

Testing amelioration of sentences with subject obviation

Authors

Abstract

Subject obviation is a restriction on having coreferential subjects in sentences like **Je veux que je parte* ‘I want that I leave-SBJV’. Since Ruwet (1984), it has been noted that this restriction is not absolute, and sentences like above can be ameliorated by a number of factors or their combination. According to Ruwet, ameliorated sentences have different degrees of acceptability, which provides an argument against an explanation of subject obviation in categorical terms. Given the theoretical importance of Ruwet’s observations, in this study, we want to explore a way of empirically testing the amelioration effect. We tested the amelioration effect in Italian for two factors: (i) the use of passives in the embedded clause, and (ii) the use of a modal in the embedded clause. We used a fully factorial within-within design. We found evidence for the amelioration effect in case of passives, but not in case of modals. Our results show that, in principle, it is possible to empirically test the amelioration effect in a highly controlled setting. The results also suggest that there is empirical support for the amelioration effect in case of passives in Italian.

1 Introduction

Subject obviation is a restriction on having coreferential subjects in attitude reports with finite complements.¹ Consider the French example in (1). In (1)a, *vouloir* ‘want’ is used with a finite (subjunctive) complement, and the subjects of the main and the embedded clause refer to different individuals. Such sentences are fully acceptable. But when the subjects of the main and the embedded clause refer to the same individual, as in (1)b, the sentence becomes unacceptable. To express coreference, an infinitive complement can be used, as in (1)c.

¹The restriction is also known as ‘subjunctive obviation’ or ‘a disjoint reference effect’. ‘Subject obviation’ is a recent label used for a range of constructions of which our restriction may be a member (Kaufmann, 2019; Stegovec, 2019). Whether a unified analysis of all subject obviation cases is possible (or desirable) is an open question. We use ‘subject obviation’ to avoid reference to subjunctives, since it has been argued that the restriction is not limited to subjunctives (Szabolcsi, 2021; Costantini, 2023).

- (1) a. Je veux que tu partes. (French)
 ‘I want that you leave-SBJV’
 b. *Je veux que je parte.
 ‘I want that I leave-SBJV’
 c. Je veux partir.
 ‘I want leave-INF’ (Ruwet, 1991)

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. These sentences have the same structure as the French sentence in (1)b and are usually judged unacceptable. We will call such sentences *obviative*. We will call sentences with non-coreferential subjects, like in (1)a, *non-obviative*.³

- (2) a. *Voglio che io parta domani. (Italian)
 ‘I want that I leave-SBJV tomorrow’
 b. *Penso che io parta domani.
 ‘I think that I leave-SBJV tomorrow’ (Costantini, 2016, 2023)
- (3) a. *Azt akarom, hogy meg-látogassam Marit. (Hungarian)
 ‘I want that I visit-SBJV Mary’
 b. *Remélem, hogy (nem) szédülök.
 ‘I hope that I (don’t) have-IND vertigo’ (Szabolcsi, 2021)

Although linguists have been interested in subject obviation since early 1980s, at present, there is no consensus about the source of the restriction. It has been attributed to different domains of grammar (syntax, semantics, pragmatics) and even to properties of “general intelligence”.⁴ Despite the lack of agreement about the source of subject obviation, many things about the restriction have been discovered. First, we now know that subject obviation is not

²E.g. Bouchard 1982; Ruwet 1984, 1991; Picallo 1985; Raposo 1986; Kempchinsky 1987, 2009; Farkas 1992; Schlenker 2005; Costantini 2009, 2016, 2023; Feldhausen 2010; Szabolcsi 2021.

³In the literature, different marks are used to signal unacceptability of obviative sentences (*, #, (?)), etc.), often hinging on the theoretical standpoint of an author. In this paper, we stay neutral to the theoretical analysis of subject obviation and therefore, use grammaticality marks only descriptively. For us, an asterisk (*) signals that a sentence is unacceptable (syntactically ill-formed, semantically deviant, or pragmatically improper). A question mark (?) signals that a sentence has been reported as being acceptable to some degree (either out of the blue or with a particular interpretation in a particular context). Fully acceptable sentences are unmarked.

⁴The analysis in Kempchinsky 1987 is a good representative of a syntactic solution. Farkas (1992) gives an influential semantic account. Kaufmann 2019 (and following her Szabolcsi 2021) is a recent analysis in terms of conflicting pragmatic requirements. Ruwet (1984) has an interesting analysis in terms of “general intelligence”, where “general intelligence” is linked to iconic relationships between sentence structures (finite vs. infinitive) and their interpretations (disjoint reference vs. coreference).

directly related to subjunctive argument clauses, as was initially thought (hence, the old name ‘subjunctive obviation’; see fn. 1). This is because subject obviation is also found with indicative argument clauses: the Hungarian example in (3)b shows that subject obviation is not limited to constructions with subjunctive, since *remélni* ‘hope’ selects the indicative mood.⁵ Second, we know that subject obviation is not directly related to bouletic modality (desires, hopes), as was initially thought; this is because subject obviation is also found with epistemic attitudes. For example, in Italian, subject obviation obtains with *volere* ‘want’, as well as with *pensare* ‘think’, as shown in (2) (Costantini, 2016, 2023). Third, we know that subject obviation is not due to a competition between infinitival and finite complements, as early analyses suggested. This is because in some languages, subject obviation obtains when there is no infinitival alternative to compete with a finite obviative construction. This point is convincingly argued by Szabolcsi (2021) using data from Hungarian where subject obviation is found with predicates like *akarni* ‘want’, which can take an infinitival complement, and with predicates like *remélni* ‘hope’, which cannot, as shown in (3).

Finally, it has been noted that subject obviation is not an absolute restriction – sentences with subject obviation can be improved by a number of factors or their combination.⁶ Ruwet (1984) was the first to discuss the importance of this *amelioration effect* for explaining subject obviation.⁷ Some of his examples of the amelioration effect in French are given in (4) (keeping his own grammaticality judgements). The sentences in (4), like the obviative sentence in (1)b, have coreferential subjects, but they are claimed to be fully acceptable or perceivably more natural than (1)b. The ameliorating factor in (4)a is the use of a passive construction in the embedded clause; the ameliorating factor in (4)b is the use of a modal in the embedded clause. The examples in (5) and (6) show the amelioration effect in Italian and Hungarian (with various ameliorating factors: the use of a passive and a modal in Italian, and the use of an accidental action and a stative predicate in Hungarian).

⁵Earlier, Raposo (1986) suggested that the opposition between subjunctive and indicative cannot be the main source of subject obviation.

⁶In the literature, amelioration of sentences with subject obviation is sometimes called *obviation weakening*. In this paper, we avoid this label because it might seem to come with an implicit theoretical assumption that unacceptability of obviative sentences is due to a restriction that can (or cannot) be weakened. We use the terms *amelioration* and *improvement* as referring to acceptability of sentences. Our interpretation of the issue is compatible with an account where subject obviation is due to a cluster of conflicting or costly demands on competence or processing.

⁷The paper was first published as Ruwet 1984 in French; it was later translated into English by John Goldsmith and published as a chapter in Ruwet 1991. We will refer to this work by its first date of publication, but use the English version for examples and citations. We will depart from the English version only in one respect: Goldsmith chooses to translate sentences with subjunctive complements using an infinitive with a *for*-phrase; for example, he translates (1)b as ‘I want for me to leave’. We keep our translations closer to the structure of original sentences in French.

- (4) a. ?Je veux que je sois enterré dans mon village natal. (French)
 ‘I want that I be-SBJV buried in the village of my birth’
 b. ?Je veux que je puisse partir dès demain.
 ‘I want that I be.able-SBJV to leave by tomorrow’ (Ruwet, 1991)
- (5) a. Penso che io sia stato autorizzato a partire. (Italian)
 ‘I think that I be-SBJV authorized to leave’
 b. Penso che io possa aver fatto molti errori.
 ‘I think that I may-SBJV have made many mistakes’ (Costantini, 2016)
- (6) a. Nem akarom, hogy (véletlenül) az egészséges lábat amputáljam. (Hungarian)
 ‘I don’t want that I (accidentally) amputate-SBJV the healthy leg’
 b. Remélem, hogy benne vagyok a csapatban.
 ‘I hope that I be-IND on the team’ (Szabolcsi, 2021)

The importance of the amelioration effect is that it supplies an argument against a simple explanation of subject obviation in terms of a categorical restriction. This is because ameliorated sentences come with varying degrees of naturalness, which would be unexpected if subject obviation is due to a categorical restriction. Initially, Ruwet (1984) discusses the amelioration effect as an argument against the explanation of subject obviation as a syntactic restriction. For him, since syntactic ill-formedness cannot be amended by degrees and ameliorated sentences come with different degrees of acceptability, sentences with subject obviation have to be syntactically well-formed and the source of subject obviation should be sought elsewhere. This argument is not limited to a syntactic explanation of subject obviation. The same argument applies to an explanation of subject obviation as a categorical semantic or pragmatic restriction.

Given the importance of the amelioration effect for explaining subject obviation, we want to investigate a way of empirically testing the effect. The difficulty is that subject obviation is a complex phenomenon and, at present, we have only rough understanding of the factors that improve sentences with subject obviation. The goal of this paper is to show how, despite the difficulties, the amelioration effect can be tested using a fully factorial design and a differences between differences measure.

In our study, we test the amelioration effect in Italian sentences with *sperare* ‘hope’. The two ameliorating factors in our study are (i) the use of a passive and (ii) the use of a modal in the embedded clause. Two factors are used in order to address two additional questions on top of the main question whether passives or modals improve acceptability of sentences with coreferential subjects. The first additional question has to do with what we will call a strong and a weak hypothesis. The strong hypothesis says that both factors lead to an improvement of the same construction (modulo the factor itself) in the same context. The weak hypothesis says that either of the factors (but not necessarily both) do. The second additional question is whether amelioration with passives and modals are to the same degree.

As we will see, we found evidence for the amelioration effect in case of passives, but not in case of modals. Our results suggest that there is empirical support for the amelioration effect in tested sentences with subject obviation in Italian, at least in case of passives. Because modals did not show reliable results, we conclude that our experiment did not show evidence for the strong hypothesis above. Also, for the same reason, the second additional question cannot be addressed. At a more general level, our results show that, in principle, it is possible to empirically test the amelioration effect in highly controlled settings.

The paper is structured as follows: In section 2, we discuss the amelioration effect, our current theoretical understanding of it, and previous empirical studies. Section 3 presents our experiment and its results. Section 4 contains a discussion of the results of our experiment. Section 5 concludes.

2 Subject obviation and the amelioration effect

2.1 Complexity of the phenomenon

Ruwet (1984) was the first to recognize the importance of the amelioration effect for explaining subject obviation. He says that “the degree of acceptability of any particular sentence [with subject obviation – Authors] depends on a variety of factors”; therefore, subject obviation cannot be thought of “in all-or-nothing terms”.⁸ The goal of this section is to describe the complexity of subject obviation and the amelioration effect as it is exposed in Ruwet 1984 for subject obviation in French.

The factors that Ruwet discusses can be roughly divided into three groups: (i) factors that affect the interpretation of the matrix clause, (ii) factors that affect the interpretation of the embedded sentence, and (iii) factors that simultaneously affect the interpretation of the matrix and the embedded clause.⁹ The most obvious element in the first group is the meaning of a matrix predicate itself. In French, predicates like *croire* ‘believe’, *penser* ‘think’, *dire* ‘say’, and *pretendre* ‘pretend (claim)’ allow coreferential subjects with finite complements, as in (7). But predicates like *vouloir* ‘want’, *souhaiter* ‘wish’, *attendre* ‘wait’, and *mériter* ‘deserve’ do not, as in (8). Although the examples in (7) and (8) might suggest that the epistemic/bouletic or the indicative/subjunctive divide plays a role in explaining subject obviation, this is not so. We saw in the introduction that there are languages in which subject obviation obtains with both attitudes and with attitudes that select either the subjunctive or the indicative. So, if we want to have a unified analysis of subject obviation, we need to look for the explanation elsewhere.

⁸(Ruwet, 1991, 48).

⁹Ruwet (1984) himself does not divide the factors into groups, but this grouping will be useful for our discussion.

- (7) a. Je crois que je suis malade. (French)
 ‘I believe that I am-IND sick’
 b. Elle prétendait qu’elle m’aimait.
 ‘She claimed/pretended that she loved-IND me’ (Ruwet, 1991, 49)
- (8) a. *J’attends que je meure. (French)
 ‘I am waiting that I die-SBJV.’
 b. *Tu mérites que tu aies une retraite dorée.
 ‘You deserve that you have-SBJV a golden retirement.’ (Ruwet, 1991, 40- 1)

The meaning of a matrix predicate is an important but not a decisive factor. With predicates, like *vouloir* and others, that give rise to subject obviation, it is nevertheless possible to have an acceptable sentence (with a finite complement and coreference). For example, one can use the conditional tense on a matrix verb in combination with other factors (to be discussed below) affecting the embedded clause to have a fully acceptable sentence, as in (9). The importance of (9) is that the factors can have a cumulative effect.

- (9) Je voudrais bien que je puisse amuser ces enfants. (French)
 ‘I certainly would like that I be.able-SBJV to amuse those children’ (Ruwet, 1991, 27)

The second group of factors that affect the degree of acceptability of obviative sentences are factors that determine the interpretation of the embedded clause. Among these factors Ruwet (1984) reports the use of (a) passives, as in (10)a, (b) the modal *pouvoir* ‘be able’, as in (10)b, (c) perfective aspect, as in (10)c, (d) stative predicates, as in (10)d, and (e) psych-verbs, as in (10)e.

- (10) a. ?Je veux que je sois autorisé à partir demain. (French)
 ‘I want that I be-SBJV authorize to leave tomorrow’
 b. ?Je veux que je puisse partir dès demain.
 ‘I want that I be.able-SBJV to leave by tomorrow’
 c. Je veux (absolument) que je sois parti dans dix minutes.
 ‘I want (absolutely) that I be-SBJV gone in ten minutes’
 d. Je veux que je sois en forme pour le match de demain.
 ‘I want that I be-SBJV in form for tomorrow’s match’
 e. ?Je veux que j’amuse ces enfants.
 ‘I want that I amuse-SBJV those children’ (Ruwet, 1991, 20-27)

In (10), we kept Ruwet’s own judgements, for whom ‘?’ indicates that a sentence does not sound fully natural but is acceptable. The examples in (10) show yet again that ameliorated sentences have varying degrees of acceptability.

In addition to these two groups, there are factors that affect the interpretation of the matrix and the embedded clause simultaneously. One such complex factor is the use of negation in the matrix clause and the use of an action that conveys “something that involves a more or less negative expression as far as the subject is concerned”, as in (11).¹⁰

- (11) a. ?Je ne veux pas que je rate une occasion pareille. (French)
 ‘I do not want that I miss-SBJV a chance like that’
 b. Je ne voudrais pas que j’oublie de donner ce coup de téléphone.
 ‘I would not want that I forget-SBJV to give a phone call’ (Ruwet, 1991, 29)

Another factor is implicit contrast, which can be achieved by the use of a pseudo-cleft, as in (12), or another mechanism (coordination, emphatic adverbs, like *absolument* (‘absolutely’), etc.). The use of a pseudo-cleft by itself does not immediately amend an unacceptable obviative sentence, as shown in (12)a. But together with a facilitating factor from the second group, such as the use of a modal, it can result in an acceptable sentence, as shown in (12)b.

- (12) a. *Ce que je veux, c’est que je parte. (French)
 ‘What I want is that I leave-SBJV’
 b. Ce que je veux, c’est que je puisse partir dès demain.
 ‘What I want is that I be.able-SBJV to leave by tomorrow’ (Ruwet, 1991, 24)

Yet another factor that affects both the matrix and the embedded clause is the use of special idiomatic constructions, such as *Pourquoi veux-tu que S?* ‘Why do you want that S?’ where S echoes a part of the interlocutor’s statement and the question means something like ‘Why in heaven’s name do you take it that S?’; see (13).

- (13) A: C’est malheureux, il va falloir que je parte. (French)
 ‘It’s too bad, I am going to have to leave’
 B: Pourquoi veux-tu que tu partes?
 ‘Why do you want that you leave-SBJV?’
idiom ‘But why should you leave?’ (Ruwet, 1991, 31)

Finally, setting up a context correctly can be enough for improving an obviative sentence. Ruwet gives the example in (14), which is similar to the paradigmatic obviative sentence in (1)b that he starts with. But, unlike (1)b, the sentence in (14) is acceptable because it is interpreted in an appropriate context (in fact, as Ruwet reports, this sentence was actually uttered by a French linguist in the given context).

- (14) *Context:* The speaker stays for a night at his friend’s house. He has to leave early the

¹⁰(Ruwet, 1991, 29)

following morning and does not want to wake his host or any member of the household. Before retiring to bed, he says:
Je veux que, demain, je parte si tôt que personne ne m’entende.
‘I want that tomorrow I leave-SBJV so early that no one hears me’ (Ruwet, 1991, 33)

Without being immersed in the theoretical discussion about the source of subject obviation, it is hard to see how the different factors above fit together. Unfortunately, being immersed in that discussion clarifies the picture only moderately. As we are about to see, our vague and fragmentary theoretical understanding of subject obviation is among the main difficulties for testing the amelioration effect.

2.2 Present understanding of the phenomenon

The intuition about what makes obviative sentences unnatural seems to be straightforward: given that the unnaturalness disappears when the two subjects have disjoint reference (even as minimal as a non-*de se* interpretation), the source of subject obviation must be an illegitimately tight relation between the subject of the matrix clause and the subject of the embedded clause. Why, then, do the facilitating factors help? Simple; they loosen the relation between the two subjects, so that, in effect, the subjects behave as if they have disjoint reference in some sense.

The difficulty is to express this intuition in theoretical terms. Ruwet (1984) tries to elucidate the intuition, but he does not make it expressible in a linguistic framework. For him, *vouloir* ‘want’ “creates an intimate link [...] between the subject and the object of the wanting”, and this “special affinity between the will and action, in particular one’s own action” is the source of subject obviation.¹¹ Other predicates, like *croire* ‘believe’ and *prétendre* ‘claim/pretend’ in (7), do not convey this “special affinity”, and therefore, they do not give rise to subject obviation. The facilitating factors help because they “introduce a distance between the expression of the will and the accomplishment of the act”.¹²

“Intimate link”, “special affinity”, and “distancing” are good locutions to capture an intuition, but they are not very adequate as theoretical terms. For example, which linguistic feature makes *croire* ‘believe’ in French less “intimate” than *pensare* ‘think’ in Italian, so that it explains the difference between (7)a and (2)b? Or what is the shared linguistic mechanism that ensures that the use of passive in (10)a and the context in (14) introduce an appropriate amount of “distancing” for the amelioration effect?

There have been attempts to make Ruwet’s intuition more precise, but all more or less formal solutions so far are partial. They are partial in two ways: First, they are stated in “all-or-nothing terms”, which, as we saw above, goes against the observation that ameliorated

¹¹(Ruwet, 1991, 18)

¹²(Ruwet, 1991, 21)

sentences have varying degrees of acceptability. And, second, they cover only some obviative constructions in some languages with only some facilitating factors.

Among the recent accounts of subject obviation that (explicitly or implicitly) follow Ruwet’s intuition, it is worth mentioning the accounts in Schlenker (2005); Szabolcsi (2021) and Costantini 2016, 2023. For Schlenker (2005), obviative sentences in French have an event *de se* interpretation (in addition to a simple *de se* interpretation), which makes them lose in a competition to similar sentences with infinitival complements; thus, unacceptability. Szabolcsi (2021) argues (based on Hungarian data) that embedded clauses of obviative sentences involve “classical predicates of direct experience”¹³ or a responsibility operator (borrowed from Farkas 1988). When such embedded clauses combine with desire attitudes, they put contradictory demands on the common ground, which explains the unacceptability of obviative sentences. Costantini (2016, 2023) looks at sentences with epistemic attitudes in Italian and observes that they are obviative when the embedded proposition “is accessible through introspection”, where introspection is described as something that gives the believer “a direct access to mental states and is highly epistemically secure”.¹⁴ For him, the unacceptability arises because introspection-based embedded clauses are in conflict with the interpretation of epistemic attitudes (in Italian) as conveying that “the information expressed in the embedded clause either [is] not reliable [...] or [is] indirect or inferred”.¹⁵

The goal of this brief excursion into recent accounts of subject obviation is not to compare or critically assess them. The goal is to show that our current theoretical understanding of subject obviation is vague and fragmentary. However, as we will show, it is possible to design an informative experiment testing subject obviation and amelioration even though we might currently not have a tight grip on the theoretical notions and their interactions.

2.3 On a previous experimental study of the amelioration effect

Given our rough understanding of the phenomenon, it does not come as a surprise that there has, to the best of our knowledge, only been one attempt to put the amelioration effect to test in an experimental setting. Feldhausen and Buchczyk 2021 designed an acceptability judgment experiment in which they used many of Ruwet’s original examples for amelioration as experimental stimuli, plus some constructed items, and a set of ungrammatical and grammatical fillers. The objective of the study was to, in a sense, “replicate” Ruwet’s (1991) result; that is, to test whether the factors assumed by Ruwet to have an ameliorating effect on subject obviation would indeed show it in a controlled experiment. Their study was exploratory in nature, as it was unclear to what extent such effects could be confirmed experimentally. Since previous

¹³(Szabolcsi, 2021, 10)

¹⁴Costantini (2023)

¹⁵Costantini (2023)

approaches to the phenomenon had primarily been based on introspective data, the contribution of Feldhausen and Buchczyk (2021) is that they subjected data that were previously tested only informally by the respective author(s) to a controlled experimental study with multiple naive participants. The levels of the AMELIORATION factor came from six different structures: passive, periphrastic past, negation, modals, psych verbs, and coordination. The resulting design was a *within-participants, between-items* realization of the 6-level AMELIORATION factor, plus a fully *within-participants*, but partly *within-items* and partly *between-items* realization of the control filler conditions. This means that repeated measurements were obtained by repeatedly testing the same item in one experimental condition on multiple participants, but not with multiple lexical realizations. The reasoning was that an amelioration effect for a given factor level is present if the mean of the respective ratings are higher than those of the ungrammatical fillers, and (possibly, but not necessarily) equal to those of the grammatical ones. The experiment was run online using LimeSurvey; sentences were presented in isolation in a pseudo-randomized order; 88 native speakers of French rated these on a 7-point scale, and the ratings were submitted to a one-way by-participants ANOVA and pairwise comparisons by means of a Tukey HSD test. The authors conclude from their results that only one of the six structures hypothesized to show an amelioration effect, coordination, does indeed show that effect in the setting tested by the authors.¹⁶

While their study constitutes an important first step towards a quantitative assessment of the amelioration effect in a controlled experimental setting, there are a number of points concerning Feldhausen and Buchczyk 2021’s experimental design decisions that our follow-up study tries to improve on. These concern (a) the assumption of an absolute acceptability scale; (b) the lack of within-items testing; and (c) an anti-conservative statistical procedure.

The present study builds upon the work of Feldhausen and Buchczyk (2021) but addresses the critical points mentioned above head-on. While the previous study aimed to replicate Ruwet’s data, the current study has the goal to identify the optimal methodological approach for detecting the amelioration effect. To this end, we will implement three aspects in a way that differs from that of Feldhausen and Buchczyk (2021).

First, concerning (a), in order to establish an amelioration effect for a given structure (say, passives), we think that it is necessary to show that a given set of passive sentences with coreference between matrix and embedded subject is judged as more acceptable than the set of its corresponding active variants. Whether the active variants are themselves judged as ungrammatical, or mildly ungrammatical, or even higher up on the scale, does not matter at all, as long as the mean for the corresponding passive sentences is reliably different from that of the active ones. In other words, the amelioration effect makes a prediction about a *relative* difference between the ameliorated and the unameliorated structure; but it does not make a

¹⁶In Buchczyk and Feldhausen (2020), the authors suggested how the results of Feldhausen & Buchczyk (2021) could be modelled in a syntactic way.

prediction about where on a given scale this relative difference is located. While there have been attempts to establish an absolute scale for linguistic acceptability (cf. Featherston (2009); Gerbrich et al. (2019)), we believe that such absolute scalings are not too suitable when the objective of the experiment is to establish a relative effect. A related issue that is inherent to this kind of scaling procedure is that it leaves open how exactly part of the alternative hypothesis is to be worded: while it is clear that the judgements for the ameliorated sentences should be above those for the ungrammatical fillers, it is not clear whether the means for the ameliorated sentences should be at the same acceptability level as the grammatical fillers (necessitating an alternative hypothesis that states an identity), or lower (but to what degree?), or even higher. Having to decide upon questions like this while formulating the prediction for the amelioration effect is a direct consequence which comes with the assumption of an absolute scale, and complicates the interpretation of the results. Moreover, as we will argue below, this assumption turns out not to be necessary at all to predict, and test, an amelioration effect.

As for (b), it is generally desirable to exert as much control over variables in an experiment as possible. That is what underlies the “Max-Min-Con” principle (Kerlinger, 1975), which states that systematic (theoretically significant) variance should be maximized, while unsystematic variance should, if possible, be minimized, or controlled. Following this principle, we employed a within-items design in the current experiment. This means that each of the four conditions of our two 2×2 subdesigns, we tested six items (lexical variants) per participant; these six realizations were controlled for the syntactic properties of their target sentences (see below for the details). While this meant that we had to deviate from Ruwet’s original examples, the fully crossed within-participants and within-items design allowed us a more reliable estimation of the unsystematic variance in our statistical analysis of the judgment data, and thus a more conservative test of the potential amelioration effect.

Finally, concerning (c), it has recently been argued rather convincingly (cf. Liddell and Kruschke 2018) that submitting data measured at an ordinal level, as the ones yielded by the 7-point judgement scale employed by Feldhausen and Buchczyk 2021, to statistical procedures like ANOVAs and/or *t*-tests can be detrimental to the generalizability of the results. Both procedures employed in Feldhausen and Buchczyk 2021 tend to give anti-conservative results with ordinal data; simplifying matters somewhat, this is because the property of non-equidistance of the ranks of the scale is not respected by these procedures. Note that this may arguably not only produce type I errors (i.e., false positives), but also type II errors (i.e., failure to detect an effect). We thus cannot exclude the possibility that one, or even both types of errors have affected the results of the previous study.

In our current study, we deal with all three methodological issues laid out above by running fully crossed design with repeated measures for participants and items and subjecting the resulting data to mixed model ordinal regressions with intercepts and slopes for both random factors (cf. Barr et al. 2013). It should be noted that this limits us to the two ameliorating

factors whose structures lend themselves to a within-items design. Also, given the problems with the explanation of the amelioration effect in theoretical terms, which we discussed in the previous section, we thought it advisable to start with a highly controlled setting testing only a small number of factors and their possible interactions, and to embed these into small context scenarios to further constrain the interpretation of the critical sentences.

3 Experimental evidence

3.1 Design and predictions

To answer the question whether passives and/or modals ameliorate the subject obviation effect in Italian, we came up with an experimental design that allowed us to quantify this improvement of acceptability in terms of *differences between differences* (see Sprouse et al. 2013). In order to be able to compare sentences with coreference between the matrix and an embedded 1st person subject with sentence without coreference (with a 1st person matrix and a 2nd person embedded subject), we created contexts in which both versions of these sentences would be equally plausible; (15) gives an example of such a context, plus the two versions of the sentence in the baseline condition (i.e., in the active voice and without a modal auxiliary):

- (15) *Context:* La compagnia teatrale di cui fanno parte Fabio, Chiara e Marco, oggi sta facendo pratica di trucco scenico. Si dovranno formare gruppi di due persone: una persona truccerà l'altra. Poco prima che il capogruppo decida chi truccerà chi, Chiara dice a Fabio: [Rough translation: *The theater group, in which Pietro, Carla and Luigi are participating, is practicing makeup today, for which groups of two are to be formed. Just before the group leader decides who gets to put makeup on whom, Carla says to Pietro.*]
- a. “Spero che io trucchi Marco.”
 Hope-1SG that I put-make-up-on Marco.
 “I hope that I put make up on Marco.”
- b. “Spero che tu trucchi Marco.”
 Hope-1SG that you put-make-up-on Marco.
 “I hope that you put make up on Marco.”

We can call the difference in acceptability between (15-a) and (15-b) the *basic obviation effect*. Most speakers of Italian will probably judge (15-a) as less acceptable than (15-b), due to the lack of coreference in the latter case. Let's call this difference in acceptability DIFF.ACT, for “difference in the active condition”. Given our research question, we can use DIFF.ACT as a baseline against which we compare the two other conditions: the passive (cf. (16)) and the

modal condition (cf. (17)), respectively.

- (16) a. “Spero che io sia truccato da Marco.”
Hope-1SG that I be-SBJV put-make-up-on-PTCP by Marco.
“I hope that I will be made up by Marco.”
b. “Spero che tu sia truccato da
Hope-1SG that you be-SBJV put-make-up-on-PTCP by
Marco.” (–COREF/PASSIVE)
Marco.
“I hope that you will be made up by Marco.”
- (17) a. “Spero che io possa truccare Marco.”
Hope-1SG that I can-SBJV put-make-up-on Marco.
“I hope that I can make up Marco.”
b. “Spero che io possa truccare Marco.”
Hope-1SG that you can-SBJV put-make-up-on Marco.
“I hope that I can make up Marco.”

If obviative sentences can indeed be improved by the passive and the modal, the difference between the (a) and (b) variants in (16) (DIFF.PASS), and in (17) (DIFF.MOD) should be smaller than the one in the baseline condition. In other words, we can make the following *strong* prediction with respect to the differences in acceptability:

(H₁) DIFF.ACT > DIFF.PASS and DIFF.ACT > DIFF.MOD

Of course, the corresponding null hypothesis is that the differences are equal. We call this a strong prediction, because it assumes that both modals and passives are able to ameliorate the basic obviation effect. It is, however, not entirely clear that this assumption should hold. It might be that modals exert a stronger amelioration effect, while passives do not show it at all; or the other way around. Given the null effect between these two conditions in the Feldhausen and Buchczyk (2021) data, it is hard to come up with a clear hypothesis about the *relative* amelioration effect of these two constructions. Thus, we might phrase a somewhat *weaker* prediction as follows:

(H'₁) DIFF.ACT > DIFF.PASS or DIFF.ACT > DIFF.MOD

Given the disjunctive form of H'₁, either of the construction might show an ameliorating effect, or both of them.

To test this prediction, we employed a 2×3 design, fully crossing the 3-level factor VERB FORM (active vs. passive vs. modal) with the 2-level factor COREF (coreferential vs. non-coreferential subjects). We thus predict an interaction between these factors carried by the two

larger differences for the COREF factor in the passive and modal conditions, as compared to the active condition. Statistically speaking, we predict two 2×2 interactions between the factors VERB FORM and COREF: one for the comparison *active vs. passive*, and one for the comparison *active vs. modal*.

3.2 Method

We performed an acceptability rating study in Italian in which we asked participants to rate the acceptability of a target sentence relative to a context, as exemplified in example (15) above. The preparation of the experimental materials (creation of lists, randomization, etc.), as well as the statistical analysis were done in R (cite R project).

Participants. We tested $n = 48$ participants, all of them native speakers of Italian and undergraduate students of linguistics at the University of Udine. After checking the benchmarking items (on which see below), eight participants had to be replaced due to apparent problems with performing the acceptability judgment task. The $n = 48$ participants that entered the analysis had a mean age of 22.8 years ($sd=2.81$); 33 of the participants identified as female, and 15 as male. They took part in the experiment as part of a course assignment.

Materials. We constructed 24 items, each consisting of a context that introduced a scenario in which the content of the target sentence was roughly equally plausible irrespective of the condition tested. This meant that each of the protagonists mentioned in the scenario could figure in the target sentence as the AGENT or the THEME of the embedded verb with roughly the same plausibility. Below, we give a further example of an actual experimental item in all six conditions.

(18) *Context:* Kim, Sascha e Alex studiano fotografia. In uno dei seminari, gli studenti faranno esercizio sulla fotografia di ritratto. Poco prima che il docente assegni i gruppi da due per l’esercizio, Kim dice a Sascha: [Rough translation: *Kim, Sascha and Alex are studying photography. In one of the seminars, the students are required to practice portrait photography. Just before the lecturer assigns the groups of two for the exercise, Kim says to Sascha:*]

- a. “Spero che io fotografi Kim.” \approx *I hope that I will take a picture of Kim.*
- b. “Spero che tu fotografi a Kim.” \approx *I hope that you will take a picture of Kim.*
- c. “Spero che io sia fotografato da Kim.” \approx *I hope that I will have my picture taken by Kim.*
- d. “Spero che tu sia fotografato da Kim.” \approx *I hope that you will have your picture taken by Kim.*
- e. “Spero che io possa fotografare Kim.” \approx *I hope that I can take a picture of Kim.*

- f. “Spero che tu possa fotografare Kim.” \approx *I hope that you can take a picture of Kim.*

Each item was followed by 7-point rating scale of the type exemplified below:

(completamente inaccettabile)	1	2	3	4	5	6	7	(completamente accettabile)
-------------------------------	---	---	---	---	---	---	---	-----------------------------

In addition to the 24 experimental items, there were three sets of filler items ($n = 48$) testing for different phenomena of Italian syntax, plus a set of $n = 12$ benchmarking items which contained a completely grammatical version of the construction tested in our items, as well as a completely ungrammatical one; here is an example:

- (19) *Context:* Daniele, Luisa e Gianni sono iscritti ad una società di calcio. Stanno aspettando di essere assegnati dal loro allenatore a una delle due squadre che schierano ex professionisti. Daniele dice a Luisa: [Rough translation: *Daniele, Luisa and Gianni have joined in a football club. They are waiting to be assigned by the coach to one of two teams who are playing against ex professionals. Daniele says to Luisa:*]
- a. “Spero che tu giochi contro Baggio.” \approx *I hope that you play against Baggio.*
b. *“Spero che tu giocare contro Baggio.” \approx *I hope that you to play against Baggio.*

These benchmarking items were included to make sure the participants understood the task correctly and that they were attentive to the properties of both the contexts and the target sentences throughout the experiment. Eight of our participants failed to differentiate between the two conditions of the benchmarking items; i.e., their overall means for the six items of the form (19-a) were not higher than those for the six items of the form (19-b), which is a strong indication that they failed to perform the task correctly. Accordingly, these eight participants were replaced.

Overall, each participant encountered 24 items and $48+12=60$ filler/benchmarking items, resulting in an item:filler ratio of 1:2.5. Since one set of the filler items tested eight conditions, we created 24 lists (24 being the least common multiple of 6 and 8) by assigning the items to the lists according to a latin square design. This means that each item appeared in one list in one condition only, and that, across lists, an equal number of conditions was tested for each item. The lists were pseudo-randomized with an R routine that made sure that two experimental items were separated by at least one filler. The 24 pseudorandomized lists were then inverted to test for any effects of ordering, resulting in an overall number of 48 lists. These item lists were embedded into a L^AT_EXtemplate that required participants to fill in a few sociographic data (native language, gender, and year of birth), and which contained a short instruction. This instruction familiarized the participants with the task by means of three sample stimuli one of the context+target type: one grammatical, one ungrammatical, and one where the target

sentence did not fit the context. For each of these, sample ratings were provided to explain the use of the scale. Participants were further asked not to apply normative rules, but to judge the sentences according to their intuition, and to make use of all the values on the scale.

The 48 questionnaires were then handed to the students in class. The filled-out questionnaires were then digitized, and the resulting judgment files were assigned to the respective item/condition files from the latin square procedure. There was one judgment missing for one of the experimental items due to a participant accidentally failing to provide a judgment for that item.

3.3 Results

In Table 1, we give an overview of the results. All analyses were run in R (R Core Team (2024)). The raw data and the R/RMarkdown script for the analysis are provided on OSF (todo).

COREF	VERB FORM	Mean	Median	SD	SE	N
+	active	3.48	3	2.06	0.15	192
-	active	5.38	6	1.84	0.13	192
+	passive	4.22	5	2.15	0.16	191
-	passive	4.94	6	2.21	0.16	192
+	modal	4.76	5	2.11	0.15	192
-	modal	5.88	7	1.79	0.13	192

Table 1: Descriptive statistics dependent on experimental condition

To assess how our hypothesis (17) fares when held against the data, we have to look at the data in terms of differences of differences. The data show a difference between the effect of the factor COREF on the active sentences ($3.48 - 5.38 = -1.9$) on the one hand, and on the passives ($4.22 - 4.92 = -0.7$) and the modals ($4.76 - 5.88 = -1.12$), on the other hand. The difference in differences is perhaps more graspable in a graphical depiction of the data pattern (given here in terms of means and ± 1 SE error bars), which we give in Figure 1.

To test for the reliability of the differences between differences, we computed a series of ordinal regression mixed models (see Liddell and Kruschke, 2018; the R package used was `ordinal`, see Christensen, 2023). We first computed a model for the 2×3 design to check for the reliability of the two 2×2 interactions (COREF \times active vs. passive, and COREF \times active vs. modal). For this analysis, and the ones following it, we followed the recommendations in Schad et al. (2020) and coded all contrasts as sum contrasts when testing for interactions, thus rendering the contrasts for our factors orthogonal. To assess the reliability of interactions, we performed a model comparison with the additive model using likelihood-ratio tests and ran the

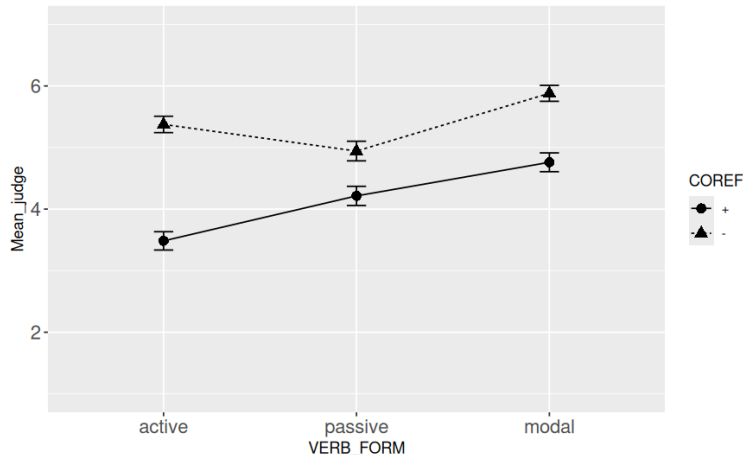


Figure 1: Means and ± 1 SE error bars dependent on condition

models with maximal random effect structure (see Barr et al. 2013), with the exception of the model for the modals; on which see below.

The full 2×3 model showed significant effects for both interactions, and the model comparison with the additive model was flagged as significant, too ($LR-\chi^2_{(24)} = 41.79, p < .05$; see online materials for the full analysis). This result, according to common practice, allows us to look into the results for 2×2 interactions.

First, we provide the model summary of the fixed effect coefficients for the subdesign COREF \times (active vs. passive):

	Estimate	Std. Error	z value	$\Pr(> z)$
COREF1	-0.6539	0.1101	-5.940	0.00000000285
VERB_FORMpassive	0.1017	0.1405	0.724	0.469251
COREF1:VERB_FORMpassive	0.3750	0.1138	3.296	0.000981

Table 2: Model output for subdesign COREF \times (active vs. passive)

As the third line in Table 2 shows, the interaction between the factor COREF and the two levels active vs. passive of the VERB_FORM factor is significant; this was corroborated by running the model comparison with the additive model ($LR-\chi^2_{(9)} = 31.47, p < .001$). Even if we correct for multiple comparisons (e.g., by applying Bonferroni’s correction), the result is still reliable: the difference in acceptability between the two coreference conditions is smaller in the passive conditions than it is in the active conditions.

Second, we checked for the reliability of the interaction in the subdesign COREF \times (active vs. modal). The model with the maximal random effect structure did not converge; we thus

removed the random factor with the least variance, COREF, from the random slopes, and reran the model. The summary is given below in Table 3.

	Estimate	Std. Error	z value	$\Pr(> z)$
COREF1	-0.55172	0.05740	-9.612	< 0.0000000000000002
VERB_FORMmodal	0.61361	0.10569	5.806	0.00000000641
COREF1:VERB_FORMmodal	0.11786	0.08095	1.456	0.145

Table 3: Model output for subdesign COREF \times (active vs. modal)

The third line in Table 3 informs us that the interaction between COREF and the two levels active vs. modal of the VERB_FORM factor is not significant; running the model comparison with the additive model again corroborates this result ($LR\text{-}\chi^2_{(1)} = 2.12, p > .10$).¹⁷

Taken together, the results falsify our strong hypothesis H_1 , and provide only evidence for our weaker hypothesis H'_1 : on the one hand, there clearly is a smaller difference in acceptability in the passive condition between the two coreference conditions: the difference participants made in their judgments between the coreferential and the non-coreferential condition was reliably smaller for the passive than for the active condition. This is in line with the first conjunct of H_1 , and can be interpreted as an instance of the amelioration effect. On the other hand, the somewhat smaller difference between the coreferential and the non-coreferential condition in the modal condition, as compared to the active condition, turned out not to be statistically reliable. Thus, we cannot uphold the second conjunct of our strong hypothesis, and thus H_1 as a whole, and have to retrench to its weaker version, H'_1 . Because the strong hypothesis is not confirmed, we could not use the differences between differences measure to compare the degree of amelioration in the passive and the modal conditions.

4 Discussion

Our experimental results show that amelioration of the basic subject obviation effect can be obtained when sentences are presented in context, and if the target coreferential structure is combined with a passive in the embedded clause: the difference between the coreferential and the non-coreferential condition is significantly smaller in the passive conditions than in the active conditions. This is in line with the original observation for French in Ruwet (1984) that passives tend to ameliorate sentences with subject obviation; and it shows that an amelioration

¹⁷We get the same result in terms of reliability if we remove the VERB_FORM factor from the random slopes. As a methodological aside, we may note that the results for by-participant and by-item ANOVAs are significant for the interaction, which might be taken to add further evidence for type I errors produced by running linear models on ordinal data; see Liddell and Kruschke (2018).

effect for passives can be obtained under highly controlled experimental conditions. However, modals did not show the predicted amelioration effect (*contra* Ruwet): the difference between the coreferential and the non-coreferential condition is not reliably smaller in the modal conditions than in the active conditions. Given the presence of the effect for passives in the same experiment, we are tempted to interpret this null effect as meaningful: at least in our experiment, modals did not show an amelioration effect. In what follows, we discuss our interpretation of the results of the experiment, its limitations, and questions for future research.

We start with the results for the modal condition. There are two points to be discussed here. The first point is a minor one that does not directly concern the purpose of our experiment. As we saw in Figure 1, the modalized sentences show overall higher acceptability than their non-modalized counterparts (in both coreferential and non-coreferential conditions). This can be a processing effect due to the addition of the modal. The second point directly relates to the purpose of our study and concerns the absence of evidence for the amelioration effect with modals. This absence can be interpreted in a number of ways. The most obvious one is to conceive of this as a null effect, following the rule “Absence of evidence is not evidence of absence.” However, given the presence of an effect for the passives, we are tempted to deviate from this rule and speculate that, at least in the experimental setting we tested, the amelioration effect of modals is at least weaker than that of the passives. This follows directly from our assumption of a *relative* acceptability measure: within one experiment, the different types of items tend to interact with each other, giving rise to what might be called “scale-gauging” effects: while modals failed to show a reliable effect in our experiment vis-à-vis the passive structures, they might show a reliable effect when run in an experiment where they have to compete with a different structure (say, e.g., negation). It lies in the very nature of linguistic acceptability that it is a highly context-dependent, relative measure; and our current results, and the way they deviate from those obtained in Feldhausen & Buchczyk’s (2021) study, bear witness to this.

However, our speculation about the possible meaningfulness of the null results for the modals is limited by the following considerations. For a start, in our experiment, we looked only at one obviative construction (with *sperare* ‘hope’) in one language (Italian). This is clearly not enough for making a far-reaching conclusion about other obviative constructions and other languages. But even for this particular obviative construction in this particular language, the absence of the effect with modals can be explained by a number of reasons, including the facets of the experiment itself. As discussed in the introduction and then in section 3, our goal was to test the amelioration effect in highly controlled settings using a fully factorial design, but this design came with some costs. For example, we had to use the same contexts for all three conditions (actives, passives, and modals). The contexts were constructed to make the interpretation of both active and passive sentences equally plausible. But the same contexts might have affected the interpretation of modals in an unpredictable way. Similarly, the lexical predicates used in

the embedded clause were chosen primarily based on how natural they are in both active and passive constructions with little regard to their interaction with modals. Finally, the fact that sentences with modals are overall judged as more acceptable might have had an impact. To determine whether modals can (or cannot) improve sentences with subject obviation, the above mentioned factors has to be disentangled and more experimental work done.

The absence of evidence for the amelioration effect with modals might be taken to cast doubt on the finding for the passives, as well. After all, the smaller difference for the passive conditions might be argued to have come about by the non-coreferential/passive condition showing a drop in acceptability; if it weren't for this drop, the passive conditions might not have showed an amelioration effect, either; or so the critical reader may argue.

We agree that, at first blush, the “drop” in acceptability seems to have acted in favour of our hypothesis. But this drop should then have affected *both* the non-coreferential and the coreferential condition in the passives, contrary to what the data seem to suggest at first sight. What turns out to be the case is that in the non-coreferential scenario, the passive forms seems less natural than the active one. That is no so surprising. It is known since Givón (1991) seminal work that passives are marked with respect to their active counterparts in all respects. They are less likely to be used, are harder to process and acquired later. In plain competition with actives, actives prevail: In out-of-the-blue scenarios the active form is indeed preferred over the passive form. If Sally is chasing Doko, it is more natural to say *Sally is chasing Doko* rather than *Doko is (being) chased by Sally*. The reason is that the subject of an active is the Agent and that of a passive is the Patient/Theme. In terms of prominence, in the active form, the Agent is more prominent, in the passive form, the Patient/Theme is.

Now in the contexts that we provided for the items, both the coreferential and the non-coreferential reading had to be equally plausible in order to avoid a confound of coreferentiality and plausibility-in-context. But given the markedness of the passives, the usage of the passive forms was independently decreased.

However, as the data show, only the non-coreferential condition actually shows this drop, while the coreferential condition does not. This actually confirms our hypothesis with respect to passives. Here, its acceptability is boosted by the amelioration effect, which counteracted the penalty for the passive. The active form was significantly degraded, but the passive form was not. If the active form was equally good as the passive, in the co-referential condition, we should have expected the same results as in the non-co-referential condition. Apparently, the co-referential active is not a salient competitor for the passive, whereas the non-coreferential active is.¹⁸

¹⁸Although it should be mentioned that a descriptive difference of 0.731 between the coreferential/active and the coreferential/passive condition might not seem like a mentionable difference on a 7-point scale, it turned out to be significant in a *post-hoc* comparison using the `emmeans` package; see Lenth (2024): $|z| = 3.347, p < .01$.

5 Conclusion

The goal of this study was to show that the amelioration effect found for sentence with subject obviation can be empirically tested in highly controlled settings using a fully factorial within-within design. We tested the amelioration effect in Italian for sentences with *sperare* ‘hope’ and two ameliorating factors: (i) the use of a passive and (ii) the use of a modal in the embedded clause. We found evidence for the amelioration effect in case of passives, but not in case of modals. However, we interpreted the absence of evidence for the amelioration effect with modals as possibly being due to the limitations of our experiment. Our results show that, in principle, it is possible to empirically test the amelioration effect using rigorous experimental methods, although given the complexity of the phenomenon, it is necessary to proceed with caution by small steps, and be aware of the dependence of possible amelioration effects on the experimental conditions under which they are obtained. We hope that further research proceeding in small, controlled steps might yield a clearer picture of the factors involved in ameliorating obviation effects in different languages. Since the amelioration effect is a key to understanding subject obviation, this might ultimately lead to a clearer theoretical picture of obviation itself.

References

- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3).
- Bouchard, D. (1982). *On the content of empty categories*. PhD thesis, Massachusetts Institute of Technology.
- Christensen, R. H. B. (2023). *ordinal—Regression Models for Ordinal Data*. R package version 2023.12-4.
- Costantini, F. (2009). *Interface perspectives on clausal complementation: the case of subjunctive obviation*. Libreria Editrice Cafoscarina.
- Costantini, F. (2016). Subject obviation as a semantic failure: a preliminary account. *Annali Di Ca’Foscari*, 50:109–131.
- Costantini, F. (2023). On some epistemic access effects. In Goncharov, J. and Zeijlstra, H., editors, *Agency and Intentions in Language*, volume 6 of *Brill Research Perspectives in Linguistics*. Brill, Leiden.
- Farkas, D. (1988). On obligatory control. *Linguistics and Philosophy*, 11:27–58.
- Farkas, D. (1992). On obviation. In Sag, I. and Szabolcsi, A., editors, *Lexical matters*, pages 85–109. CSLI.
- Featherston, S. (2009). A scale for measuring well-formedness: Why syntax needs boiling and freezing points. In Featherston, S. and Winkler, S., editors, *The Fruits of Empirical Linguistics I. Process*, volume 1, Process, pages 47–74. Mouton de Gruyter, Berlin, New York.
- Feldhausen, I. (2010). *Sentential form and prosodic structure of Catalan (Linguistik Aktuell/Linguistics Today 168)*. John Benjamins.
- Feldhausen, I. and Buchczyk, S. (2021). Revisiting subjunctive obviation in French: a formal acceptability judgment study. *Glossa: A Journal of General Linguistics*, 6:59.
- Gerbrich, H., Schreier, V., and Featherston, S. (2019). Standard items for english judgment studies: Syntax and semantics. In Featherston, S., Hörnig, R., von Wietersheim, S., and Winkler, S., editors, *Experiments in Focus*, pages 305–328. De Gruyter, Berlin, Boston.
- Givón, T. (1991). Markedness in grammar: Distributional, communicative and cognitive correlates of syntactic structure. *Studies in Language*, 15.

- Kaufmann, M. (2019). Who controls who (or what). In Blake, K., Davis, F., Lamp, K., and Rhyne, J., editors, *Proceedings of SALT 29*, pages 636–664.
- Kempchinsky, P. (1987). The subjunctive disjoint reference effect. In Neidle, C. and Nuñez Cedeño, R., editors, *Studies in Romance Linguistics*, pages 123–140. Foris.
- Kempchinsky, P. (2009). What can the subjunctive disjoint reference effect tell us about the subjunctive? *Lingua*, 119:1788–1810.
- Lenth, R. V. (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.10.2.
- Liddell, T. M. and Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, 79:328–348.
- Picallo, C. (1985). *Opaque domains*. PhD thesis, City University of New York.
- R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Raposo, E. (1986). Some asymmetries in the binding theory in Romance. *The Linguistic Review*, 5:75–110.
- Ruwet, N. (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*, 7:75–138.
- Ruwet, N. (1991). Je veux partir/*Je veux que je parte: On the distribution of finite complements and infinitival complements in french. In Goldsmith, J., editor, *Syntax and Human Experience*. University of Chicago Press.
- Schad, D. J., Vasishth, S., Hohenstein, S., and Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, 110:104038.
- Schlenker, P. (2005). The lazy Frenchman’s approach to the subjunctive. In Geerts, T., van Ginneken, I., and Jacobs, H., editors, *Romance languages and linguistic theory 2003*, pages 269–309. John Benjamins.
- Sprouse, J., Schütze, C. T., and Almeida, D. (2013). A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010. *Glossa: A Journal of General Linguistics*, 134:219–248.

Stegovec, A. (2019). Perspectival control and obviation in directive clauses. *Natural Language Semantics*, 27(1):47–94.

Szabolcsi, A. (2021). Obviation in Hungarian: what is its scope, and is it due to competition? *Glossa: A Journal of General Linguistics*, 6:1–28.