Open Library of Humanities

The discursive function of additives in interaction

Alexandra Lorson, University of Groningen, NL, a.lorson@rug.nl

Hannah Rohde, University of Edinburgh, UK, hannah.rohde@ed.ac.uk

Chris Cummins, University of Edinburgh, UK, ccummins@exseed.ed.ac.uk

The insertion of additives such as *too* has been argued to be obligatory (in affirmative sentences) if the immediate context contains a suitable antecedent such that the presupposition triggered by additives is satisfied. However, the obligatoriness of additives has been found to be gradient and their insertion to depend on contextual factors. While most research has focused on comprehension, the present study examines the production of additives and the extent to which they are obligatory by manipulating the factors Similarity and Turn Distance. We furthermore explored whether not using additives even in obligatory environments could be an instance of diverging (i.e. socially distancing) from the antecedent speaker. For this purpose we investigated whether speakers would omit additives when interacting with an impolite antecedent speaker. Overall, the results of our two experiments suggest that (i) in line with previous results on similarity, speakers tend to utter additives more frequently when their utterance's content more closely matches the content of a previously formulated utterance; and (ii) speakers use additives more frequently when the matching utterance directly precedes their utterance. Furthermore, the results of experiment II suggest that (iii) speakers deliberately drop the use of additives when doing so would allow them to signal divergence from an impolite speaker. Our findings lend support to models in which speakers use additives as a discourse managing tool to organise the discourse and maintain discourse coherence.

1 Introduction

This study investigates the production of additive discourse particles such as *too*. Specifically, this study tests what contexts (if any) necessitate the presence of these additives and how sensitive their production is to discourse factors. Additives, many of which are often characterised as additive presupposition triggers, have been argued to be obligatory (in affirmative sentences) as soon as their presupposition is met in the context. For example in (1), the antecedent is (1a) which satisfies the presupposition that someone salient other than Donald watched Dune. Omitting *too* in the host sentence (1b) was claimed to lead to a marked discourse/pragmatic oddness, indicated by # (e.g. Bade 2016).

(1)   a.   Speaker1: Lisa saw Dune.
        b.   Speaker2: Donald saw Dune, #(too).

The obligatoriness of additives has been argued to be gradient (Kaplan 1984), and their insertion was posited to depend on context, discourse and information structure. For instance, the degree to which additives were judged to be obligatory was found to depend on the degree to which antecedent and host are similar (Spenader 2002; Amsili et al. 2016) and the distance between antecedent and host (Kim 2015; Chen & Husband 2018).

With one exception to our knowledge (i.e. Eckhardt & Fränkel 2012), the obligatoriness of additives has mainly been studied from the perspective of comprehension. The present study examines the production of additives in dialogue-like structures when the antecedent turn in the discourse context varies along two dimensions: Similarity and Turn Distance. Furthermore, this study extends previous research by investigating whether social factors play a role in the speaker's production choices. More specifically, we explored whether signalling similarity between antecedent and host utterance via additive use comes with converging, i.e. socially aligning (Giles 1973), with the antecedent speaker. We therefore investigated Politeness to see whether speakers omit additives more frequently when speaking with an impolite antecedent speaker (to avoid convergence) as compared to a neutral antecedent speaker. These three factors were examined by conducting two production experiments, one with an open-choice and one with a forced-choice paradigm. In the following we introduce accounts on the obligatoriness of additives, the discursive function of additives, and the potential connection between additive production and social convergence/divergence.

2 The obligatoriness of additives
2.1 Exhaustivity

The production of additives in affirmative sentences is said to be obligatory if there is a suitable antecedent in the preceding context (cf. Krifka 1998; Zeevat 2003; Sæbø 2004). For example in

(2), speaker2 should utter the particle *too* (2b), since the preceding utterance (2a) is a suitable antecedent. In contrast, omitting *too* leads to pragmatic oddness.

(2)   a. Speaker1: Lisa finished her PhD in Linguistics.
        b. Speaker2: EXH[Mattia]$_F$ finished his PhD in Linguistics, #(too).

Bade's 2016 account (based on Krifka 1998; Sæbø 2004) – henceforth the Obligatory Implicature account – explains this pragmatic oddness by assuming that (2b) gives rise to an exhaustivity implicature (i.e. that Mattia is the only salient individual who finished their PhD) which contradicts speaker1's utterance. In order to circumvent this markedness, *too* has to be inserted to block the exhaustivity implicature of (2b). The exhaustivity implicature arises due to the presence of focus – here on *Mattia* evoked by the contrastive phrase (2a). Focus causes the insertion of an exhaustivity operator (EXH) excluding all alternatives that are not entailed by (2b) making *Mattia* the exhaustive answer to the Question Under Discussion (QUD Roberts 1996) *Who finished their PhD?*[1]

The degree to which additives were judged to be obligatory has been found to depend on the extent to which exhaustivity is enforced by the context (Bade 2016; Bade & Renans 2021). Bade & Renans (2021) carried out two experiments in Ga (Kwa) and German, in which participants were asked to judge the acceptability of sentences with or without additives. Bade & Renans (2021) manipulated the strength of exhaustivity via context-sentence type (SVO, cleft), whereby exhaustivity should be stronger with cleft constructions (3b) than SVO (3a) structures. Here we give an example of the manipulation in English taken from Bade & Renans (2021):

(3)   Peter just came home. He is thinking about what he ate today and remembers suddenly.
        a. He ate a pear. (SVO)
        b. It was a pear that he ate. (cleft)

(4)   He ate a pineapple, too./He ate a pineapple.

In line with the Obligatory Implicatures account, Bade & Renans (2021) found that omitting additives for both languages was less accepted in contexts with strengthened exhaustivity (cleft) than in contexts with moderate exhaustivity (SVO). Exhaustivity, and with that the extent to

---

[1] An alternative account that is sometimes discussed in light of additives is the Maximise Presupposition (MP) account, (c.f. Amsili & Beyssade 2010). We did not discuss MP because there is an ongoing debate whether the insertion of additive particles follows from MP at all and if so, how different issues with this assumption can be dealt with. Our understanding is that this debate is based on the fact that there are a number of environments where additives have already been argued and shown to not be as obligatory as MP predicts them to be: antecedents of conditionals, negation and in the presence of conjunction and clefts (c.f. Bade 2016; Bade & Renans 2021). Similarly, some experimental work suggests that obligatory additives and expressions giving rise to MP effects, such as *the*, behave differently in processing (Aravind & Hackl 2017). To keep the paper streamlined we do not discuss the MP account in more detail.

which additives are perceived as obligatory, may additionally be influenced by discourse factors such as the degree to which antecedent and host are similar (Spenader 2002; Amsili et al. 2016) and the distance between antecedent and host (Kim 2015; Chen & Husband 2018) which we discuss below. Thus, while we did not test exhaustivity explicitly, we investigated factors that may influence the strength of the exhaustivity implicature; more specifically, we looked at similarity and proximity between antecedent and host utterance.

2.2 Similarity

In a corpus study using the London-Lund Corpus of Spoken English (LLC), Spenader (2002) found that *too* mainly served the purpose of signalling parallel information. Similarly, Amsili et al. (2016) found that *aussi* (French *too*) signals similarity of the sentence containing *aussi* – the host sentence – and its antecedent. Use of *aussi* increases the more similar the host and antecedent are, whereby increased similarity is measured by the presence of ellipsis or anaphora in the host. For example, in one condition the host sentence was identical to the antecedent (see (5a) for the English equivalent), whereas in the more reduced manipulations the post-verbal arguments were replaced gradually by pronouns (5b–5c).

(5)   Jean has shown his car to Paul and Lea…
          a.   …has shown her car to Paul, too.
          b.   …has shown her car to him, too.
          c.   …has shown it to him, too.

Amsili et al. (2016) found that the more reduced the host sentence was – and arguably more similar to the antecedent – the more *aussi* was preferred by addressees. Their results are based on the argument that while the full form *Paul* in (5a) may refer to a second individual named Paul (other than the Paul in (5)), the pronominalised form *him* in (5b) refers more clearly to the same individual as Paul in (5). Thus, due to pronominalisation, (5b) is more similar to (5) than (5a). Note that this argumentation is slightly different from Kaplan's (1984) account, who first observed the increased obligatoriness of additives in anaphorically reduced host sentences. According to Kaplan (1984), reduced host sentences introduce greater prominence of contrastive topic than full host sentences, and thereby higher pressure to use additives to emphasise the similarity between antecedent and host. From the speaker's perspective this would mean that with increasing similarity the pressure to utter additive particles such as *too* increases.

Winterstein & Zeevat's (2012) argument proceeds along similar lines whereby the insertion of *too* depends on argumentative identity. Winterstein & Zeevat (2012) claim that the more similar the antecedent and host are, the more they can be seen as being directed towards the same conversational goal, which increases the pressure to insert *too*. For example, (6b) seems less felicitous than (6a), since (6a) seems closer to the conversational goal of (6) than (6b). More

specifically, for (6b), it is argued that although *being almost on time* entails *being late*, (6) credits John with almost achieving timeliness, whereas (6b) gives no such credit to Mary. Assuming that *too* requires argumentative similarity between host and antecedent, (6b) is degraded because the two sentences involve contradictory predicates.

(6)   John was almost on time.
       a.   Mary was almost on time, too.
       b.   ?Mary was late, too.

In the current study, we extend previous findings on anaphorically reduced antecedents by investigating similarity in terms of mostly content-based similarity, in contrast to previous work that has emphasized similarity of surface form between host and antecedent. Including similarity in our production study is crucial to determine in what way similarity plays a role in the speaker's considerations when contributing to an ongoing discourse in comparison to the effect of similarity on the comprehension or acceptance of linguistic material.

2.3 Turn Distance

Another factor that was found to play a role for the insertion of additives was the distance between suitable antecedents and the host. This factor is related to what Kripke (2009) calls the anaphoric component of presuppositions triggered by additives: additives trigger a propositional alternative that is required to be pre-mentioned (in the form of an antecedent) and additives cannot just be uttered out of the blue. Similar to other anaphoric elements of language, whether content is perceived as a suitable antecedent may be related to locality or distance between antecedent and host.

    Kim (2015) found that the comprehension of additives is influenced by the linear distance between antecedent and host[2] such that comprehenders interpreted material that was most recent to be the antecedent of the additives. For (7f), the sentence (7e) was most frequently chosen as antecedent, yielding the interpretation that Andy bought a croissant and nectarines as opposed to celery, a croissant and nectarines. However, one may argue that while (7c) could in principle be interpreted as the antecedent as well, the usage of *usually* may endorse an interpretation in which Andy usually buys celery but did not do so today thereby removing "Andy bought celery" as a potential antecedent. Thus, the reason for which (7e) was the most frequent option may not solely have to do with locality.

(7)   a.   The roommates went to the farmer's market together.
       b.   Beth always buys bread.

---

[2] Kim (2015) also found evidence that the comprehension of additives is influenced by the hierarchical distance between antecedent and host. For brevity these results won't be discussed here.

    c.   Andy usually buys some celery.

    d.   His doctor told him he needs to eat more vegetables.

    e.   Today Andy treated himself to a croissant.

    f.   He also bought some NECTARINES.

Chen & Husband (2018) similarly manipulated linear distance between antecedent and host with a binary acceptability judgement task. In contrast to Kim (2015), Chen & Husband manipulated distance within an if-clause:

(8)   Near (satisfied):
        If the editor *resigned*, then the critics *resigned* too.

(9)   Far (satisfied):
        If the editor *resigned*, then everyone from the publishing house would be shocked to hear that the critics *resigned* too.

Chen & Husband also manipulated the context such that the context either satisfied (8/9) or violated (10/11) the presuppositions. They found that participants were more accurate – accepting the sentence in the satisfied condition (8/9) and not accepting the sentence in the violated (10/11) condition – in the near condition than in the far condition.

(10)   Near (violated):
        If the editor *plagiarised*, then the critics *resigned* too.

(11)   Far (violated):
        If the editor *plagiarised*, then everyone from the publishing house would be shocked to hear that the critics *resigned* too.

Both studies illustrate the potential role that proximity between antecedent and host may play in comprehension using written texts. It still remains to be investigated how proximity between a potential antecedent and host as measured by turn distance in dialogue-like structures influences additive production. By looking at mostly content-based similarity and proximity (turn distance), our study looks at two crucial factors that contribute to the structure of a discourse.

2.4 Social convergence/divergence

Considering that dialogues are placed in social settings which in turn influence speakers' production choices, we investigated the influence of social considerations on additive production. More specifically, the role of similarity and distance on the production of additives may extend to the social level in terms of similarity between speakers and social proximity. Thus, we focused on another possible function of additives that has not been investigated yet – at least to our knowledge – which is social convergence/divergence. More specifically, speakers who utter *too*

may not only signal similarity between their utterance and the previously mentioned antecedent utterance but may also draw parallels between themselves and the antecedent speaker. This line of reasoning builds on Winterstein & Zeevat's (2012) account of argumentative identity, which we introduced earlier as a way of construing the distance between antecedent and host in terms of whether they are directed towards the same conversational goal.

The concept of social convergence/divergence is introduced first within the framework of Communication Accommodation Theory (e.g. Giles 1973; 1977; 1979). The account is built on the assumption that interlocutors negotiate personal and social identities through linguistic, discursive and non-linguistic tools (Gallois et al. 2005). Thus, interlocutors may adjust their way of communicating to appear more alike (converge) or distinct (diverge) from other interlocutors. A possible motive for speakers to align their communicative behaviour could be the desire for approval (Gallois et al. 2005). In contrast, divergence may result from a speaker's wish to reinforce their social identity and/or dissociate themself from other interlocutors. Whereas converging speakers are most commonly received as cooperative, the reception of diverging speakers may vary. Divergence may be evaluated negatively, especially when it is judged to threaten social norms, e.g. politeness norms. However, divergence may likewise be received positively when diverging from one speaker entails converging with other interlocutors who share a common, positively valued group membership (Giles & Hewstone 1982).

Converging and diverging are very much intertwined with social norms such as politeness, a concept construed within Politeness Theory as interlocutors' aim to maintain their positive or negative face (Brown & Levinson 1987). Whereas positive face reflects the interlocutors' maintenance of a positive self-image, negative face reflects the interlocutors' freedom to act on their own terms. Face-threatening actions can damage the face of either the speaker or addressee. For example, speakers may aim to converge – e.g. by choosing similar topics as the other interlocutors – to maintain their positive face. In contrast, diverging from someone – e.g. by changing the topic – may maintain the speaker's negative face but may likewise threaten their positive face, since diverging can come across as impolite.

Applied to additive particles, using *too* to pick out an antecedent uttered by another speaker seems not only cooperative on the ontological level but also means that speakers converge with the antecedent speaker by signalling similarity between their utterances. This convergence move may come across as particularly strong if both antecedent utterance and host utterance convey values or opinions of both speakers, as in (12). In line with Communication Accommodation Theory, the motive for uttering *too* instead of leaving it out may reflect a general tendency to converge with speakers that share the same values (Gallois et al. 2005). However, one could also say that convergence is just a by-product of uttering *too* in obligatory environments. In this sense, uttering *too* is simply the easiest option to avoid making a controversial or face-threatening discourse move.

(12) *Context: Four colleagues are at lunch discussing which parties they have voted for in the last election. Rachel is the newest member of the department and has not made a good impression so far and the team generally tries to avoid including her in their lunch dates.*

    a.   Ross: I voted for Labour.
    b.   Joey: I voted for the Liberal Democrats.
    c.   Rachel: I voted for the Green party.
    d.   Monica: ?I voted for the Greens.

The derivation from the default of uttering *too* may be an instance of divergence. For example, one could assume that speakers, such as Monica in (12d), omit *too* and avoid acknowledging the parallelism between their utterance and the antecedent utterance (e.g. Rachel's utterance (12c)) in order to distance themselves from the antecedent speaker (e.g. Rachel). Or to phrase it differently, speakers may choose not to block the contradictory exhaustivity implicature with the insertion of additives, and leave the resulting contrast between their utterance and the antecedent utterance to highlight a contrast on the social level. A possible motive for diverging from the antecedent speaker could involve contrasting values/opinions or that the antecedent's behaviour threatens politeness norms. In this sense, speakers may wish to maintain their own negative face by not being associated with the antecedent speaker. In our study, we aim to explore whether omitting additives in potentially obligatory environments can be understood as a means to diverge from an interlocutor's turn. In this way we are able to see the extent to which speakers consider or disregard discourse constraints such as proximity depending on social factors such as politeness.

3 The present study

This paper investigates the circumstances under which speakers produce additive particles such as *too* when the antecedent turn in the discourse context varies along three dimensions: similarity, proximity and politeness. While similarity and proximity have been investigated in comprehension, this study tests both factors from the perspective of production and by looking at different senses of similarity and proximity, namely content-based similarity and turn distance. Additionally, we test the influence of politeness on additive production to see whether intentionally omitting additives can be seen as a speaker's attempt to diverge from an impolite antecedent speaker.

For this purpose, participants were asked virtually to attend a fictional work dinner. Their task was to interact with their colleagues and a waiter based on a sequence of prompts and visually presented cues about the intended content to convey. We manipulated similarity by asking participants to formulate an utterance whose content either perfectly matched or did not perfectly match the utterance of a previous speaker. Turn distance was manipulated by the participant's turn either immediately following the antecedent speaker's turn (0 intervening turns)

or following after three intervening turns. We acknowledge that we could have treated Similarity as a continuous variable ranging from identity to complete dissimilarity by letting an independent sample of participants judge similarity between antecedent and host utterance. Likewise, Turn Distance could have been treated as ordinal variable. While we do see these points as important considerations for future research, we did not have a sufficiently refined theory of the role of either factor to quantify how far along the relevant axis additional items would lie. Therefore, we chose to include Similarity and Turn Distance as binary factors and thereby compare their 'endpoints' (i.e. perfect similarity/dissimilarity, 0/3 turns distance) as an initial investigation. In order to test divergence we tested two groups, one in which participants conversed with polite/ neutral speakers, and one in which the antecedent speaker would be impolite. This study consists of two experiments: in experiment I participants were invited to produce free text responses; in experiment II, they selected what they would say from a set of options (informed by the results of experiment I). Data and material for all experiments as well as the pre-registrations can be found here: https://osf.io/az2uf/.

3.1 Experiment I

The first experiment investigates the production of additives in an open choice paradigm in which participants were asked to type in their response in a text box. We analyse the speaker's binary choice to include or omit an additive particle as our dependent variable. For the coding of our dependent variable, the first and second authors coded the participants' responses independently and discussed deviations in their coding together with the third author. Participants' responses were coded as containing an additive (= 1) if participants used the additive in their response to refer back to the antecedent utterance. Thus, we coded responses as not containing an additive (= 0) if a response didn't contain an additive to begin with, or if the additive did not unambiguously refer back to the antecedent utterance, such as in (13) and (14).

(13)   I'd like a Chardonnay, please. Can I **also** have some tap water?

(14)   I contributed £15 and bought him some chocolates **too**

We considered the following additives: *too, also, as well, same as* and *another,* such as in (15) and (16).[3]

(15)   'Order' item examples (participants were asked to order food/drinks)
        a.   chardonnay for me **too** please chilled if you would
        b.   Can I **also** get a Heineken, please

---

[3] We initially pre-registered an analysis of the presupposition triggers *too, as well* and *also.* After data inspection and in response to an anonymous reviewer we extended the discourse particles under investigation to the present set.

    c.   I'd like a glass of Chardonnay **as well**

    d.   Can I get the **same as** omar please

    e.   **Another** breaded Brie & prawn linguine, please

(16)   'Talk' item examples (participants were asked to converse about a range of topics)

    a.   I love Korean food **too**, especially Kimchi Jjigae

    b.   I **also** got him £15, wine and a food voucher

    c.   Yep Wednesday at 2 is what I've saved **as well**

    d.   Omar, I've done the **same as** you! £15 and a bottle of wine.

Besides additive production, participants used other linguistic tools to pick up the antecedent utterance such as in (17) and (18). We considered these instances in a further analysis which is part of Appendix 2.

(17)   Make that two please

(18)   Haha, yeah I was like you Robert

First, experiment I tests whether our experimental set-up elicits the use of additives and which additives participants produce most frequently. We analyse the speaker's binary choice to include or omit an additive particle (dependent variable) to see how that choice is affected by Similarity, Turns, and Politeness. In experiment II we used a forced-choice paradigm based on the responses of experiment I.

3.1.1 Participants

Participants (N = 78) were recruited from the crowd-sourcing platform Prolific. We recruited only those with an approval rate above 90. Participants were paid an average of £7.60/hour (the average duration of the experiment was 18 minutes). The age of the participants ranged from 18 to 83 years, with a mean of 37 years (median = 32, mode = 31). 40 participants stated their preferred pronoun as she/her, 37 chose he/him, and 1 chose they/them.

3.1.2 Design and Materials

Participants were asked to engage with their colleagues and a waiter by typing their responses in a text-field. Half of the critical items ('order' items) were about ordering dishes/drinks from a waiter, and half of the critical items ('talk' items) were about content related to work or daily life, e.g. the date and time of a meeting. Each of the eight critical items started with an introduction to the upcoming conversation; participants were presented with context pictures

together with a question about their content. See **Figure 1** for an 'order' item picture (always a menu) and **Figure 2** for a 'talk' item picture (here a planner). In this way, participants were able to familiarise themselves with the visual cues (the menu/planner) they would later need for formulating their response. For example, the 'order' item context already prepared participants that they would be asked to order a glass of Chardonnay when it was their turn to order. The visual cues were then presented to participants again (Menu with 'Chardonnay' highlighted in **Figures 5** and **6**). This set-up is an attempt to mimic actual conversation, since interlocutors in an actual conversation do already have some knowledge about what they can contribute about a given topic when listening to the contributions of other interlocutors. It also allowed us to constrain the participant's message content so that we could test what factors affected how they formulate an utterance about that content.

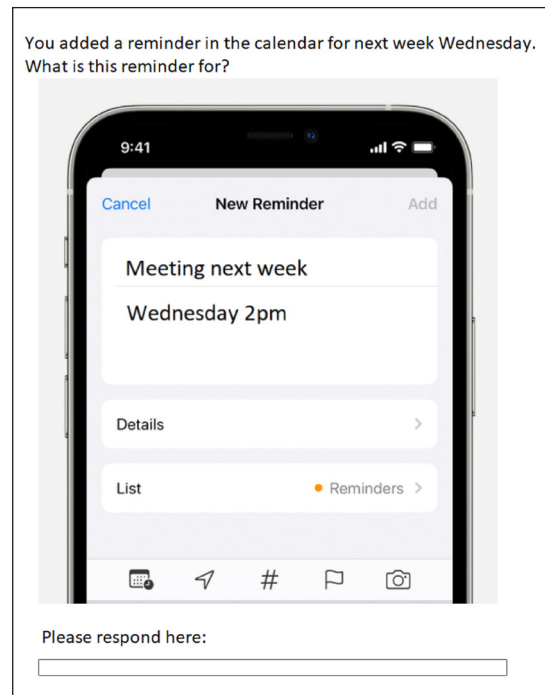**Figure 1:** Example of an 'order' item context.

Figure 2: Example of a 'talk' item context.

After having answered the context question, the conversation started. Each picture had a headline reinforcing the topic of the conversation. For 'order' items, the headlines would signify what interlocutors were ordering, e.g. 'Ordering drinks'; for 'talk' items the headline would say what the conversation was about, e.g. 'Discussion about the time of next week's meeting'. Each conversation started with Omar's turn followed by Lee, Amber and Robert, see **Figures 3** and **4** for the first three turns. Each conversational turn was displayed separately and participants had to click a button to proceed to the next one. The last turn (Robert) was always displayed together with the visual cue and together with the request for the participants to write their response, see **Figures 5** and **6**.

Turn Distance was manipulated such that the antecedent utterance either immediately preceded the participants' turn (0 turn distance, see **Figure 3**), or the antecedent and participants' turn were separated by the remaining colleagues (3 turn distance, see **Figure 4**). Thus, the antecedent speaker for the 0 turn condition was always Robert and the antecedent speaker for the 3 turn condition was always Omar.[4]

---

[4] As pointed out by a reviewer, when inspecting the seating arrangements in the picture, one can see that Turn Distance is confounded with the distance between where the antecedent speaker and the participant are sitting. We decided on this particular way of displaying the interlocutors due to practical reasons, e.g., to incorporate the waiter into the picture without changing the gaze or orientation of the interlocutors for the order versus talk items.

Figure 3: Example of the first three turns for an 'order' item in which the participant will order Chardonnay with no intervening turn between the antecedent speaker and the participant.

Figure 4: Example of the first three turns for an 'order' item in which the participant will order Chardonnay with three intervening turns between the antecedent speaker and the participant.
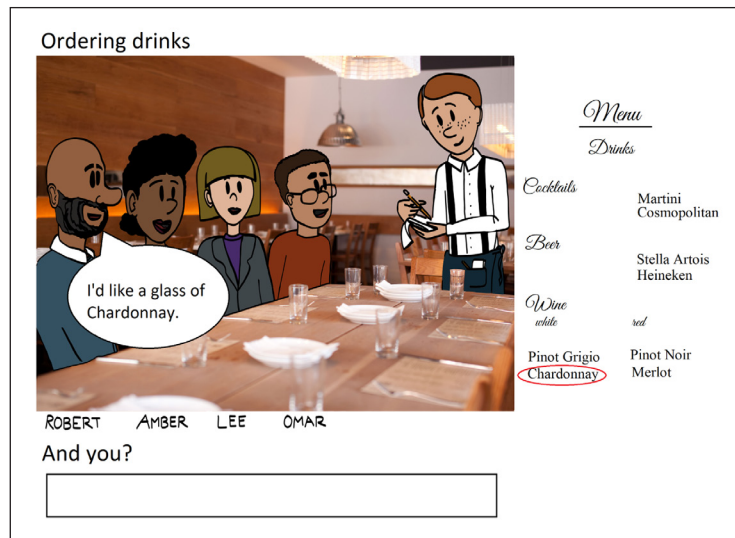
Figure 5: Example of a last turn for the 'order' items in the perfect similarity ('Chardonnay' is highlighted), no intervening turn condition. In the reduced similarity condition 'Pinot Grigio' would be highlighted.
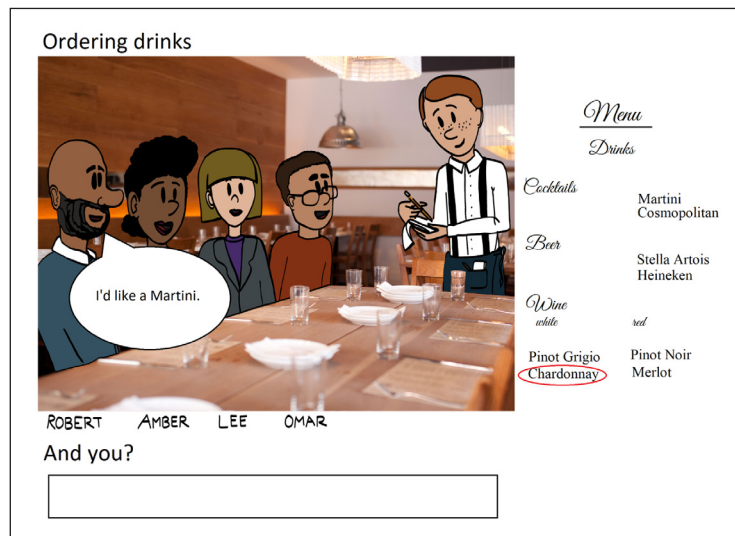

Figure 6: Example of a last turn for the 'order' items in the perfect similarity ('Chardonnay' is highlighted), three intervening turns condition. In the reduced similarity condition 'Pinot Grigio' would be highlighted.

To manipulate the factor Similarity, the picture cues either encouraged participants to give a response that was perfectly similar or only vaguely similar to the antecedent utterance produced by either Robert or Omar. Taking the 'order' items as an example, the drink/dishes we asked participants to order either completely corresponded (perfect similarity) or did not correspond/

only partially corresponded to the antecedent order (reduced similarity). For example, in **Figures 5** and **6**, participants were encouraged to order Chardonnay which perfectly matched Robert's or Omar's order. For the reduced similarity condition the participants' order would have been Pinot Grigio which did not match Robert's or Omar's order. This was slightly different for orders in which participants were asked to order two dishes: in the perfect similarity condition, both the highlighted starter and main matched the antecedent order, as opposed to the reduced similarity condition in which only the starter but not the main matched the antecedent order (partial match).[5] The focus of this study was primarily to investigate the difference between environments in which the additive usage is supposedly obligatory (perfect similarity) versus environments in which it is not (partial/no similarity). However, we briefly discuss how these different kinds of dissimilarities influenced the participants' use of additives for experiment I in Section 3.1.6 and more thoroughly in Section 3.2.6 for experiment II. We have illustrated the similarity manipulation for the 'talk' items in Appendix 1.[6]

We additionally investigated social divergence and manipulated the way the antecedent speakers phrased their orders: either neutrally, see **Figure 5**, or impolitely, see **Figure 7**. Participants were randomly assigned to either the neutral or impolite condition. In the neutral condition, all four colleagues behaved neutrally; in contrast, in the impolite condition, the two antecedent speakers (Omar and Robert) would make impolite remarks towards the waiter or colleagues. For this we included impolite material, see the bolded part of (19), to precede the utterance of the neutral condition, see the italic part of (19a) and (19b).

(19) a. Item 1, neutral condition:
    I'd like a glass of Chardonnay.
  b. Item 1, impolite condition:
    Although you served me warm white wine last time, *I'd like a glass of Chardonnay*.

Placing the impolite material first was a way to keep the neutral and impolite conditions comparable: placing impolite material after or within the utterance *I'd like a glass of Chardonnay* would have induced more distance between antecedent and participant turn as well as increased the dissimilarity between the turns. However, we do acknowledge that the impolite material increases the complexity of the antecedent turn overall. The antecedent speakers' impolite behaviour was restricted to the critical items and the attention checks stayed the same for both

---

[5] This manipulation was the same for the 'talk' items. Either both utterances were completely dissimilar (different favourite cuisines) or matched partially (same date but different month). In total, half of the critical items were manipulated to be dissimilar to the antecedent utterance and the other half to be a partial match.

[6] A limitation of our design is that we cannot overgeneralise our findings to a wide range of possible conversations/ conversational topics, since we targeted specific ones.

conditions. This resulted in a $2 \times 2 \times 2$ design, with two within-subjects factors (Similarity and Turn Distance) and one between-subjects factor (Politeness). We decided to test Politeness as a between-subjects manipulation to be able to investigate Similarity and Turn Distance in a completely neutral antecedent speaker set-up which more closely resembles what has previously been done.

Figure 7: Example of a last turn for the 'order' items in the perfect similarity ('Chardonnay' is highlighted), no intervening turn condition in which the antecedent speaker is impolite.

We tested whether interlocutors indeed perceive the added material in the impolite condition as being impolite by running an additional norming study using an independent sample. Participants (N = 30, British English speakers) were asked how they would rate the antecedent speakers' utterances choosing between the options *impolite*, *neutral* or *polite*. The utterances in question were presented as part of the same four turn dialogue used in the main study with the only exception that both antecedent utterances directly preceded the participants' task. Our results suggest that our politeness manipulation was successful: utterances in the impolite condition were most frequently judged as being impolite (88.3%) and much less so as being neutral (10.8%) or polite (0.83%). In contrast, utterances in the neutral condition were mostly judged as being either neutral (53.3%) or polite (41.5%) and only infrequently as being impolite (4.16%). The data is available in our OSF repository https://osf.io/az2uf/.

The 8 attention checks were similar to the 'talk' items. Participants were introduced to the visual cue, see **Figure 8**, followed by the colleagues' turns and their own turn, see **Figures 9** and

10. The only difference was that the colleagues' turns were not manipulated to induce participants to use additive particles. Participants who failed to answer the context question correctly, and also failed to base their conversational contribution on the visual cue, were considered as having failed the attention check.

Figure 8: Example of an attention check context.

The experiment consisted of four blocks, each containing two attention checks, one 'talk' item, and one 'order' item (block1: drinks, block2: appetizer & main, block3: dessert, block4: drinks). Thus, each participant saw 8 critical items.[7] Although the item order within each block was randomised, the order of the blocks was always the same.[8]

---

[7] We only tested 8 critical items because we are of the strong belief that carrying out short experiments that are engaging improves data quality, since participants pay attention throughout and don't have the time to develop any response strategies. In order to collect a dataset of comparable size to that in other studies in the psycholinguistic literature, we increased our dataset size by recruiting more participants.

[8] In order to see whether block order influenced the probability of producing additives (i.e. whether the participants became more or less likely to produce additives as the experiment went on) we conducted models with block order as a predictor for experiments I and II. We did not find such a trend. The only weak trend that the order effects model revealed is that the frequency of choosing additives drops for Block3 in comparison to the other blocks, which suggests that the items of Block3 deviated from the rest of the items: see Appendix 3 for more information.
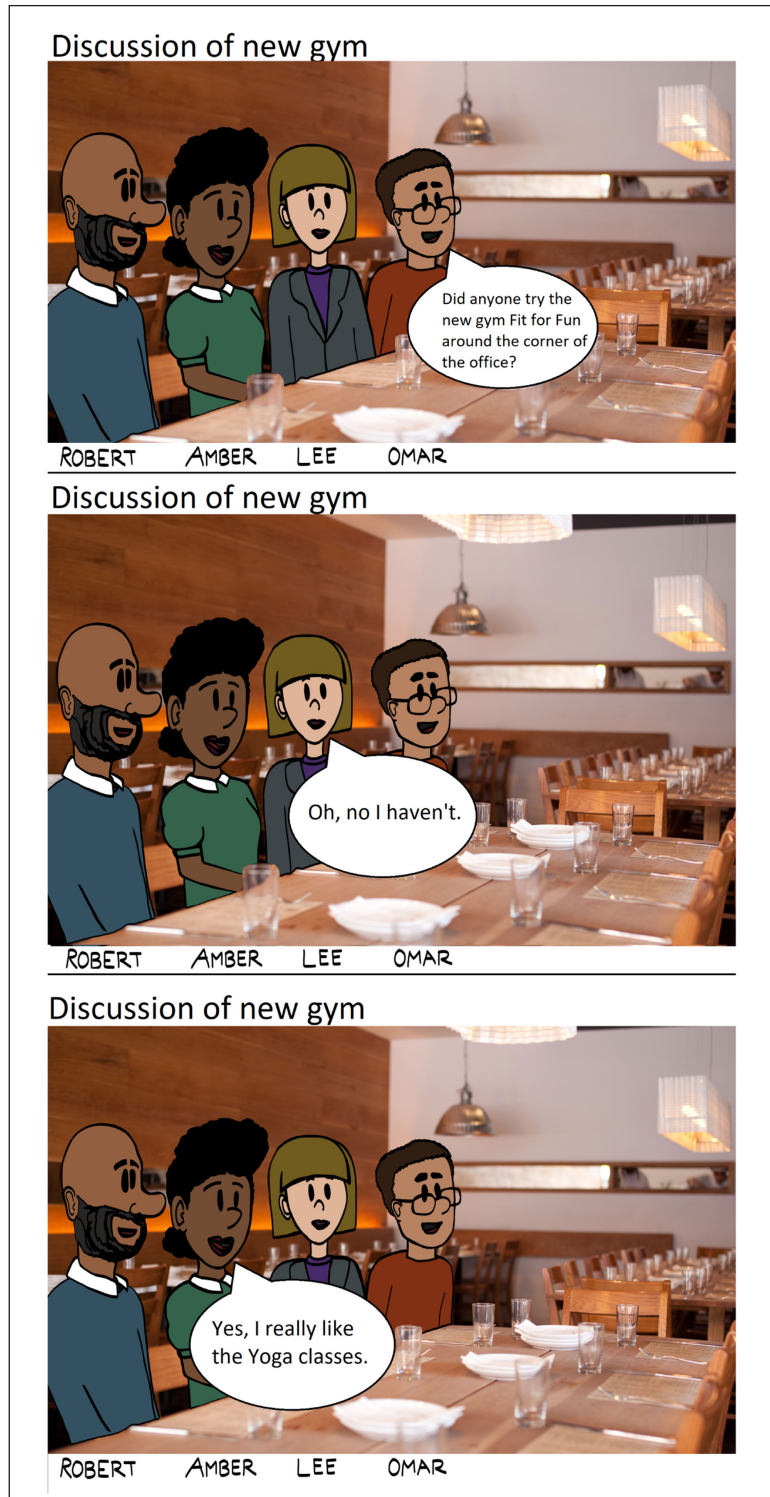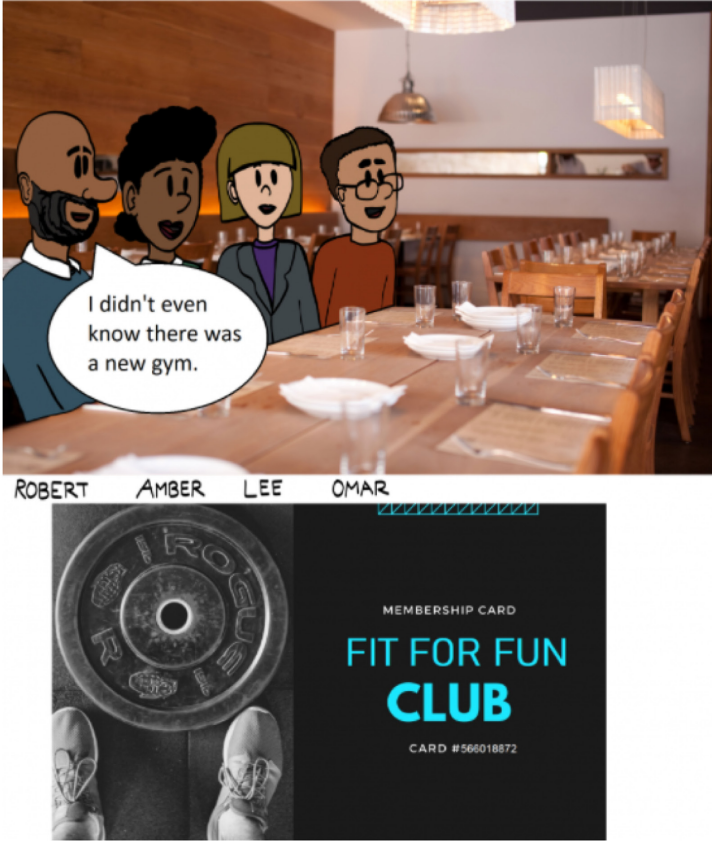
Figure 9: Example of an attention check: first three turns.

Figure 10: Example of an attention check: last turn.

3.1.3 Procedure

First, participants were asked to give informed consent to participate in the task. We then informed them about the structure of the experiment which was as follows: (1) engaging with different items of information (e.g. a to-do list) and interacting with fictional colleagues and a waiter, and (2) a question about the participants' understanding of the task and a demographic questionnaire. After that we introduced participants to the task and their four colleagues. We explained that each conversation would start with them being asked to answer a question about visual material (e.g. a menu) that is related to the upcoming conversation. Participants were instructed to follow the conversation and type their own conversational contribution when prompted based on the

visual material (e.g. a menu with highlighted dishes). We emphasised that there was no need to memorise anything at the beginning because the visual material that cues their contribution would be available throughout. Furthermore, we told participants that we were interested in the way they engage with others and asked participants to write their responses in such a way that they reflect their preferred way of speaking. The experiment was followed by a question about the participants' general understanding of the task – a measure ranging from 'did not understand at all' to 'completely understood' – and the voluntary demographic questionnaire which helped to characterise our sample.

3.1.4 Predictions

Our predictions are based on the previously discussed theoretical accounts and experimental studies. Overall, following the Obligatory Implicatures account, participants are predicted to most frequently use additives for highly similar and recent antecedents, as such a context would yield the strongest 'contrast' or exhaustivity effect. More specifically, we predict that with increasing similarity between antecedent and host sentence, the preference to use additives should increase. Thus, participants are predicted to produce additives more in the perfect similarity condition than in the reduced similarity condition. Furthermore, with increasing turn distance between antecedent and host sentence, the preference to use additives should decrease. Hence, participants are predicted to produce fewer additives in the 3 turns condition (three intervening turns) than in the 0 turns condition (zero intervening turns). If the use of additives reflects Politeness considerations, participants are predicted to use additives for emphasizing alignment with a polite speaker and in turn omit them for signalling divergence with an impolite speaker. More specifically, when antecedent speakers behave impolitely, participants are predicted to produce additives less in the impolite condition than in the neutral condition. We also explored the interactions of these three factors. For example, politeness could completely cancel a speaker's considerations regarding one or both of the discourse factors. More concretely, it was plausible that speakers consider discourse factors such as similarity and proximity only when interacting with polite as opposed to impolite antecedent speakers. This part is exploratory and informed our second experiment.

3.1.5 Analysis

We coded participants' responses for both experiments as either containing (1) or not containing (0) any of the following additives: *too, also, as well, another* and *same as*. Those instances that were coded with 1 where those in which the additive was used to refer to the antecedent utterance. The data was analysed fitting a Bayesian logistic regression model with varying by-item and by-participants intercepts and slopes using the R (R Core Team 2020) package brms (Bürkner 2018), which provides an interface to fit Bayesian mixed models using Stan (Stan Development

Team 2017).[9] The experimental factors Similarity, Turn Distance, Politeness and their interactions were included to predict the probability of producing additives. All three factors were sum-coded, whereby the levels perfect similarity, zero intervening turns and neutral politeness were the reference categories coded with 1 (the other levels were coded with –1).[10] The model included varying intercepts and slopes for participants and items, assuming that the effects of Similarity, Turn Distance and their interaction vary between participants, and the effects of Similarity, Turn Distance, Politeness and their interactions vary between items.

We used weakly regularising priors, which allowed a reasonably wide range of parameter values and penalised very extreme values. The priors for the by-expression intercepts were normal distributions with mean 0 and standard deviation 3. For both fixed effects, normal priors with a mean of 0 and a standard deviation of 1 were used. Random effects were modelled as a correlation matrix and a vector of standard deviations. The standard deviations were assigned half-normal priors with a mean of 0, and a standard deviation of 1. For the correlation matrix, a LKJ(2) prior was used such that smaller correlations are favoured over extreme values such as +/–1 (Stan Development Team 2017; Sorensen et al. 2016).[11]

Samples were drawn from the posterior distributions of the model parameters using the NUTS sampler (Hoffman & Gelman 2014). We ran four sampling chains, each collecting 4000 iterations whereby the first 1000 iterations were disregarded as part of the warm-up phase leading to 12000 iterations available for analysis. Chains mixed well (all $\hat{R}$ = 1.0).

Unlike the frequentist analysis, the Bayesian analysis does not produce point estimates but instead posterior distributions over parameters quantifying the probability of each possible parameter value given the data. We report the posterior mean $\hat{\beta}$ and the 95% credible interval (95%-CrI). The 95%-CrI is the range around the posterior mean within which the true value of the parameter lies with a probability of 0.95. We interpret the evidence as reliable if zero lies outside the parameters' 95% credible interval (Kruschke et al. 2012).

---

[9] Our model is similar to frequentist mixed effects logistic regression models. We chose the Bayesian framework for different reasons, one being that Bayesian models allow us to quantify the uncertainty around our estimates (intercept and slopes) directly.

[10] We had pre-registered both experiments with treatment coding but used sum-coding instead. This was due to the fact that (i) sum-coding helped the models to converge (especially for experiment I), (ii) using sum-coding, the associations of parameters correspond intuitively to a main effect of an independent variable being the average effect of changing between levels of the independent variable (Levy 2014).

[11] In line with our preregistration, we ran two models with more uninformative/wider priors, see Appendix 3 for details. Applying wider priors yielded coefficients with wide standard deviations and credible intervals. One of the model outputs suggests that the chains did not mixed well, which has to do with the combination of number of observations and too broad priors that assign probability mass to values that are not likely considering the model and data.

3.1.6 Results

Participants were above 96% accurate for the attention checks, which suggests that they paid attention during the experiment. When asked, participants indicated that they understood the task (mean = 83, sd = 20.6, mode = 100, median = 90). Overall, participants used the additive particles *too, also, as well, another,* and *same as* for target items throughout the experiment (in total 108 times out of 624 responses on target trials), and thus, 17% of the participants' utterances contained additives. The additive *same as* was the most frequent choice, see **Table 1** for more details.[12]

Additive	Item Type	Absolute Freq.	Overall
same as	Order Item	28	47
same as	Talk Item	19	
too	Order Item	10	23
too	Talk Item	13	
also	Order Item	13	23
also	Talk Item	10	
as well	Order Item	12	16
as well	Talk Item	4	
another	Order Item	3	3
another	Talk Item	0	

Table 1: Absolute frequencies of additives for the different item types and overall.

There were differences in how frequently additives were used across different types of reduced similarity items. Recall that half of the critical items were manipulated to be a categorical mismatch between antecedent and participant utterance, and the other half was manipulated to be a partial match. Participants would still occasionally use additives when there was a partial match –5% of all partial match trials contained additives – but not when there was no match. In contrast, 31% of all perfect match trials contained additives.

As **Figure 11** illustrates, additives were most frequently used when host utterance and antecedent utterance were similar and when there were zero intervening turns. Frequency was

---

[12] The sum of the absolute frequencies in Table 1 does not add up to 108 because the table counts account for the few instances in which participants would use two additives in one utterance, e.g. *I'm on the **same** brain wave **as** you there Robert ill **also** have the breaded brie and king prawn linguine please.* We only considered these instances as a single additive occurrence in the analysis.

reduced when three turns intervened. In contrast, the difference between the neutral and impolite antecedent speaker groups is less clear.
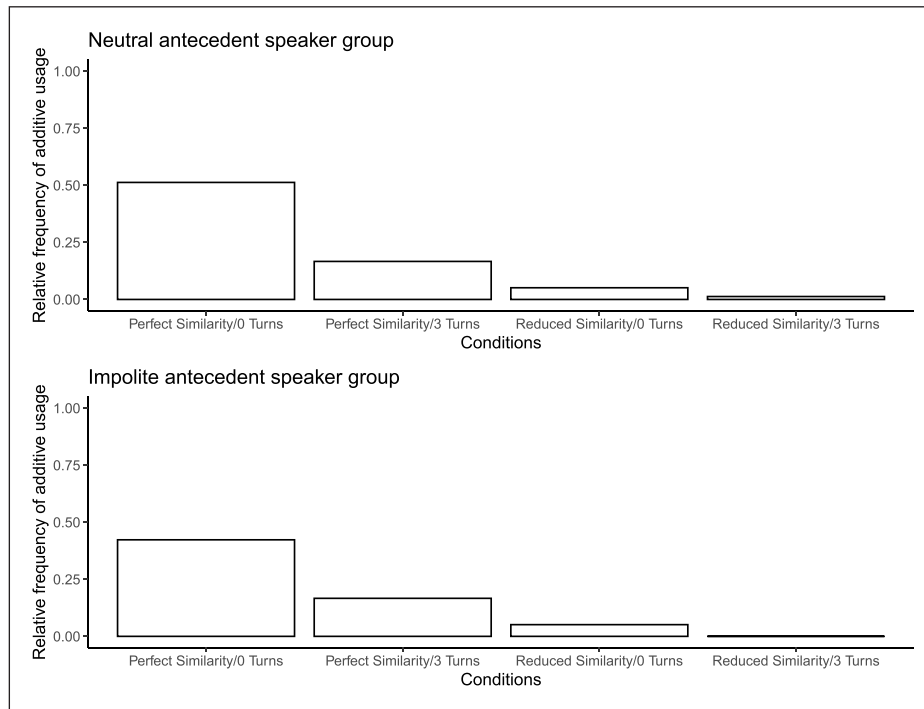

Figure 11: Experiment I: relative frequency of choosing additives in each of the Similarity × Turn Distance conditions, contrasting participants confronted with a neutral (top) versus impolite (bottom) antecedent speaker.

These observations are supported by the outcome of the analysis, see **Table 2**. Averaging over Turn Distance and Politeness, a change to the perfect similarity condition meant an increase in log-odds of additive production ($\hat{\beta} = 2.00$, CrI:[1.23, 2.78]). Similarly, averaging over Similarity and Politeness, a change to the zero turns condition meant an increase in log-odds of additive production ($\hat{\beta} = 1.10$, CrI:[0.49, 1.77]). Both effects appear reliable, since the credible intervals of both effects do not include zero. However, we did not find reliable evidence that the antecedent speakers' politeness influenced additive production. In **Figure 12**, we illustrated the predicted probabilities for speakers to produce additives given our data and model. Speakers are predicted to produce additives most frequently for perfect similarity and when no turns intervene between their utterance and the antecedent utterance and least frequently for reduced similarity when three turns intervene. We found almost identical effects when considering all linguistic tools besides additive particles (such as *make that two* constructions), see Appendix 2 for the complete model output.

Coefficient Estimates	posterior mean	Est.Error	I-95% CrI	u-95% CrI
Intercept	–3.52	0.73	–4.99	–2.11
Similarity	**2.00**	**0.39**	**1.23**	**2.78**
Turn Distance	**1.10**	**0.32**	**0.49**	**1.77**
Politeness	0.23	0.34	–0.42	0.90
Similarity*Turn Distance	0.06	0.32	–0.60	0.69
Similarity*Politeness	–0.02	0.33	–0.67	0.64
Turn Distance*Politeness	–0.18	0.31	–0.82	0.40
Similarity*Turn Distance*-Politeness	0.25	0.31	–0.34	0.87

Table 2: Population-level estimates of the logistic regression model in log-odds with the standard errors and 95% credible intervals. In the table the intercept (i.e. the grand mean) is listed first, then the estimates for Similarity, Turn Distance, Politeness and their interactions follow. Similarity is the change in log-odds for perfect similarity (–1 reduced, 1 perfect), Turn Distance is the change in log-odds for no turn (–1 three turns, 1 no turns), and Politeness is the change in log-odds for a neutral antecedent speaker (–1 impolite, 1 neutral). Slope coefficients whose 95% credible intervals do not include zero and are therefore treated as reliable effects are highlighted in bold.
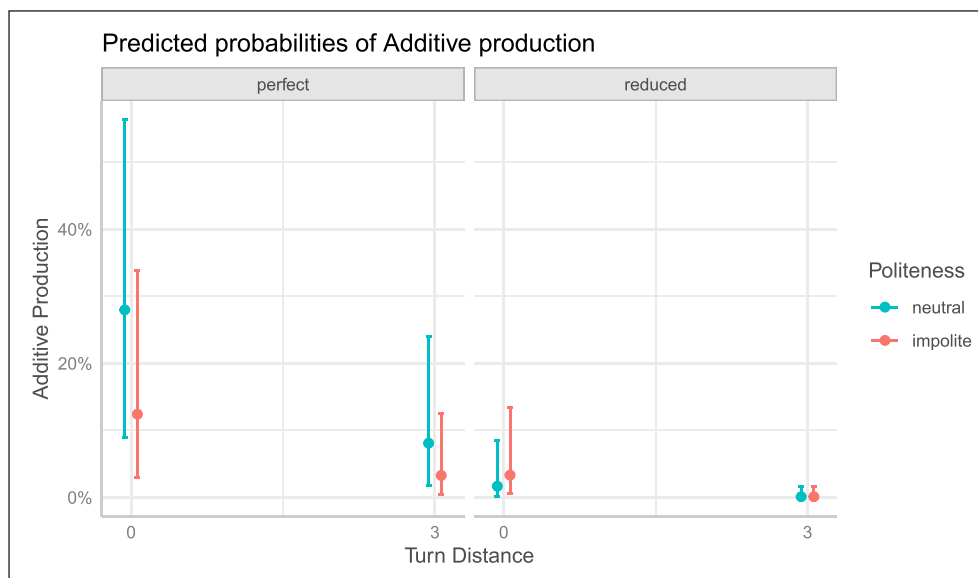

Figure 12: Predicted probabilities of producing additive particles given the model and the data for perfect (left) and reduced similarity (right) and 0 and 3 intervening turns. The figure contrasts the probability of producing additives when faced with a neutral antecedent speaker (blue) versus an impolite antecedent speaker (red).

The credible intervals of all interaction terms include zero which suggests that our three predictors did not interact in a systematic way. We compared the expected log predictive density of the full interaction model to models with reduced interaction terms, a main effects only model and a model that only included Similarity and Turn Distance as predictors. Model comparison was carried out via PSIS-LOO approximation (Pareto smoothed importance sampling leave-one-out approximation; Vehtari et al. (2015; 2017)). The approximation showed that a simple main effects model excluding any interaction terms had the highest predictive accuracy, followed by a model that included all three models as main effects and an interaction between Similarity and Politeness. However, the differences between the models were small, as shown in **Table 3**.

	elpd_diff	se_diff
Similarity + Turn Distance + Politeness	0	0
Sim * P + TD	–1.1	1.0
Sim * TD	–1.3	1.4
Sim + TD * P	–1.9	0.8
Sim * TD + P	–4.9	1.7
Sim * TD * P	–6.8	2.1

Table 3: Experiment I: the table shows the difference in the models' expected predictive accuracy (elpd_diff) and the standard error of the difference in elpd (se_diff) with the preferred model listed first.

3.1.7 Discussion

The results of experiment I suggest that speakers use additives more frequently for perfect similarity than reduced similarity and when the antecedent turn directly precedes the speakers' turn as opposed to when it does not. There was no reliable evidence that speakers' additive production is influenced by the antecedent speaker's politeness. The overall use of additives was rather low (approx. 17% and 51% in the neutral/perfect similarity/0 turn distance condition), which is lower than expected if one considers that additives should be used to block the exhaustivity implicature which should most strongly arise in the perfect similarity condition with zero intervening turns. Thus, at least for the condition with perfect similarity and zero intervening turns, participants should have used additives much more frequently, according to this account of exhaustivity.

3.2 Experiment II

The second experiment investigates the use of additives in a forced-choice paradigm in which participants can choose between the bare assertion (20), the assertion plus additive (21), an

incorrect response (22) and the option to type their own response ('other'). Note that Similarity is enforced in the sense that now participants are not only presented with more or less similar content but also formulations that are similar to the antecedent speaker's turn.

(20)   I'd like a glass of Chardonnay, please.

(21)   I'd like a glass of Chardonnay too, please.

(22)   I'd like a Martini, please.

(23)   Other: _____

We decided to focus on *too*, since *too* was after *same as* the most frequently produced additive. For this decision we considered the participants' overall production preference including ambiguous cases in which the additive did not unambiguously refer back to the antecedent utterance.[13] We initially were interested in presupposition triggers which is why we did not use *same as* as the pre-determined choice for experiment II. In order to account for participants wanting to use additives other than *too* or preferring to produce their own utterances in general, we included the participants' formulated utterances (option 'other') in our analysis. We applied the same coding as for experiment I.

The experimental design of experiment II allows us to circumvent issues of scope ambiguity by creating stimuli in which the additive more unambiguously refers back to the antecedent utterance. However, in restricting the participants' responses, experiment II more closely resembles acceptability judgement tasks which are frequently conducted to analyse the obligatoriness of additives: participants are shown an antecedent utterance and see the direct contrast between a following utterance with or without an additive. Thus, one could argue that this makes our claim about the novelty of our experimental design being a production study less appealing. However, we argue that experiment II contributes additional insights: since participants were explicitly offered to use a host utterance either containing or not containing an additive in the forced-choice paradigm, exhaustivity should be enhanced in experiment II as opposed to experiment I. We therefore expect additives to be chosen more frequently in experiment II than in experiment I. The aim of experiment II was furthermore to replicate the effects of Similarity and Turn Distance and to shed more light on possible interactions between the three factors under investigation by increasing the number of participants as well as the expected number of additive-containing continuations. In addition to investigating Similarity as a binary factor, we take an exploratory look at the distinction between partial similarity and no similarity.

---

[13] Frequencies: *same as* (N = 47), *too* (N = 28), *also* (N = 24), *as well* (N = 17), *another* (N = 2).

3.2.1 Participants

Participants (N = 141) were recruited from the crowd-sourcing platform Prolific. We recruited only those with an approval rate above 90. Participants were paid an average of £7.60/hour (the average duration of the experiment was 12 minutes). The age of the participants ranged from 18 to 75 years, with a mean of 36 years (median = 34, mode = 32). Three participants stated their preferred pronoun as they/them, 16 participants as he/him and 122 as she/her.

3.2.2 Design and Materials

We used the same experimental design as for experiment I, see Section 3.1.2. The only difference was that participants were asked to engage with their colleagues and a waiter by either choosing one of the three utterance choices or by typing their own response (choice 'other'), see (24–26). Crucially, we presented different utterance choices for partial similarity items (25) as opposed to no similarity items (26). This decision was based on the results of experiment I, which showed that participants still produced additives for partial similarity items as opposed to no similarity items. In these instances participants used additives in constructions such as (25b). We decided to use this formulation instead of 'I gave £15 and some chocolate, too', since we did not want to gloss over the fact that *too* can still felicitously be used for parts of the antecedent utterance in the reduced similarity condition. Being aware of an enforced difference between partial and no similarity we explored the factor Similarity as a three level factor in a second analysis.

(24) Perfect Similarity (antecedent speaker: glass of Chardonnay)
    a. I'd like a glass of Chardonnay, please.
    b. I'd like a glass of Chardonnay too, please.
    c. I'd like a Cosmopolitan, please.
    d. Other: _____

(25) Partial Similarity (antecedent speaker: £15 and wine)
    a. I gave £15 and some chocolate.
    b. I gave £15 too but bought him some chocolate instead.
    c. I gave £25 and whisky.
    d. Other: _____

(26) No Similarity (antecedent speaker: a pint of Heineken)
    a. I'd like a pint of Stella Artois, please.
    b. I'd like a pint of Stella Artois too, please.
    c. I'd like a Martini, please.
    d. Other: _____

Furthermore, while experiment II still targets content-based similarity, using pre-determined choices also contributed to an increase in syntactic similarity between antecedent and participant utterance. For example, for a short order such as in (24) the antecedent speaker's utterance was: *I'd like a glass of Chardonnay*, which is identical to the participants' turn not only in content but also in the structure of the sentence. Similarly for (25) the antecedent utterance was *I contributed £15 and got him wine*. The syntactic structure deviated more for items where we had to control for scope ambiguities such as in (27) with the corresponding antecedent order *I'd like the sticky toffee pudding and an espresso*. In these cases we used *for me* constructions to make it more salient that *too* does not only scope over espresso (with the reading *I order espresso as well as something else*) but both over espresso and sticky toffee pudding (i.e. *sticky toffee pudding and espresso for me as well as for someone else*), see Appendix 5 for a full set of the predetermined responses of experiment II.

(27)   Perfect Similarity (antecedent speaker:)
    a.   The sticky toffee pudding and an espresso for me, please.
    b.   The sticky toffee pudding and an espresso for me too, please.
    c.   The berry Pavlova and an Earl Grey tea for me, please.
    d.   Other: _____

Since syntactic-based similarity is something that may enhance the similarity manipulation in experiment II as opposed to experiment I, we will discuss what this meant for additive usage in our experiment to inform future research. However,we do not aim to disentangle content- and syntactic-based similarity at this point.

3.2.3 Procedure

The same procedure as in experiment I was followed: (1) informed consent, (2) experiment, (3) questionnaire. The only difference was that we added a question to the questionnaire about whether they thought that the utterance choices reflected their way of speaking. Participants could respond by adjusting a slider from 0 ('Definitely not') to 100 ('Definitely yes'). The experiment consisted of four blocks that were presented in the same order to all participants.

3.2.4 Predictions

Similar to experiment I, following the Obligatory Implicatures account, participants are predicted to most frequently use additives for highly similar and recent antecedents, as such a context would yield the strongest 'contrast' or exhaustivity effect. If additive use reflects social goals, we furthermore predict that participants will choose additives less in the impolite condition than in the neutral condition. We also further explored the interactions between these three factors. This

part was again exploratory, since the use of additives in combination with other predictors in experiment I was too sparse to inform any concrete predictions here.

3.2.5 Analysis

We analysed the data as in experiment I, by fitting a Bayesian logistic regression model with varying by-item and by-participant intercepts and slopes. The experimental factors Similarity, Turn Distance and Politeness and their interactions were included to predict the probability of using *too*. All factors were sum-coded as in the analysis of experiment I. The model included varying intercepts and slopes for participants and items, assuming that the effects of Similarity, Turn Distance and their interaction vary between participants, and the effects of Similarity, Turn Distance, Politeness and their interactions vary between items. The choice of priors and the sampling process were the same as for the experiment I analysis; see Section 3.1.5 for a detailed description.[14]

3.2.6 Results

Again, participants were above 96% accurate for the attention checks, which suggests that they paid attention during the experiment. Participants also seemed to have understood the task (mean = 85, sd = 19.6, mode = 100, median = 94) and rated the predetermined choices as mostly natural (mean = 79, sd = 20.88, mode = 100, median = 82). Overall, participants used additive formulations throughout the experiment and more often than in experiment I: 32.3% of the participants' utterances contained additives. 'Order' items again elicited more uses of additives (35%) than 'talk' items (29%). The option 'other' was chosen 73 times (out of 1128). After inspection of the 'other' responses, 9 instances were included that contained additives of which *too* was used 7 times, *as well* and *also* 1 time, respectively. There are two observations for which participants used other linguistic formulations to refer back to the antecedent utterance such as *I'm with Omar on this, I love Korean food*. 'Other' responses containing *too* often entailed further interaction with either a neutral (28a) or impolite (28b–28c) antecedent speaker.

(28)   a.   Great minds think alike Robert. I'll have the sticky toffe pudding and an expresso too.
       b.   i'd like a pin of heineken too please. Robert you should'nt speak to the waiter in that way it is disrespectful
       c.   To waiter: The Breaded brie and king prawn linguine for me too, please. To Omar (once the waiter has left): Omar, are you okay? I think you might have been a bit abrupt with the waiter and that's not like you.

---

[14] In line with our preregistration, we ran two models with more uninformative/wider priors yielding similar posteriors; see Appendix 3 for details. Overall, the models with wider prior distributions yielded higher posterior means but also introduced more uncertainty, specifically for the posterior distributions of the interaction coefficients.

Figure 13 illustrates that, similarly to experiment I, additives were most frequently used when host utterance and antecedent utterance were similar and when there were zero intervening turns. In contrast to experiment I, participants used additives in the reduced similarity/three intervening turns condition. Moreover, the frequency of additive use seems to be less when speaking after an impolite speaker than after a neutral speaker both for perfect similarity conditions and in the reduced similarity with three intervening turns. The opposite seems to be the case for the reduced similarity condition with zero intervening turns.
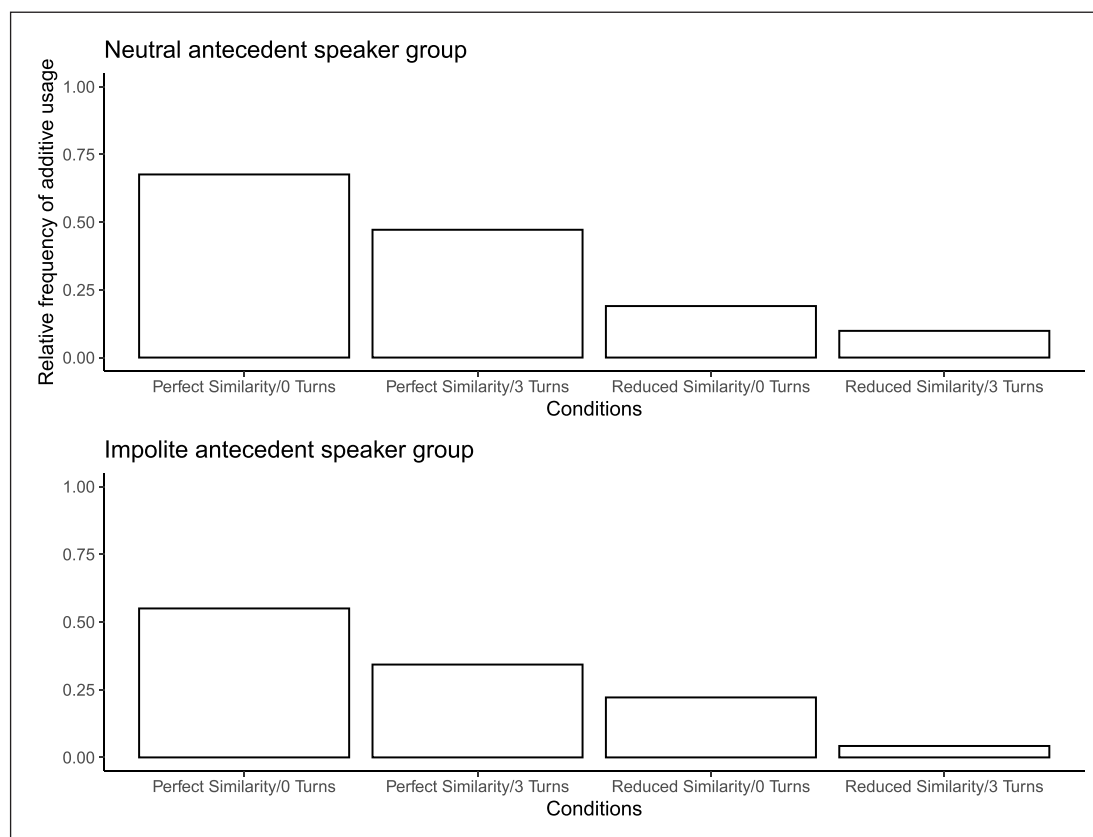

Figure 13: Experiment II: relative frequency of choosing additives in each of the Similarity × Turn Distance conditions, contrasting participants confronted with a neutral (top) versus impolite (bottom) antecedent speaker.

These observations are supported by the outcome of the analysis, see **Table 4**. Holding everything else constant, perfect similarity (as opposed to reduced similarity) between antecedent utterance and host utterance led to an increase in log-odds of additive production ($\hat{\beta} = 1.32$, CrI:[0.53, 2.02]), as did the antecedent utterance directly preceding the host utterance ($\hat{\beta} = 0.72$, CrI:[0.29, 1.16]), and speaking after a neutral antecedent speaker ($\hat{\beta} = 0.34$, CrI:[0.03,

0.69]). Overall, Similarity seems to affect the production of additives to a greater extent than Turn Distance and Politeness. The credible intervals of all effects do not include zero.

Coefficient Estimates	posterior mean	Est.Error	I-95% CrI	u-95% CrI
Intercept	–1.42	0.34	–2.09	–0.76
Similarity	**1.32**	**0.37**	**0.53**	**2.02**
Turn Distance	**0.72**	**0.21**	**0.29**	**1.16**
Politeness	**0.34**	**0.17**	**0.03**	**0.69**
Similarity*Turn Distance	–0.14	0.18	–0.50	0.22
Similarity*Politeness	0.02	0.15	–0.27	0.32
Turn Distance*Politeness	–0.23	0.17	–0.59	0.10
Similarity*Turn Distance*Politeness	0.23	0.13	–0.03	0.49

Table 4: Population-level estimates of the logistic regression model in log-odds with the standard errors and 95% credible intervals. In the table the intercept (i.e. grand mean) is listed first, then the estimates for Similarity, Turn Distance, Politeness and their interactions follow. Similarity is the change in log-odds for perfect similarity (–1 reduced, 1 perfect), Turn Distance is the change in log-odds for no turn (–1 three turns, 1 no turns), and Politeness is the change in log-odds for a neutral antecedent speaker (–1 impolite, 1 neutral). Slope coefficients whose 95% credible intervals do not include zero and are therefore treated as reliable effects are highlighted in bold.

The credible intervals of all interaction terms include zero. The most reliable interaction coefficient is the three way interaction coefficient ($\hat{\beta} = 0.23$, CrI:[–0.03, 0.49]). In fact, the coefficient is 96% likely to be greater than zero predicting an increase in log-odds for perfect similarity, zero intervening turns and a neutral antecedent speaker. A model for which additive usage included all linguistic tools (besides additives) that were used to refer back to the antecedent utterance revealed almost identical results, see Appendix 2. We compared the expected log predictive density of the full interaction model to models with reduced interaction terms, a main effects only model and a model that only included Similarity and Turn Distance and their interaction as predictors. Model comparison was carried out via PSIS-LOO approximation (Vehtari et al. 2015; 2017)) which disclosed that the model including all predictors and their interactions had the highest predictive accuracy, followed by a model that included Similarity and Turn Distance and their interaction plus Politeness as main effect. The model only including Similarity and Turn Distance had the lowest predictive accuracy, as shown in **Table 5**. This suggests even more that the interactions between the three predictors play a crucial role in the way additives are produced.

	elpd_diff	se_diff
Similarity * Turn Distance * Politeness	**0.0**	**0.0**
Sim * TD + P	–0.2	1.8
Sim + TD + P	–0.4	4.8
Sim + TD * P	–0.6	4.4
Sim * P + TD	–1.9	4.5
Sim * TD	–2.3	4.2

Table 5: Experiment II: the table shows the difference in the models' expected predictive accuracy (elpd_diff) and the standard error of the difference in elpd (se_diff) with the preferred model listed first.

In **Figure 14**, we illustrated the predicted probabilities for speakers to utter additives given our data and model. Speakers are predicted to produce additives more frequently for perfect similarity and when no turns intervene between their utterance and the antecedent utterance and least frequently for reduced similarity when three turns intervene. There seems to be a more pronounced difference between neutral (blue) and impolite conditions (red) than in experiment I for the perfect similarity conditions such that speakers are predicted to produce additives more frequently when talking to a neutral antecedent speaker as opposed to an impolite antecedent speaker.
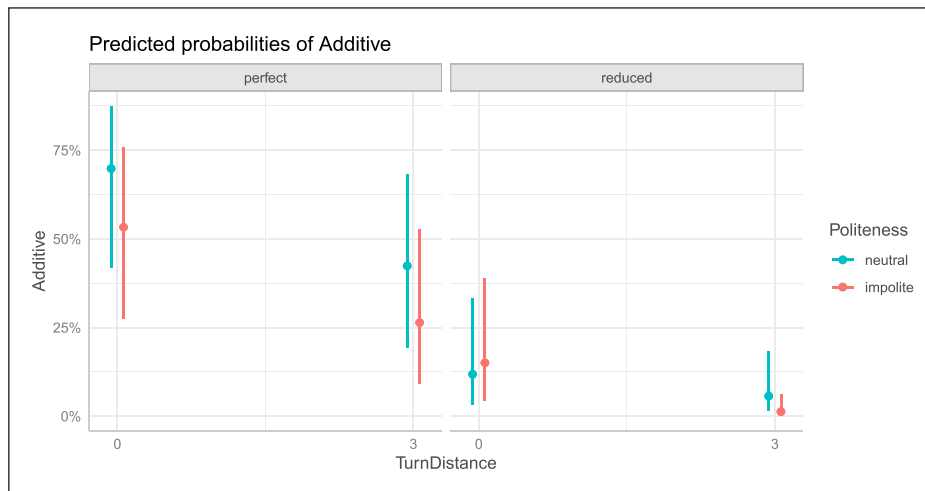

Figure 14: Predicted probabilities of producing additive particles given the model and the data for perfect (left) and reduced similarity (right) and 0 and 3 intervening turns. The figure contrasts the probability of producing additives when faced with a neutral antecedent speaker (blue) with an impolite antecedent speaker (red).

As was previously mentioned, using a forced-choice task, we can no longer speak about solely content-based similarity, since presenting participants with formulations also introduces some degree of syntactic similarity. Looking at the relative frequencies of additives in the perfect similarity condition, it seems that syntactic similarity is taken into account by participants. For example, for Order Item 1, see (24), for which participants were asked to order a Chardonnay using the same formulation as the antecedent speaker, participants used additives in 61% of the cases. In contrast, for Order Item 4, see (27), for which the formulation of the pre-determined choices deviated slightly from the antecedent utterance, participants used additives only 47% of the time. We have to leave it for future research to look into the roles that content- versus syntactic-similarity play in additive production.

Again, the focus of this study was primarily to investigate the difference between environments in which additive usage is obligatory (perfect similarity) versus environments in which it is not (reduced/no similarity). This distinction is however coarse, and we looked into differences within the group of non-obligatory environments. As was mentioned above, we ran a further analysis in which we treat the factor Similarity as a three level predictor (perfect similarity, partial similarity, no similarity). By doing so, we are able to look at differences between environments in which the production of additives is obligatory (perfect similarity), optional (partial similarity) or not needed (no similarity) according to the Obligatory Implicature account. We re-ran the same model as before only this time with Similarity as a three level factor which was sum-coded, as shown in **Table 6**.

(29) Partial Similarity (antecedent speaker: £15 and wine)
   a. I gave £15 and some chocolate.
   b. I gave £15 too but bought him some chocolate instead.
   c. I gave £25 and whisky.

(30) No Similarity (antecedent speaker: a pint of Heineken)
   a. I'd like a pint of Stella Artois, please.
   b. I'd like a pint of Stella Artois too, please.
   c. I'd like a Martini, please.

Recall that for previous analyses, the category 'reduced similarity items' contained partial similarity and no similarity items. Participants had the option to choose between three utterances: an utterance with *too*, an utterance without *too*, and an incorrect response, as shown in (29–30). Partial similarity and no similarity differ from each other in that *too* can still be felicitously used in the partial condition but not in the no similarity condition.

The output suggests that, averaging over the other effects, perfect similarity led to an increase in log-odds of additive production ($\hat{\beta}$ = 1.91, CrI:[1.03, 2.69]), and in contrast, dissimilarity led to a decrease in log-odds of additive production ($\hat{\beta}$ = –2.07, CrI:[–2.96, –1.03]), as shown in **Table 7**. Partial similarity led to an increase in additive production but less so than perfect similarity (approximate increase: 0.16 in log odds). The 95% credible interval for Politeness includes zero and is no longer reliable.

	Similarity1	**Similarity2**
perfect	1	0
dissimilar	0	1
partial	–1	–1

Table 6: Experiment II: Contrast coding for the three-level Similarity factor.

Coefficient Estimates	**posterior mean**	**Est.Error**	**l-95% CrI**	**u-95% CrI**
Intercept	–2.07	0.35	–2.75	–1.38
Similarity1	**1.91**	**0.42**	**1.03**	**2.69**
Similarity2	**–2.07**	**0.48**	**–2.96**	**–1.03**
Turn Distance	**0.74**	**0.26**	**0.23**	**1.28**
Politeness	0.34	0.22	–0.09	0.79
Similarity1*Turn Distance	–0.17	0.25	–0.67	0.33
Similarity2*Turn Distance	–0.31	0.36	–1.03	0.40
Similarity1*Politeness	0.01	0.24	–0.46	0.48
Similarity2*Politeness	0.07	0.35	–0.61	0.80
TurnDistance1*Politeness	–0.37	0.23	–0.84	0.08
Similarity1*Turn Distance*Politeness	0.36	0.23	–0.09	0.81
Similarity2*Turn Distance*Politeness	–0.37	0.36	–1.10	0.32

Table 7: Experiment II of Similarity as a three-level factor. Population-level estimates of the logistic regression model in log-odds with the standard errors and 95% credible intervals. In the table the intercept is listed first, then the estimates for Similarity, Turn Distance, Politeness and their interactions follow. Slope coefficients whose 95% credible intervals do not include zero and are therefore treated as reliable effects are highlighted in bold. Similarity1 is the difference between intercept and the perfect similarity condition, Similarity2 is the difference between intercept and the dissimilar condition. The difference between the intercept and the partial similarity condition can be calculated as follows: $1.91 * (-1) + (-2.06) * (-1) = 0.15$.

3.2.7 Discussion

We were able to replicate parts of the findings of experiment I: speakers use additives more for perfect similarity than reduced similarity and when the antecedent turn directly precedes the speakers' turn as opposed to when it does not. Overall, additive production was again lower than assumed in potentially obligatory environments (68% in the neutral/perfect sim./0 TD condition). Furthermore, in contrast to experiment II we found evidence that speakers more frequently use additives when speaking after a neutral antecedent speaker and more frequently refrain from producing additives when the antecedent speaker is impolite.

By investigating Similarity as a three level factor we discovered differences between the way partial similarity and dissimilarity influenced participants' additive use. With regards to the absent effect of politeness, we conjecture that the words *but* and *instead* in (29b) for partial similarity may already express a dissociation between the antecedent speaker and the participant and their contributions so that dropping *too* is no longer necessary. However this explanation is speculative and has to be investigated further.

4 General Discussion

4.1 Summary

In line with the comprehension results of Kim (2015) and Chen & Husband (2018), our results suggest that distance influences the extent to which additives are obligatory. Extending prior findings on distance that use measures of textual distance, we found that speakers use additives more frequently as the distance in dialogue turns decreases: They use more additives when their conversational turn immediately precedes the antecedent turn. These results additionally shed more light on how immediately a suitable antecedent has to precede a speaker's utterance in order for additives to be used felicitously. While it is true that immediately preceding antecedents did elicit a higher rate of additives, speakers still occasionally used additives for an antecedent separated by three intervening turns.

With regards to Similarity, our results are in line with what has been found for comprehension (Amsili et al. 2016): additives were used more in the perfect similarity condition. Taking Turn Distance and Similarity together, additives were most frequently produced when the antecedent directly preceded the host and both were highly similar – which should be a context with the strongest exhaustivity implicature. However, even in this context, the probability of speakers producing additives was rather low: on average only 68% (in experiment II).

Regarding the factor Politeness, in experiment II we found that speakers use additives more frequently when speaking after a neutral antecedent speaker and omit additives more frequently when speaking after an impolite antecedent speaker. This may indicate that omitting additives in contexts that give rise to an exhaustivity implicature can be a strategic way to diverge from

other interlocutors. The contrast or contradiction that arises by not blocking the exhaustivity implicature may highlight a contrast on the social level. One explanation for why Politeness was overall the weakest and most inconsistent factor in our study could be due to the way we implemented the politeness factor. As suggested by an anonymous reviewer, by adding impolite material as a subordinate clause or a separate clause preceding for example the order, the additive might simply be avoided because the participant (henceforth Speaker B) does not want the subordinate clause ('Although you served me white wine last time') to be part of the presupposition (not because it is impolite but because it is not true). Consequently, a response such as 'I'd like Chardonnay, too.' could have been ambiguous between two readings: speaker B is committing themselves or not to the subordinate clause.

For a start, we think that there are several reasons to doubt that the preceding 'although' clause can be part of this presupposition. If this were to be available, utterances such as 'Me too' – uttered by Speaker B after Speaker A's utterance 'Although you served me warm white wine last time, I'd like a glass of Chardonnay' – would in principle also be ambiguous between readings in which speaker B presupposed that speaker A had been served warm white wine last time and one in which speaker B presupposed that speaker B had been served warm white wine last time, which we think is not evident. Moreover, if A were to utter 'Although speaker B thinks I've given up drinking, I'd like a glass of Chardonnay', for B to utter 'Me too' would be paradoxical, which again conflicts with our intuitions. If our judgment on this point is incorrect and the 'although' clause can be part of the presupposition of *too* in our materials, then we agree that speaker B may refrain from using *too* if they do not endorse the content of the 'although' clause. However, in the 'warm white wine' example, speaker A is reporting a (claimed) fact in the 'although' clause, and it is not entirely clear why speaker B should refuse to endorse this fact other than as an expression of social divergence.

Overall, the effect of Politeness on the production of additives needs further investigation. While the different outcomes of experiments I and II may be due to sparse data in experiment I, replication studies are necessary not only to explore Politeness further but also to shed more light on the possible effect that open versus forced-choice tasks have on production data. Furthermore, we only investigated one potential reason for speakers to diverge from an antecedent speaker, namely when they behave impolitely. There may be other reasons with a potentially greater influence on whether speakers diverge or not such as when the antecedent speaker expounds controversial, false or immoral opinions. However, such reasons are difficult to test in an experimental setting due to ethical considerations.

4.2 Low additive frequency

For both experiments, the additive usage in obligatory environments was rather low even when the antecedent utterance directly preceded the host utterance and both were highly similar in

content (experiment I) or content and formulation (experiment II). The low frequency might not be as surprising for experiment I, for which we used a free-text task. Here, participants could use other linguistic means to block the exhaustivity implicature, e.g. phrasing their utterance differently from the antecedent speaker utterance or using lexical items such as *and* at the beginning of their utterance to flag a similar addition to the dialogue. The relatively low additive frequency in obligatory environments in experiment II contrasts however more directly with predictions of the Obligatory Implicature account. Recall that in experiment II participants were directly presented with utterance choices and offered the use of additives. Thus, in obligatory environments participants saw the antecedent utterance followed by two syntactically and semantically identical response options either containing or not containing *too*.

We want to briefly address possible reasons as to why additive production may have been relatively low, which might reconcile our findings with the predictions of the Obligatory Implicature account. One potential reason could be the dialogue setting. Bade (2016) claimed that exhaustivity may be reduced in settings in which more than two interlocutors are involved, based on two studies investigating *auch* (German *too*). More specifically, in study 1 Bade (2016) investigated the acceptability of sentences containing versus not containing *too* in a dialogue setting, see (31).

(31)   a.   A: Peter had a party at his house last night.
       b.   B: Cool./Who came to the party?
       c.   A: Mary came to the party.
       d.   C: Julia came to the party./Julia came to the party, too.

In contrast, in study 2 Bade (2016) tested the acceptability of continuations containing or not containing *too* following a context sentence, see (32), a setting which could mirror a narrative or the contribution of a single speaker.

(32)   a.   Context:
            Rita came to work, made coffee, greeted Stefan and Sabine and sat down at
            her desk.

       b.   Continuation:
            She greeted Lisa.
            She greeted Lisa, too.

Bade found that sentences not containing *too* were judged less acceptable than sentences containing *too* in both studies (study 1: mean = 3,62 on a 5-point scale, study 2: mean = 2,1 on a 7-point scale), but most crucially that sentences without *too* were judged less acceptable in study 2 than study 1. Besides arguing that contradictions between conversational partners are more acceptable than contradictions within a single discourse, exhaustivity may be reduced in

dialogue settings because multiple contributions of more than one speaker are expected to be possible partial answers to the QUD.

If the distinction between dialogue contexts and single speaker contribution contexts holds up for production, one would expect higher additive production in studies looking at additive production in narration produced by one speaker – such as in Eckhardt & Fränkel's 2012 study – than in studies looking at additive production in dialogue like settings – such as in our study. Eckhardt & Fränkel (2012) investigated the production of *auch* and *wieder* (German *too* and *again*) by asking participants to describe four pictures of a comic strip one after another. The pictures illustrated two protagonists that would occasionally perform similar actions at different times such as brushing their teeth. Participants were either instructed to describe the comic strip as if they were telling a children's story or they were instructed to describe what happened at specific times, similar to a protocol. Focusing on the relevant narration condition, participants seemingly produced additives and iteratives on average 55% of the time which is similar to the additive frequency of experiment I (~51% in the neutral antecedent speaker/perfect similarity/0 turn distance condition) and lower than the additive frequency of experiment II (~67% in the neutral/perfect sim./0 TD condition). Thus, it remains open whether additive insertion is more obligatory in narration or for the contribution of a single speaker versus in dialogue settings. Since the experimental designs of Eckhardt & Fränkel's 2012 study and the current study differ to a great extent, we leave it to future research to investigate differences in additive production for different contexts.

Another point that is worth highlighting is the role of focus in our experiment. Pragmatically, focused material generally addresses the explicit or implicit *Question Under Discussion* (QUD). Following Roberts (1996), the QUD corresponds to the current discourse topic. Focus may also be introduced by prosodic stress such that the kind of exhaustivity implicature that arises may depend on which constituent of the host sentence is prosodically stressed and thereby focused (e.g. Rooth 1985). For example, in (33a) *too* associates with the prosodically stressed 'I' making the speaker the exhaustive answer to the QUD *Who noted down that she's due to start in January?* In contrast, in (33b), 'she' is prosodically marked making 'her' the exhaustive answer to the QUD *Who's due to start in January?*

(33) a. [I]$_F$ noted down that she's due to start in January, too.
      b. I noted down that [she]$_F$'s due to start in January, too.

We did not provide participants with audio recordings in our experiments. In experiment II with the pre-determined choices, participants therefore would have had to assign any implicit prosody individually, e.g. participant1 may have placed focus on 'I' (33a) and participant2 on 'she' (33b). Giving rise to different implicatures may have led to different assessments by participants of whether there was a suitable antecedent in the preceding discourse which would in turn induce some variation in additive use between participants.

While the placement of individual focus is possible, we argue that by presenting the overall discourse topic, we reduced such variation to some extent. Recall that in our study we set the discourse topic for each experimental item headlining each picture of a given conversation, such as 'Ordering drinks' or 'Discussion about when the new colleague is starting' but not an explicit QUD. This might have generated a variety of corresponding QUDs such as *Who orders what?* In general, QUDs that permit one correct answer lead to a stronger exhaustivity effect than QUDs that permit partial answers or multiple contributions. The majority of possible QUDs in our study do permit partial answers; in fact, partial answers are expected. For example, the possibly generated QUDs *Who orders what?* for 'order' items permits partial answers and multiple contributions. In contrast, for the 'talk' items (items 7 and 8) possible implicit QUDs may be associated with one correct answer. Both are about the discussion of when a specific event takes place, which gives rise to implicit QUDs such as *When will the new colleague start?* (item 7) and *At what time is next week's meeting?* (item 8). Just looking at the relative frequencies in **Figure 15**, however, there is no indication of a trend in the expected direction, i.e. more frequent additive production for items 7 and 8 than for the remaining conversations. In fact, item 7 seems to be an outlier for other reasons, since additive production was lower than for other items.[15]
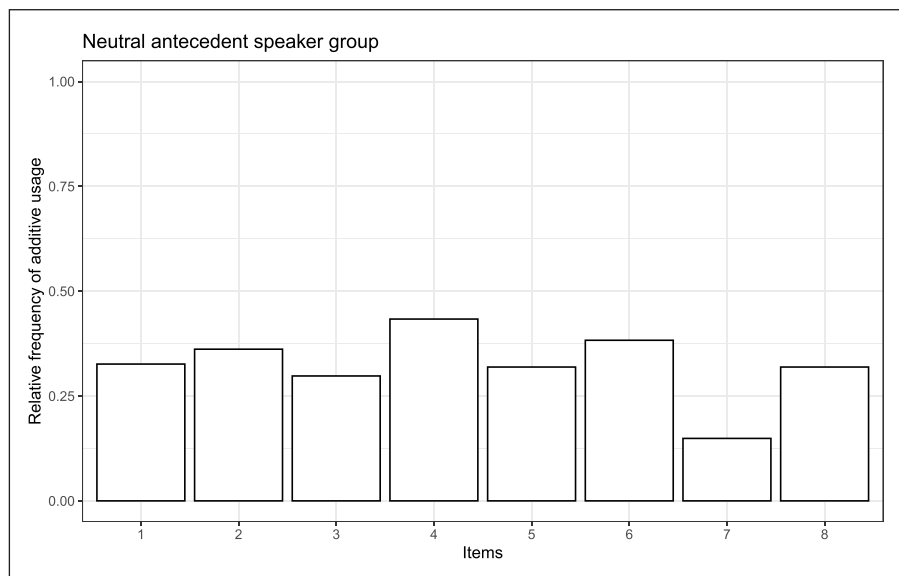

Figure 15: Relative Frequencies (y-axis in proportion) of additive production for experiment II by item (x-axis).

---

[15] Note that one could argue that item 7 is an outlier in experiment II, since the formulations of the antecedent and participant utterances deviated here the most (antecedent: *She'll start next year in January.*, participant: *I noted down that she's due to start in January.*), and thereby reducing syntactic similarity. However, even in experiment I, item 7 was an outlier. In fact, for experiment I, item 7 was the only item for which participants did not use any additives in any of the conditions.

On a related note, multiple *Wh*-questions such as *Who orders what?* may give rise to contrastive topic and focus contours (c.f. Büring 2016; Tomioka 2009). More specifically, for a question such as (34a) *I* would be marked as a contrastive topic and *glass of Chardonnay* as focus. The contrastive topic + focus contour leads to the following reading: whereas the focus excludes all possible alternatives (there is nothing else that the speaker would like to order other than a glass of Chardonnay), contrastive topic implies that other speakers might want to order a glass of Chardonnay. Such a reading would make the insertion of additives not necessarily obligatory and could explain the relatively low additive production.

(34) a. Who orders what?
      b. I_{CT} would like a [glass of Chardonnay]$_F$.

While we can only speculate about why additive production was rather low, this discussion opens up exciting possibilities for future research, e.g. investigating the effect of exhaustivity on additive production in dialogue-like structures by manipulating the number of speakers and QUD-type.

4.3 A speaker model of additive production

The following section aims to introduce a possible way to model the production of additives in dialogue-like structures by emphasising the importance of their (i) anaphoric component and (ii) discourse managing properties. Regarding (i), the processing or comprehension of additives may involve resolving the anaphoric dependency between host and a pre-mentioned antecedent. From the production perspective, the anaphoric component of additives makes the presence of a suitable antecedent necessary and involves keeping track of what material has previously been mentioned. Keeping track of what has been previously mentioned should be effortless when the antecedent directly precedes the speaker's utterance. However, producing additives may come with a working memory cost, such that keeping track of what information has already been mentioned may be more difficult with increasing conversational turns/time. Note that this may be due to memory decay and/or interference of contributions/sentences between antecedent and host (see Chen & Husband (2018) for an investigation of the processing of additives in terms of cue-based retrieval). Thus, not producing additives may happen unintentionally and may happen more if the turn with the suitable antecedent happened longer ago. The observed effect of Turn Distance seems to corroborate this: speakers more frequently used additives when their turn immediately preceded the antecedent turn.

Independently of the number of turns intervening antecedent and host, based on Obligatory Implicatures, we can assume that further contextual factors (e.g. Similarity, QUD-Type, sentence structure) determine the strength of the exhaustivity implicature and thereby the pressure on speakers to produce additives. However, we argue that (ii) the discourse managing properties

should be taken into account to fully capture additive production. By discourse managing properties of additives we mean that additives can be used by speakers to signal ontological control (i.e. the awareness of how content is organised in the discourse). The idea of ontological control is based on Eckhardt & Fränkel (2012)'s study which found that interlocutors more frequently produced additives when writing a story than when writing a protocol. Eckhardt & Fränkel (2012) claimed that the obligatoriness of additives depends to some extent on the contextual requirements for the speaker to signal control over ontology – i.e. speakers are aware of the way the content of different utterances is related and structured in the discourse. Omitting additives in textual environments in which the requirements to signal control over ontology are high, e.g. in narratives, is claimed to lead to utterances which are detached from the preceding discourse and the undertone that the speaker did not plan their narrative beforehand.

One could assume that spontaneous speech does not require the same standard of signalling control of ontology as is necessary for narratives. Instead, engaging explicitly in discourse management in conversation may be seen as an occasional, cooperative effort of interlocutors. For example, to make it easier for the waiter who is noting down the orders, speakers use additives to signal that something similar has been ordered before. Seeing additives as discourse managing tools may explain why additives were not always used in contexts that should give rise to a relatively strong exhaustivity implicature (i.e. when a highly similar antecedent directly preceded the host): speakers might not always deem it necessary to explicitly engage in discourse management.

Regarding the factor Politeness, the use of additives as a cooperative attempt to participate in discourse managing may be perceived as a convergence move that happens by default. Disengaging from this process may be due to a speaker's attempt to diverge from the antecedent speaker. Thus both Turn Distance and Politeness may account for instances when speakers choose not to produce additives. Whereas Turn Distance may explain unintentional omission of additives, Politeness may explain intentional omission of additives even with contexts usually associated with giving rise to an exhaustivity implicature.

A few responses in experiments I and II suggest that, for some participants, discourse management had priority over diverging from the antecedent speaker by omitting additives: after having formulated an utterance containing an additive, some participants continued to express their disagreement with the antecedent speaker's behaviour in a second sentence, as seen in (28). In this way participants remained cooperative towards the waiter and the other interlocutors. This way of dealing with an impolite speaker may be one of the reasons for why we did not find an effect of Politeness in experiment I. Since participants formulated their own utterances in experiment I, they were possibly more likely to follow this strategy than participants in experiment II who would most likely choose pre-formulated utterances.

4.4 Limitations and open questions

Potential limitations of our study include the experimental design and sample size. As previously mentioned, we used only a small number of experimental items (eight), and the observed effects could be due to the unique characteristics of these items and may not generalise to other conversational settings and topics. We also used a relatively low number of experimental trials: each participant saw each condition only twice, whereas typically psycholinguistic studies use around four trials per participant per condition. We chose this low number of trials to create a realistic interaction scenario (comprising one coherent conversation) with the aim of gathering data representative of actual language use. We would also argue that the use of a low number of trials increases participant engagement and limits the opportunity for the participants to develop task strategies, which would be detrimental to statistical power (e.g. Arehalli & Wittenberg 2021; Brehm et al. 2021). Recent single-trial experiments (e.g. von der Malsburg et al. 2020; Laurinavichyute & von der Malsburg 2022; 2024; Winter & Duffy 2020; Thibodeau & Boroditsky 2024; Hassemer & Winter 2018) or experiments with low number of trials (e.g. Rohde et al. 2021) combat this issue. However, it must be acknowledged that recent single-trial experiments have typically used appreciably more participants ($\sim$25000 (von der Malsburg et al. 2020), $\sim$5500 (Laurinavichyute & von der Malsburg 2024), $\sim$200–300 (Winter & Duffy 2020), $\sim$185–485 (Thibodeau & Boroditsky 2024), $\sim$185–345 (Hassemer & Winter 2018) to $\sim$170 participants (Rohde et al. 2021)). Further research is needed to establish optimal sample sizes for experiments with few items, and hence our results should be interpreted with caution in this regard.

5 Conclusion

In this paper we investigated the circumstances under which additives are used. The findings of experiments I and II suggest that the production of additives is influenced by discourse factors: our findings show that increased similarity and close proximity between antecedent and host sentences favour increased use of additives (but do not make their use obligatory). Furthermore, our findings suggest that omitting additives (even with a highly salient antecedent present) may be understood as a divergence move away from the antecedent speaker. However, this has to be investigated further. We introduced a model in which speakers use additives as a discourse managing tool to organise the discourse and maintain discourse coherence. Discourse managing itself could be understood as a cooperative process by which speakers converge with their interlocutors.

Data accessibility statement

Data and materials for all experiments can be found here: https://osf.io/az2uf/.

Supplementary files

Supplementary information for this article can be found in three Appendix files at DOI: 10.17605/OSF.IO/AZ2UF.

This Appendix contains the following materials:

- Appendix 1: sample experimental items for experiment I
- Appendix 2: additional analysis (including all linguistic tools referring back to antecedent utterance) for experiment I
- Appendix 3: additional analyses (uninformative priors, order effects) for experiments I and II
- Appendix 4: norming study for politeness manipulation
- Appendix 5: pre-determined responses of experiment II

Ethics and consent

This research has been approved by the PPLS Research Ethics Committee (School of Philosophy, Psychology and Language Sciences, University of Edinburgh).

Competing interests

The authors have no competing interests to declare.

References

Amsili, Pascal & Beyssade, Claire. 2010. Obligatory presuppositions in discourse. In Kühnlein, Peter & Benz, Anton & Sidner, Candace L. (eds.), *Constraints in discourse 2, of pragmatics & beyond*, 105–123. Benjamins Publishers. DOI: https://doi.org/10.1075/pbns.194.06ams

Amsili, Pascal & Ellsiepen, Emilia & Winterstein, Grégoire. 2016. Optionality in the use of *too*: The role of reduction and similarity. *Revista da ABRALIN* 15(1). 229–252. DOI: https://doi.org/10.5380/rabl.v1i15

Aravind, Athulya & Hackl, Martin. 2017. Against a unified treatment of obligatory presupposition trigger effects. *Semantics and Linguistic Theory* 27. 173–190. DOI: https://doi.org/10.3765/salt.v27i0.4141

Arehalli, Suhas & Wittenberg, Eva. 2021. Experimental filler design influences errorcorrection rates in a word restoration paradigm. *Linguistics Vanguard* 7(1). DOI: https://doi.org/10.1515/lingvan-2020-0052

Bade, Nadine. 2016. *Obligatory presupposition triggers in discourse – empirical foundations of the theories maximize presupposition and obligatory implicatures*: University of Tübingen dissertation. DOI: https://doi.org/10.15496/publikation-10551

Bade, Nadine & Renans, Agata. 2021. A cross-linguistic view on the obligatory insertion of additive particles – maximize presupposition vs. obligatory implicatures. *Glossa: A Journal of General Linguistics* 6(1). DOI: https://doi.org/10.5334/gjgl.727

Brehm, Laurel & Jackson, Carrie N. & Miller, Karen L. 2021. Probabilistic online processing of sentence anomalies. *Language, Cognition and Neuroscience* 36(8). 959–983. DOI: https://doi.org/10.1080/23273798.2021.1900579

Brown, Penelope & Levinson, Stephen. 1987. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511813085

Büring, Daniel. 2016. (Contrastive) topic. In Féry, Caroline & Ishihara, Shin (eds.), *Handbook of information structure*, 64–85. Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199642670.013.002

Bürkner, Paul-Christian. 2018. Advanced Bayesian multilevel modeling with the R package brms. *The R Journal* 10(1). 395–411. DOI: https://doi.org/10.32614/RJ-2018-017

Chen, Sherry Yong & Husband, Matthew E. 2018. Comprehending anaphoric presuppositions involves memory retrieval *too*. In *Proceedings of the Linguistic Society of America*, vol. 3. 1–11. DOI: https://doi.org/10.3765/plsa.v3i1.4288

Eckhardt, Regine & Fränkel, Manuela. 2012. Particles, maximize presupposition and discourse management. *Lingua* 122(15). 1801–1818. https://www.sciencedirect.com/science/article/pii/S0024384112001726. DOI: https://doi.org/10.1016/j.lingua.2012.08.014

Gallois, Cynthia & Ogay, Tania & Giles, Howard. 2005. Communication accommodation theory: A look back and a look ahead. In Gudykunst, B. (ed.), *Theorizing about intercultural communication*, 121–148. Thousand Oaks: Sage.

Giles, Howard. 1973. Accent mobility: A model and some data. *Anthropological Linguistics* 15. 87–105. https://www.jstor.org/stable/30029508.

Giles, Howard. 1977. Social psychology and applied linguistics: Towards an integrative approach. *ITL* 35(1). 27–42. DOI: https://doi.org/10.1075/itl.35.02gil

Giles, Howard. 1979. A new theory of the dynamics of speech. *Diogenes* 27(106). 110–127. DOI: https://doi.org/10.1177/039219217902710607

Giles, Howard & Hewstone, Miles. 1982. Cognitive structures, speech and social situations: Two integrative models. *Language Sciences* 4(2). 187–219. https://www.sciencedirect.com/science/article/pii/S0388000182800053. DOI: https://doi.org/10.1016/S0388-0001(82)80005-3

Hassemer, Julius & Winter, Bodo. 2018. Decoding gestural iconicity. *Cognitive Science* 42(8). 3034–3049. DOI: https://doi.org/10.1111/cogs.12680

Hoffman, Matthew D. & Gelman, Andrew. 2014. The No-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* 15(1). 1593–1623. https://dl.acm.org/doi/10.5555/2627435.2638586.

Kaplan, Jeff. 1984. Obligatory *too* in English. *Language* 60(3). 510–518. DOI: https://doi.org/10.2307/413989

Kim, Christina S. 2015. Presupposition satisfaction, locality and discourse constituency. In Schwarz, Florian (ed.), *Experimental perspectives on presuppositions. Studies in theoretical psycholinguistics*, vol. 45, 109–134. Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-07980-6_5

Krifka, Manfred. 1998. Additive particles under stress. In Strolovitch, Devon & Lawson, Aaron (eds.), *Proceedings of semantics and linguistic theory (SALT) 8*, vol. 2. 111–128. Ithaca: CLC Publications. DOI: https://doi.org/10.3765/salt.v8i0.2799

Kripke, Saul A. 2009. Presupposition and anaphora: Remarks on the formulation of the projection problem. *Linguistic Inquiry* 40(3). 367–386. https://www.jstor.org/stable/40284322. DOI: https://doi.org/10.1162/ling.2009.40.3.367

Kruschke, John K. & Aguinis, Herman & Joo, Harry. 2012. The time has come: Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods* 15(4). 722–752. DOI: https://doi.org/10.1177/1094428112457829

Laurinavichyute, Anna & von der Malsburg, Titus. 2022. Semantic attraction in sentence comprehension. *Cognitive Science* 46(2). DOI: https://doi.org/10.1111/cogs.13086

Laurinavichyute, Anna & von der Malsburg, Titus. 2024. Agreement attraction in grammatical sentences and the role of the task. *Journal of Memory and Language* 137. DOI: https://doi.org/10.1016/j.jml.2024.104525

Levy, Roger. 2014. Using R formulae to test for main effects in the presence of higher-order interactions. https://arxiv.org/abs/1405.2094.

R Core Team. 2020. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Roberts, Craige. 1996. Information structure: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics* 5(6). 1–69. DOI: https://doi.org/10.3765/sp.5.6

Rohde, Hannah & Futrell, Richard & Lucas, Christopher G. 2021. What's new? A comprehension bias in favor of informativity. *Cognition* 209. 104491. DOI: https://doi.org/10.1016/j.cognition.2020.104491

Rooth, Mats E. 1985. *Association with focus*: UMass Amherst dissertation.

Sæbø, Kjell J. 2004. Conversational contrast and conventional parallel: Topic implicatures and additive presuppositions. *Journal of Semantics* 21(2). 199–217. DOI: https://doi.org/10.1093/jos/21.2.199

Sorensen, Tanner & Hohenstein, Sven & Vasishth, Shravan. 2016. Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. *The Quantitative Methods for Psychology* 12(3). 175–200. DOI: https://doi.org/10.20982/tqmp.12.3.p175

Spenader, Jennifer. 2002. *Presuppositions in spoken discourse*: Stockholm University dissertation. https://www.ai.rug.nl/~spenader/public_docs/SpenaderDissertation.pdf.

Stan Development Team. 2017. *Stan modeling language users guide and reference manual, version 2.17.0*. https://mc-stan.org/.

Thibodeau, Paul H. & Boroditsky, Lera. 2024. Metaphors we think with: The role of metaphor in reasoning. *PloS one* 6(2). DOI: https://doi.org/10.1371/journal.pone.0016782

Tomioka, Satoshi. 2009. Contrastive topics operate on speech acts. In *Information structure: Theoretical, typological, and experimental perspectives*. Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199570959.003.0006

Vehtari, Aki & Gelman, Andrew & Gabry, Jonah. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* 27(5). 1413–32. DOI: https://doi.org/10.1007/s11222-016-9696-4

Vehtari, Aki & Simpson, Daniel & Gelman, Andrew & Yao, Yuling & Gabry, Jonah. 2015. Pareto smoothed importance sampling. DOI: https://doi.org/10.48550/arXiv.1507.02646

von der Malsburg, Titus & Poppels, Till & Levy, Roper P. 2020. Implicit gender bias in linguistic descriptions for expected events: The cases of the 2016 United States and 2017 United Kingdom elections. *Psychological Science* 31(2). 115–128. DOI: https://doi.org/10.1177/0956797619890619

Winter, Bodo & Duffy, Sarah E. 2020. Can co-speech gestures alone carry the mental timeline? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46(9). 1768–1781. DOI: https://doi.org/10.1037/xlm0000836

Winterstein, Grégoire & Zeevat, Henk. 2012. Empirical constraints on accounts of *too*. *Lingua* 122(15). 1787–1800. DOI: https://doi.org/10.1016/j.lingua.2012.08.003

Zeevat, Henk. 2003. Particles: Presupposition triggers, context markers or speech act markers. In Blutner, Reinhard & Zeevat, Henk (eds.), *Optimality theory and pragmatics*, 91–111. Palgrave-McMillan. DOI: https://doi.org/10.1057/9780230501409_5