

Parameters on the obligatoriness of *too**

Pascal Amsili, Emilia Ellsiepen, and Grégoire Winterstein

Laboratoire de Linguistique Formelle, CNRS–Université Paris Diderot
pascal.amsili@linguist.jussieu.fr,
emilia.ellsiepen@linguist.jussieu.fr,
gregoire.winterstein@linguist.jussieu.fr

Abstract. This paper deals with the property of *obligatoriness* that is often described as characteristic of the class of *additive* particles (e.g. the adverb *too*). In a nutshell, these particles are said to be obligatory because omitting them in a discourse creates either an infelicity or the derivation of unwanted inferences.

A common trend of analysis treats this property as boolean. The general consensus used to be that if an additive particle can be used in a discourse then it **has** to be used. However some examples show that this property is probably best treated as a gradient. In this work we investigate some of the conditions that affect this obligatoriness. First, based on a small experiment with corpora, we evaluate the frequency of the obligatory *vs.* optional uses of *too*. Then, having established that the optional cases are not the exception, we present the results of an experiment that evaluates the effect of ellipsis and anaphora on the obligatoriness of additive particles.

This paper deals with the property of *obligatoriness* that is often described as characteristic of the class of *additive* particles (e.g. the adverb *too*). In a nutshell, these particles are said to be obligatory because omitting them in a discourse creates either an infelicity or the derivation of unwanted inferences.

A common trend of analysis treats this property as boolean. The general consensus used to be that if an additive particle can be used in a discourse then it **has** to be used. However some examples show that this property is probably best treated as a gradient. In this work we investigate some of the conditions that affect this obligatoriness. First, based on a small experiment with corpora, we evaluate the frequency of the obligatory *vs.* optional uses of *too*. Then, having established that the optional cases are not the exception, we present the results of an experiment that evaluates the effect of ellipsis and anaphora on the obligatoriness of additive particles.

The corpus and experimental parts of this work deal with the French equivalent of *too*, namely the adverb *aussi*. For the purpose of this paper we consider

* The authors wish to thank Barbara Hemforth as well as the audiences of the “Workshop on Anaphora” in Paris and the “Journées d’études - Marquage et non-marquage des relations de discours” in Lyon for their valuable input and comments on this work.

that the two elements are sufficiently similar to apply to *aussi* the theoretical considerations that have been proposed for *too* and vice-versa. We do not however claim that one is the perfect equivalent of the other. For one thing *aussi* can be used in environments that would not allow *too* in English (e.g. comparative constructions, cf. *infra*).

1 Empirical domain: the obligatoriness of *too*

The obligatoriness of additive particles is illustrated with the particle *too* in the pairs (1) and (2):

- (1) a. Jo had fish, and Mo did too.
b. #Jo had fish, and Mo did. *Green (1968)*

Although (1-b) might be acceptable with a specific prosody, there is a clear contrast with (1-a). The observation is rather surprising since it seems that *too* is not informative at all in (1-a). Obviously, Mo is not the only one who had fish, yet it appears more natural to redundantly mark this information with the additive particle.

In (2), the version without *too* is not ill-formed, but it yields a very different inference than the version with *too*.

- (2) [The 5000 m race was won by Gianni Romme.]
a. The 1500 m race was won by a Dutch skater.
∴ G. Romme is **not** Dutch.
b. The 1500 m race was won by a Dutch skater too.
∴ G. Romme is Dutch. *Sæbø(2004)*

If it is not known whether Gianni Romme is Dutch, then (2-a) conveys that he has a different nationality. On the other hand, if it is otherwise known that Gianni Romme is Dutch, then the pressure to use *too* is high, and (2-a) might even appear degraded because of a clash between an inference and world-knowledge.

Many analyses of *too* describe its semantics in simple presuppositional terms (mostly following the initial proposal by Horn (1972)):

- (3) a. *too* has no asserted content by itself: it does not change or contribute to the main content of its host sentence.
b. *too* presupposes that there exists an element distinct from its associate that satisfies the same predication.

Thus in (1-a), the use of *too* presupposes that someone different from Mo had fish, which is satisfied by the first segment of the discourse. Characterizations such as (3) led to a description of the distribution of *too* that can roughly be summarized as follows (usually motivated by principles such as *Maximize Presupposition!*, cf. Singh (2008)):

- (4) a. If the presupposition of *too* is already part of the common ground then its use is mandatory.

- b. If the presupposition is not fulfilled, it cannot be accommodated (unlike most presuppositions) and hence the use of *too* is not licensed.

There are a number of reasons why this account is too simple. We mention only two here. First, there are cases where the presupposition of *too* is satisfied and yet the use of *too* is not felicitous (5).

- (5) #John was almost on time. Paul was late too. *Winterstein (2010)*

In (5) the use of the additive *too* is degraded. Yet an antecedent is available for satisfying *too*'s presupposition in the form of a conventional implicature attached to *almost* (cf. Jayez & Tovenà (2008) on the semantics of this item).¹

We won't have much to say about those cases, and we will rather focus on a second group of problems for the accounts summarized earlier, namely the cases where *too* is optional.

For instance, in (6), it seems that *too* is not obligatory:

- (6) Jean est malade, Marie est malade, Paul est malade, tout le monde est malade alors !
John is sick, Marie is sick, Paul is sick, everybody is sick then!

In such a case, though, it seems that discourse structure plays a role to explain that *too* is not obligatory (cf. Amsili & Beyssade (2009)). But there are other cases where the discourse structure does not seem to be what matters.

- (7) Hartmann's joy was apparent in his beautifully cut hair, his expensive suit, his manicured hands, the faint aura of cologne that heralded his approach; in his mild and habitually smiling face, *too*, his expressive walk, in which the body, leaning slightly forward, seemed to indicate amiability.
Winterstein & Zeevat (in press)

A second important observation has been made by Kaplan (1984): not only is *too* optional in some cases, but it seems that the degree of obligatoriness varies with certain parameters, so that we can conclude that the obligatoriness of *too* is not a boolean property, but rather a gradable property.

In this work we investigate this property in two ways. First, we try to demonstrate the reality of the optional cases in a more rigorous way by relying on corpus studies (Sect. 2). Second, we present the results of an experiment based on a proposal by Kaplan (1984) about the parameters that influence the obligatoriness of *too*, namely on the effect of ellipsis in the sentence that hosts *too* (Sect. 3).

¹ An example like (i) shows that non-main contents such as conventional implicature and presupposition can function as antecedents for the presupposition of *too*:

- (i) a. John, that idiot, forgot my birthday. And Paul is an idiot too.
 b. John regrets selling his car. Paul sold his too.

Therefore the infelicity of (5) cannot be attributed to the "non main content" status of the antecedent of the presupposition of *too*.

2 Corpus study: is *too* removable?

2.1 Motivation

We took inspiration from a corpus study to which Winterstein & Zeevat (in press) only allude. The idea is that we can test the obligation to put *too* in the positions where it is found by trying to leave it out and judge whether the omission has any effect.

“A small probe of this kind by one of the authors on the English utterances of the Oslo Parallel Corpus gives obligatory cases and optional cases in roughly the same frequencies”

Winterstein & Zeevat (in press).

More precise figures are not given, so it was decided to perform a similar study with a French corpus.

The chosen corpus is a novel from Jules Verne, *Cinq semaines en ballon*, published in 1863 (J. Hetzel et Compagnie), which contains roughly 82 000 words. The number of additive particle proved surprisingly low: in total, 10 occurrences of additive *aussi* (*too*)², one occurrence of its negative polarity counterpart *non plus* (*either*), 7 occurrences of *également* (*~also*), and 9 occurrences of *de nouveau* (*~again*). Other additives, which occur even less frequently, were left aside. We chose to classify our samples, distinguishing among the cases where:

- there is no difference between the sentences with or without *too* (**optional** cases) ex. (8)
- the sentence without *too* becomes agrammatical (**obligatory** cases) ex. (9)
- removing *too* gives rise to new inferences, i.e. inferences that can be properly qualified and do not belong to the original text (**inferential** cases) ex. (10)

² The french adverb *aussi* can also occur in comparative constructions (i-a), and such constructions were not taken into account; other cases where arguably *aussi* can be seen as additive were discarded: cases where *aussi* is a discourse connective, usually extraposed at the beginning or at the end of a sentence (i-b), as well as cases where *aussi* is part of a correlative construction (i-c).

- (i) a. Il est bon de voir par ses propres yeux un homme d’une audace aussi extraordinaire !
It is good to see a man of such/AUSSI audacity with one’s own eyes.
- b. C’est au moyen de ces documents que des essais de cartes ont été tentés. Aussi vais-je suivre notre route sur l’une d’elles, et la rectifier au besoin.
Attempts at drawing maps were made with these documents. Thus (AUSSI) will I follow our path on one of those and correct it if need be.
- c. les trafiquants s’abritent non seulement contre les bêtes fauves, mais aussi contre les tribus pillardes de la contrée.
the smugglers protect themselves not only against wild beasts, but also/AUSSI against the pillaging tribes of the country.

- (8) Il se munit de trois ancres en fer bien éprouvées, ainsi que d’une échelle de soie légère et résistante, longue d’une cinquantaine de pieds. Il calcula également le poids exact de ses vivres;
He took three solid iron anchors as well as a silk ladder both light and resistant, some fifty feet long. He ÉGALEMENT calculated the exact weight of his food.
- (9) — Si nous étions à bonne portée, dit le chasseur, je m’amuserais à les démonter les uns après les autres.
 — Oui-da ! répondit Fergusson; mais ils seraient à bonne portée aussi, et notre Victoria offrirait un but trop facile aux balles de leurs longs mousquets ;
 — *If we were at good range, said the hunter, I would enjoy taking them down each at a time.*
 — *Truly! answered Fergusson; but they would be at good range AUSSI, and our Victoria would offer too easy a target for the bullets of their long rifles;*
- (10) — (...) ces peuplades sont considérées comme anthropophages.
 — Cela est-il certain ?
 — Très certain; on avait aussi prétendu que ces indigènes étaient pourvus d’une queue comme de simples quadrupèdes; mais on a bientôt reconnu que cet appendice appartenait aux peaux de bête dont ils sont revêtus.
 — (...) *these folks are said to be anthropophagous.*
 — *Is this certain?*
 — *Quite certain; it was AUSSI said that these natives had a tail like simple quadrupeds; but soon it was recognized that this appendage belonged to the skins that they wear.*

The results of the manual study are summarized in table 1.

Optional	9	33 %
Obligatory	ill-formed 11	} 66 %
	inference 7	

Table 1. Manual classification of 27 samples with an additive

This manual study revealed a number of issues: firstly, since we do not find a similar number as the one given by Zeevat, a reasonable estimation of the distribution among optional and obligatory cases remains to be done. Secondly, it was noticed that to classify these samples is not an easy task. That’s why we undertook an annotation study, to try to get a better understanding of the actual distribution and the relevant parameters.

2.2 Annotation study

The annotation study should be described as a pilot study: we asked 10 raters to classify 17 samples taken from the previous corpus. The three classes that we presented earlier were proposed, and we asked participants not only to attribute a category to every sample, but in addition to rate their degree of confidence on a 10 point scale.

The inter-annotator agreement turned out to be very poor: $\kappa = 0.22$. This confirms that the annotation task is (too) hard, and also suggests that the three classes that had been proposed might not be appropriate.

A closer inspection of the annotation data shows several facts:

- Since the overall agreement is so low, we don't get reliable figures for the distribution of the 17 samples into our 3 classes;
- however, there are cases where the inter-annotator agreement is reasonably high: it concerns about half of the items, and 6 of them are labelled **optional**, while 3 are labelled **obligatory**.
- Our category **inferential** is clearly the most problematic one, and this is confirmed by the confidence scores: there is a significant difference between the means for this category and the means for the other two, strongly suggesting that the annotators were not at ease with this category.

2.3 Conclusion

We can only draw weak conclusions from the two studies we have presented here. The first one is that the number of optional cases is far from being negligible. We still cannot decide whether, in literary texts, they represent 30 or 60%, but their massive presence has to be accounted for by any theory dealing with obligatoriness of additives. Obviously, it is worth investigate further on this issue, both by adding more corpora (of different genres), and by devising additional experiments. We plan to start an bigger annotation study, with a larger range of samples, a better control on their length, and only two categories instead of three. We have also started pure quantitative corpus studies.

The second conclusion we can draw is that it is not established that the inferential category even exist. Interviewing the raters led to the observation that for some of them, inferential cases were very close to obligatory case (implicitly assuming, as we did, that inferences were not wanted), but for many raters, the inferential cases were rather considered as similar to optional cases. Our further studies will not consider such a category, and will either use only two or a 10 point acceptability scale to address the issue.

3 The variability of obligatoriness: experimental investigation

3.1 Kaplan's parameters

In one of the first papers about *too*'s obligatoriness, Kaplan (1984) notes that the property of being obligatory is not boolean and depends on at least four

different parameters, one of them being the degree of anaphoric reduction of the host sentence of *too*. In the following we will consider the case of two clauses coordinated by *and* that differ in one argument (referred to as the contrastive topic_{ct}) while the remainder of the clause that we shall call comment, is repeated as in (11).

- (11) a. ?J_{ct} sent Helen a note and M_{ct} sent Helen a note
 b. *J_{ct} sent Helen a note and M_{ct} sent her one

While in the full form *too* does not seem to be fully obligatory (11-a), the sentence becomes much less acceptable without *too*, if part of the comment is reduced (11-b). Although Kaplan adopts the notion of variability of *too*'s obligatoriness here, his examples are labeled in a traditional system, suggesting that any reduction of the comment leads to ungrammaticality (marked by a star) without *too*, while the full form is more acceptable (marked by a question mark). We conducted an acceptability rating study to investigate whether in French the degree of reduction of the comment predicts the acceptability of the sentence if the French counterpart of *too* (*aussi*) is omitted.

3.2 Method

Twenty-four French test sentences similar to (11) were constructed. The two factors manipulated were Reduction and Aussi. Aussi had two levels (+,-) and Reduction had 6 levels: in *full* the comment was identical to the antecedent, in *cpt* the direct object was replaced by a pronoun, in *obl* the indirect object was replaced by a pronoun, in *pro* both complements were replaced by pronouns, in *vpe* the comment was replaced by the generic VP *l'a fait* ("did it") and in *vid* the comment was dropped completely. The crossing of Reduction and Aussi resulted in 12 conditions:

- (12) Jean a montré sa voiture à Paul, et Léa...
Jean has shown his car to Paul and Lea...
 full ... a montré sa voiture à Paul aussi
 ... a montré sa voiture à Paul
 ... *has shown her car to Paul (too)*
 cpt ... l'a montré à Paul aussi
 ... l'a montré à Paul
 ... *it has shown to Paul (too)*
 opl ... lui a montré sa voiture aussi
 ... lui a montré sa voiture
 ... *him has shown her car (too)*
 pro ... la lui a montré aussi
 ... la lui a montré
 ... *it him has shown (too)*

vpe ... l'a fait aussi
 ... l'a fait
 ... *it has done (too)*
 vid ... aussi
 ...
 ... (*too*)

These conditions were tested in two versions of the experiment: both contained the two baseline levels *full* and *vid*. Version A contained additionally *cpt* and *vpe*, while version B included *obl* and *pro*. Inside the two versions, the test sentences were distributed to eight lists using a latin square technique (i.e., each list contained each test sentence in only one condition and every list contained an equal number of test sentences in each condition).

The experiment was set up online using IbexFarm. Participants were presented the test sentence and were to judge on a ten-point scale how acceptable the sentence was. Our sentences were mixed with two other judgment experiments that functioned as fillers. All participants were native speakers of French.

3.3 Predictions

We firstly predict that in all levels of Reduction, the Aussi+ version is rated more acceptable because the repeated content licenses the use of *aussi*. Following Kaplan, we also predict for the Aussi- versions a higher acceptability in the full form than in the reduced forms. Generalizing Kaplan's prediction to increasing obligatoriness of *aussi* with increasing degree of reduction of the comment, we finally predict the difference in acceptability between the two Aussi versions to increase with the level of Reduction. For the canonical forms including *aussi*, a beneficial effect of Reduction is expected that penalizes the literal repetition.

3.4 Results

For the descriptive graph in Figure 1, the raw ratings were normalized by participant to account for the fact that people tend to use different portions of the scale. For the inferential analysis, we used generalized linear mixed effect models with random factors for participants and sentences and assessed the contribution of the factors and the interaction through model reduction. The factor Reduction was recoded as a numerical predictor: We scored one point of reduction for each pronominalization and two points for a complete drop, resulting in the following mapping: *ful* = 0, *cpt*, *obl* = 1, *pro* = 2, *vpe* = 5, *vid* = 6.

As apparent in Figure 1, Aussi+ versions were in general rated higher than their counterparts. This observation is confirmed by a highly significant main effect of Aussi in the inferential analysis ($\chi(1) = 415.08, p < .001$). The factor Reduction showed no main effect ($\chi(1) < 1$). Importantly, however, there was a significant interaction of Aussi and Reduction ($\chi(1) = 74.31, p < .001$): while in the conditions including *aussi* Reduction led to an increase of acceptability,

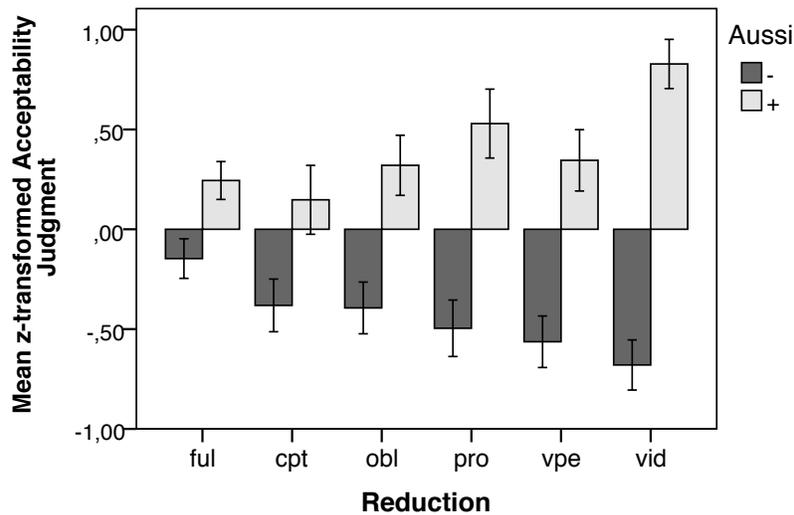


Fig. 1. Mean Judgments of Acceptability normalized by participant: 0 denotes average answer, positive values indicate higher acceptability with 1 being one standard deviation better than the average sentence.

the versions without *aussi* became less and less acceptable the more the comment was reduced. Especially the last part is nicely observable in Figure 1. The benefit of reduced forms including *aussi*, which was not the main focus of this experiment, seems to be less regular, especially for conditions *cpt+* and *vpe+*.

The results indicate that reduction of the comment indeed leads to a gradual decrease in acceptability, which is not captured by a boolean grammaticality property. The influence of reduction on the pressure to use *aussi* can be explained in two ways: Kaplan attributes this effect to the greater prominence of the contrastive topic in the reduced comment. Additionally, we see a connection to one of the other parameters he recognizes, namely the identity of reference. While the full form *Paul* could refer to a second individual with the same name in the comment, the pronominalized form forces the interpretation that the same person is referred to.

4 Conclusion

The two experiments we presented in this paper confirm that the question of the obligatoriness of *too* should not be treated as a boolean property, but rather as a gradient. The factors that affect this obligatoriness are probably numerous, and we only focused on one specific case where the host and the antecedent of *too* contain similar material.

Having shown that the reduction of the host has a deep impact on the obligatoriness of *too*, we have to explain why this is so. One explanation is to consider, as proposed by Amsili & Beyssade (2009) and Winterstein & Zeevat (in press), that the semantics of *too* goes beyond its sole presupposition and that it also functions as a marker of discourse similarity. One way to model this is to consider that *too* marks that its host and antecedent answer the same question under discussion. In the cases where the similarity is already salient in the discourse, the pressure to use *too* is high because its absence would convey that the elements that appear similar should not be treated as such (it is for such cases that analyses *à la* “Maximize Presupposition” are relevant).

Therefore to account for the results of the experiment of Sect. 3, we have to explain why this similarity appears higher in the cases of reduction than in the cases where there is a repetition of the lexical material. As already mentioned, when using a pronoun rather than the repetition of a proper name, there is no place for referential ambiguity, which can be seen as a stricter form of similarity. More generally if some linguistic material *A* is repeated with a form *A'* rather than referred to by anaphoric means, then this leaves more leeway to interpret *A'* as referring to a distinct token of whatever *A* denotes than in the anaphoric case. This means that the use of anaphora, and ellipsis as well, can be seen as a stronger mark of similarity than the lexical repetition of material. This however does not mean that lexical repetition cannot be interpreted as a mark of similarity.

The hypothesis that the use of *too* is linked to a notion of a gradient of similarity is partially confirmed by the experimental results presented in Winterstein (2011) where the acceptability of using *too* when relating quantifiers is shown to be linked to the similarity of the quantities denoted: the closer the quantities, the better the use of *too*.

- (13) a. Paul drank all his beer, and John drank all of his too.
b. ?Paul drank all his beer, and John drank some of his too.

In (13) even though *too* has an antecedent in both versions (by logical entailment in the second case) its use is better in (13-a) because the antecedent of *too* is more similar to *too*'s associate in (13-a) than in (13-b).

In future work we intend to pursue our work on corpora along the lines mentioned at the end of Sect. 2 and to further investigate the parameters mentioned by Kaplan: besides the degree of reduction he mentions three other parameters that could shed some light on the linguistic correlates of the notion of similarity that seems central in the distribution of *too*.

References

- Pascal AMSILI, Claire BEYSSADE (2009). “Obligatory presuppositions in Discourse”. In: Anton BENZ, Peter KUEHNLEIN, Candace SIDNER (eds.), *Constraints in Discourse*, Amsterdam and Philadelphia: Benjamins Publishers, vol. 2 of *Pragmatics and Beyond new series*.

- Georgia M. GREEN (1968). “On too and either, and not just too and either, either.”
In: B.J. DARDEN, C.J. BAILEY, A. DAVISON (eds.), *Papers from the 4th Regional Meeting*, Chicago: Chicago Linguistic Society, vol. 4, pp. 22–39.
- Larry HORN (1972). *On the Semantic Properties of Logical Operators in English*. Ph.D. thesis, Yale University.
- Jacques JAYEZ, Lucia TOVENA (2008). “Presque and almost: how argumentation derives from comparative meaning”. In: Olivier BONAMI, Patricia Cabredo HOFHERR (eds.), *Empirical Issues in Syntax and Semantics*. Paris: CNRS, vol. 7, pp. 1–23.
- Jeff KAPLAN (1984). “Obligatory too in English”. In: *Language* 60, 3, pp. 510–518.
- Kjell Johann SÆBØ (2004). “Conversational contrast and conventional parallel: Topic implicatures and additive presuppositions”. In: *Journal of Semantics* 21, 2, pp. 199–217.
- Raj SINGH (2008). *Modularity and Locality in Interpretation*. Ph.D. thesis, Massachusetts Institute of Technology.
- Grégoire WINTERSTEIN (2010). *La dimension probabiliste des marqueurs de discours. Nouvelles perspectives sur l’argumentation dans la langue*. Ph.D. thesis, Université Paris Diderot–Paris 7.
- (2011). “The Meaning of ’Too’: Presupposition, Argumentation and Optionality”. In: Nick BEZHANISHVILI, Sebastian LÖBNER, Kerstin SCHWABE, Luca SPADA (eds.), *Logic, Language, and Computation. TbiLLC 2009.*, Heidelberg: Springer, pp. 322–341.
- Grégoire WINTERSTEIN, Henk ZEEVAT (in press). “Empirical Constraints on Accounts of too”. In: *Lingua* .