $\alpha \in \phi$, then $s(\phi, \alpha) = \alpha$).

(27) a. 'if it were up to me whether I am jumping around, I would not be jumping around' is true in x iff
s(it is up to me whether I am jumping around, x) $\in \{ y : \neg$ I am jumping around in $y \}$

For convenience, let us assign the RESP-induced implication 'it is up to me whether to be jumping around' to ϕ and the factive implication 'I am jumping around' to ψ. We can gloss over the status of factivity and RESP-induced implications and say that for the sentence in (22)a to be assertable in the context of utterance c, the context set C_c (which includes the actual world α) must entail ϕ and ψ. That is, $C_c \subseteq \phi$ and $C_c \subseteq \psi$. Since the actual world α is a member of C_c, both ϕ and ψ are true in α. Since ϕ is also the antecedent in the conditional proposition in (27), the selection function s is required to select the actual world as the nearest world where the consequent is true. Since the consequent is equivalent to $\neg\psi$, we arrive at a contradiction that both ψ (the factive implication) and $\neg\psi$ (the consequent of the conditional) are true in α. This reasoning is schematically shown in (28).

(28) a. $\alpha \in C_c; C_c \subseteq W$ (by the definition of a context set)
b. $C_c \subseteq \phi$ (by RESP-induced implication)
c. $C_c \subseteq \psi$ (by factivity)
d. $\alpha \in \phi$ (from a,b)
e. $s(\phi, \alpha) = \alpha$ (from d and the requirement on s)

[41] Note that the counterfactual strategy does not involve a counterfactual conditional.

[42] An alternative can be to use Lewis (1973) sphere analysis. It will lead to some complications but the main point will not change.

f. $\alpha \in W - \psi$ (from e and the meaning of the conditional)
g. $\alpha \in \psi$ (from a,c)

The counterfactual strategy seems to successfully derive a contradiction that may be the source of subject obviation with factives. However, the contradiction that emerges from the reasoning in (28) is a contradiction for a wrong reason. What the strategy aims at (I think) is a conflict between what the agent intended or wished to do (if it were up to them) and what the agent actually did. For this, we need a link between the RESP-induced implication and the factive implication. However, no such link is required to derive a contradiction in (28). According to (28), sentences like 'If it were up to me whether I am eating cheese, I wouldn't be jumping around' are contradictory for the same reason as our target example in (26)a. Consider (29) in a context in which it is up to the speaker whether they are eating cheese but they decide not to and, instead, they are jumping around. To see how the contradiction is derived all we need is to substitute ϕ in (28) with 'it is up to me whether I am eating cheese'.

(29) a. If it were up to me whether I am eating cheese, I wouldn't be jumping around
b. It is up to me whether I am eating cheese (RESP-induced implication)
c. I am jumping around (factive implication)

Of course, in the actual 'regret'-sentence in (22)a the RESP-induced implication and the factive implication are generated based on the same material, so they are guaranteed to be about the same action. My point here is not that the counterfactual reasoning *plus* the link between the RESP-induced implication and the factive implication cannot explain subject obviation with factives. As we saw, it can. My point is rather that the counterfactual reasoning *by itself* does not produce the explanation we aim for. The fact that the reasoning in (28) derives a contradiction also when there is no link between a RESP-induced implication and a factive implication, as in (29), shows that counterfactual reasoning is not enough for our explanation. We also need to establish a link between a RESP-induced implication and a factive implication. Only then can we obtain a contradiction for a right reason.

In my explanation of subject obviation, there is no connection between the presuppositional meaning of an attitude and obviation. The tension that is the source of the unacceptability of obviative sentences is created by the need to satisfy the principle of non-triviality and the requirement on a rigid attitude. Consider again the obviative example from section 5 in which Alice says *#I hope I am having vertigo.* Alice's epistemic situation in this example can be represented as in Figure 7 repeated from above (where Va = Alice is having vertigo, Vm = Mabel is having vertigo, H = Heads, T = Tails). Nothing in this setup prevents the speaker to presuppose that Alice is having vertigo in every world in the context set. Assuming that factivity can be treated as speaker presupposition (e.g., Stalnaker 2014; Roberts and Simons 2022), the setup depicted in Figure 7 can satisfy factivity and sill give rise to subject obviation.[43]

Cross-linguistic variation

Now let me briefly comment on another problematic point for the semantic-pragmatic accounts discussed in the previous section, namely, cross-linguistic variation of subject obviation. Consider, for example, the verb 'believe'. Subject obviation obtains with 'believe' in Italian, but not in Spanish

[43]To capture factivity in case of non-obviative and ameliorated sentences as in Figure 6 and 5, yet again we need a more expressive system with two I-concepts, see fn. 31.

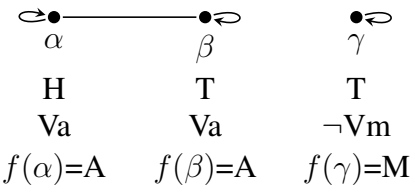

Figure 7: Alice *"#I hope I am having vertigo"*

or French, see (30). Note again that the mood choice is irrelevant here.

(30) a. #Credo che io la convinca. (Italian)
'I believe I convince-SBJV her'
b. Creo que me marcho mañana. (Spanish)
'I believe that me leave-SBJV tomorrow'
c. Je crois que je suis malade. (French)
'I believe that I be-IND ill'

To account for this variation, the semantic-pragmatic accounts have to say that 'believe' in Italian is semantically different from 'believe' in Spanish or French. For example, that 'believe' in Spanish and French do not convey indirect access to knowledge or that they lack an uncertainty presupposition. It remains to be seen whether such conjectures can be supported by evidence. The point that I want to make here is that according to my proposal, we can reformulate the question about the property that requires an attitude predicate to give rise to obviation. Instead of asking which property all such predicates have in common, such that it is responsible for subject obviation (the question to which it will be difficult to find one-size-fit-all answer), we will ask which property (possibly different in each case) disallows an interpretation with a non-rigid I-concept. I believe the latter question is less demanding and will be easier to find an answer to.

8 Conclusion

In this paper, I defended the view that subject obviation is a property of "general intelligence" rather than linguistic competence. I started with the explanation of subject obviation proposed by Nicolas Ruwet who was the first to articulate this view. In my proposal, the role of "general intelligence" was played by the principle of non-triviality that disallows ascription of a proposition whose content does not exclude any possibility from a belief state. I showed how this principle can be formulated in the framework developed by Robert Stalnaker. I also shared Ruwet's intuition (equally present in a number of recent accounts of subject obviation) that obviative sentences are unnatural because there is an illegitimately tight "self-to-self" relation between the two coreferential subjects. I discussed that this illegitimately tight "self-to-self" relation can be captured as phenomenal (or self-locating) information and it is this kind of information that makes embedded propositions in obviative sentences non-ascribable. Self-locating knowledge was also captured in Stalnaker's system. I concluded by discussing how my proposal allows us to see in a fresh light some phenomena that are problematic for grammar-based accounts of subject obviation.

References

Bouchard, Denis. 1983. "The Avoid Pronoun Principle and the Elsewhere Principle." In *Proceedings of NELS 13*, 29–36. Amherst: Univeristy of Massachusetts, GLSA.

Costantini, Francesco. 2006. "Subjunctive Obviation: An Interface Perspective." PhD thesis, Ca' Foscari University of Venice.

———. 2016. "Subject Obviation as a Semantic Failure: A Preliminary Account." *Annali Di Ca' Foscari. Serie Occidentale* 50: 109–31. https://doi.org/10.14277/2499-1562/AnnOc-50-16-5.

———. 2023. "On Some Epistemic Access Effects." In *Agency and Intentions in Language*, edited by Julie Goncharov and Hedde Zeijlstra. Vol. 6. Brill Research Perspectives in Linguistics. Leiden: Brill. https://brill.com/display/title/68330?contents=editorial-content.

Farkas, Donka. 1988. "On Obligatory Control." *Linguistics and Philosophy* 11: 27–58.

———. 1992. "On Obviation." In *Lexical Matters*, edited by Ivan A. Sag and Anna Szabolcsi, 85–110. Stanford, Calif.: Center for the Study of Language; Information.

Godfrey-Smith, Peter. 2020. *Metazoa: Animal Minds and the Birth of Consciousness*. Harper Collins.

Heim, Irene. 1992. "Presupposition Projection and the Semantics of Attitude Verbs." *Journal of Semantics* 9: 183–221.

Higginbotham, James. 2003. "Remembering, Imagining, and the First Person." In *Epistemology of Language*, edited by A. Barber, 496–533. Oxford: Oxford University Press.

Jackson, Frank. 1982. "Epiphenomenal Qualia." *Philosophical Quarterly* 32: 127–36.

Kaufmann, Magdalena. 2019. "Who Controls Who (or What)." In *Proceedings of SALT 29*, edited by Katherine Blake, Forrest Davis, Kaelyn Lamp, and Joseph Rhyne, 636–64. https://doi.org/https://doi.org/10.3765/salt.v29i0.4643.

Kempchinsky, Paula. 1986. "Romance Subjunctive Clauses and Logical Form." PhD thesis, UCLA.

———. 2009. "What Can the Subjunctive Disjoint Reference Effect Tell Us about the Subjunctive?" *Lingua* 119 (12): 1788–1810.

Lewis, David. 1973. "Causation." *Journal of Philosophy* 70 (17): 556–67. https://doi.org/10.2307/2025310.

———. 1979. "Attitudes de Dicto and de Se." *The Philosophical Review* 88 (4): 513–43. https://www.jstor.org/stable/2184843.

Oikonomou, Despina. 2016. "Covert Modals in Root Contexts." PhD thesis, MIT.

Perry, John. 1977. "Frege on Demonstratives." *The Philosophical Review* 86 (4): 474–97. https://www.jstor.org/stable/2184564.

———. 1999. *Knowledge, Possibility and Consciousness*. Cambridge, Mass.: MIT Press.

Picallo, Carmen. 1985. "Opaque Domains." PhD thesis, CUNY.

Roberts, Craige, and Mandy Simons. 2022. "Preconditions, Presuppositions, and Projection." In *Proceedings of the 23rd Amsterdam Colloquium*, edited by Giorgio Sbardolini Marco Degano Tom Roberts and Marieke Schouwstra, 429–36.

Ruwet, Nicolas. 1984. "Je Veux Partir/* Je Veux Que Je Parte. À Propos de La Distribution Des Complétives à Temps Fini Et Des Compléments à l'infinitif En Français." *Cahiers de Grammaire* 7: 74–138.

———. 1991. "*Je Veux Partir/Je Veux Que Je Parte*: On the Distribution of Finite Complements and Infinitival cOmplemetns in French." In *Syntax and Human Experience*, edited by John Goldsmith, 1–55. Chicago: University of Chicago Press.

Schlenker, Philippe. 2005. "The Lazy Frenchman's Approach to the Subjunctive." In *Romance Languages and Linguistic Theory*, 269–309.

Simons, Mandy, David Beaver, Craige Roberts, and Judith Tonhauser. 2016. "The Best Question: Explaining the Projection Behavior of Factives." *Discourse Processes* 54 (3): 187--206. https://doi.org/10.1080/0163853x.2016.1150660.

Stalnaker, Robert C. 1968. "A Theory of Conditionals." In *Studies in Logical Theory (American Philosophical Quarterly Monographs 2)*, edited by Nicholas Rescher, 98–112. Blackwell.

———. 1984. *Inquiry*. Cambridge: MIT Press.

———. 1999. *Context and Content*. Oxford: Oxford University Press.

———. 2008. *Our Knowledge of the Internal World*. Oxford: Oxford University Press.

———. 2014. *Context*. Oxford: Oxford University Press.

Stegovec, Adrian. 2019. "Perspectival Control and Obviation in Directive Clauses." *Natural Language Semantics* 27 (1): 47–94.

Szabolcsi, Anna. 2010. "Infinitives Vs. Subjunctives: What Do We Learn from Obviation and from Exemptions from Obviation?"

———. 2021. "Obviation in Hungarian: What Is Its Scope, and Is It Due to Competition?" *Glossa: A Journal of General Linguistics* 6 (1): 1–28. https://doi.org/10.5334/gjgl.1421.

Tonhauser, Judith, David Beaver, Craige Roberts, and Mandy Simons. 2013. "Toward a Taxonomy of Projective Content." *Language* 89 (1): 66–109. https://doi.org/10.1353/lan.2013.0001.

von Fintel, Kai. 1999. "NPI Licensing, Strawson Entailment, and Context Dependency." *Journal of Semantics* 16 (2): 97–148.