Overlapping mechanisms in implying and inferring

Alice Rees
Lewis Bott
School of Psychology, Cardiff University

Address for correspondence

Alice Rees, PhD
School of Philosophy, Psychology, and Language Sciences
Linguistics and English Language,
University of Edinburgh
3 Charles Street
EH9 8AD
Al.b.rees.13@gmail.com

Abstract

Prior psychological work on Gricean implicature has revealed much about how listeners infer (comprehension) but little about how speakers imply (production). This is surprising given the inherent link between the two. The current study aimed to obtain a more integral understanding of implicatures by investigating the processes that are shared between inference and implication. In two experiments, a participant and a confederate engaged in a dialogue game that invited the use of implicatures. In each there was a global priming manipulation, in which a confederate predominantly used implicit or explicit utterances, and a local priming manipulation, in which the utterance structure varied from trial to trial. Participants could choose whether to imply or use an explicit expression. Our results revealed that speaker and listener align on their use of implicatures. We interpret the local priming results as providing evidence of shared implicature representations between speaker and listener, and the global priming results as a form audience design. We also present a model of implicature production that explains our findings.

Overlapping mechanisms in implying and inferring

Utterances communicate much more than the literal meaning of the words. Consider the examples below.

(1) A: I hear Helen's husband is rich and intelligent.
B: Well, he's rich.
(2) John ate some of the cookies.

In (1) Speaker B *says* that Helen's husband is rich, but *implies* that he is not intelligent. In (2) the speaker *says* that John has eaten more than one of the cookies but *implies* that he did not eat them all.

Implications like these illustrate that communication requires speakers and listeners to reason about each other's goals and intentions, rather than simply "decode" the input. There are no words in (1), for example, that mean *not intelligent.* Instead, information about Helen's husband is communicated because the listener reasons about what the speaker could have said ("he's intelligent"), but didn't (the *alternative*), and the speaker recognizes that the listener is able to do this. Grice (1975) thus argued that speakers and listeners must cooperate. Cooperation between interlocutors is seen as fundamental to communication and has been intensely studied in the cognitive sciences (e.g., Chierchia, 2004; Levinson, 2000; Degen & Tanenhaus, 2015; Geurts, 2010; Katsos & Bishop, 2011; Breheny, Katsos & Williams, 2006; Huang & Snedeker, 2009). However, this research has focused on how the listener derives the correct inference, not how the speaker produces an utterance that communicates the correct implication. Here we take an initial step into understanding the speaker's perspective and test whether implying and inferring[1] have overlapping mechanisms.

Gricean inferences

Inferences like those in (1) and (2) can be derived by assuming that the listener reasons about what the speaker could have said but didn't (Grice, 1975). By way of illustration, consider the following account of the listener's reasoning process in (2), adapted from Geurts (2010, described as the "standard recipe").

(i) What could the speaker have said instead of (2)? What are the alternatives? One salient possibility is that the speaker could have used a stronger term and said, "John ate all of the cookies." Why didn't they?
(ii) Presumably, the speaker does not believe that "John ate all of the cookies" is true.
(iii) The speaker has an opinion about whether "John ate all of the cookies" is true.
(iv) Combining (ii) and (iii) yields the inference that the speaker believes *John ate some but not all of the cookies.*

Reasoning with alternatives is generally accepted as the most plausible account of how listeners derive these sorts of enrichments. This is true from a psychological perspective (e.g. Bott & Chemla, 2016; Chemla & Bott, 2014; Rees & Bott, 2018; van Tiel & Schaeken, 2016) just as much as from a linguistic perspective (e.g. Fox & Katzir, 2011; Gazdar, 1979; Geurts, 2010; Horn, 1984; Katzir, 2007). However, while implying and inferring are two sides of the same coin for linguistic theories, the psychological processes used by the listener are fundamentally different from those used by the speaker. Implying works in one direction (from concept to sound i.e. language production) and inferring works in the other (from sound to concept i.e. language comprehension). While the standard recipe above might translate into a good mechanistic account of the inference process, it does not capture the implication process.

[1] We adopt standard linguistics pragmatics terminology whereby the speaker (production) is said to *imply* and the listener (comprehension) to *infer*.

How, then, would an implication be produced? We know of no work that has addressed this question but as a starting point, we suggest the following (based on Levelt, 1989):

(1) A message, F, is conceived and represented in non-linguistic form. Consider two examples, one that could be expressed using an implication, F_1, and one that could not, F_2. The first, F_1, is a message with an upper-bound quantifier meaning, $\exists x \mathrm{C} x[\text{John ate } (x)] \wedge \neg \forall x \mathrm{C} x[\text{John ate } (x)]$ (*John ate some but not all of the cookies*). The second, F_2, is a message with a lower-bound quantifier meaning, $\exists x \mathrm{C} x[\text{John ate } (x)]$ (*John ate some and possibly all of the cookies*],
(2) The message is transformed into a linguistic representation involving lexical elements and syntax, L, as in "John ate some but not all of the cookies", L_1, or "John ate some of the cookies", L_2.
(3) L is checked to establish whether it has an implication structure, $[S \wedge \neg S']$, where S is a shorter sentence than L, and S' is an alternative[2] to S. L_1 is consistent with this structure but L_2 is not.
(4) If L is consistent with $[S \wedge \neg S']$, as in L_1, it is reformulated as S, *John ate some of the cookies*. Then, if the contextual constraints needed to make an implication are satisfied, such as sufficient listener knowledge and shared goals, the sentence is articulated, "John ate some of the cookies". If L is not consistent with $[S \wedge \neg S']$, as in L_2, L is articulated directly without reformulation.

In summary, we assume that the speaker first conceptualises and formulates an enriched message but then later reduces it to a sentence that can be enriched by the listener (we consider alternative models in the General Discussion).

Our hypothesis is that the representation $[S \wedge \neg S']$ is shared across inference and implication, much like syntactic and semantic frames are shared across production and comprehension (Branigan, Pickering & Cleland, 2000; Branigan & Pickering, 2017). The listener uses $[S \wedge \neg S']$ as an enrichment frame to guide comprehension and the speaker uses the same representation to guide production. Pragmatic theories typically describe Gricean implicatures as processes rather than representations, as in (i) – (iv) above, but because processes are unidirectional, none of those suggested in the literature seem appropriate to be shared across production and comprehension. We therefore present our study as a test of shared representations, namely $[S \wedge \neg S']$.

Structural priming

We employed a structural priming task. Structural priming occurs when participants are exposed to a particular linguistic structure on one trial (the *prime*) and then recovery of the same structure is facilitated on a subsequent trial (the *target*). For example, in Bock (1986), participants repeated a prime sentence that could be in active form (e.g., "One of the fans punched the referee") or passive form (e.g., "The referee was punched by one of the fans"), and then had to describe a picture. Participants were more likely to describe a picture in passive form after they had repeated a passive prime sentence than an active prime sentence. Evidence exists that a large range of linguistic structures can be primed. These include passive and active constructions (Bock, 1986; 1992; Bock & Lobell, 1990; Branigan et al., 1995), scoping with *every* (Chemla & Bott, 2014; Raffray & Picking, 2010; Feiman & Snedeker, 2016) and double and single object constructions (Pickering & Branigan, 1998).

Structural priming is often used as a technique to establish the existence of abstract representations because in order for priming to occur the stimuli must share some aspect of linguistic structure within the language processor (see Branigan and Pickering, 2018). Thus, encountering one particular syntactic structure increases the activation of that representation which, in turn, increases the likelihood of that representation being reused. For example, Pickering and

[2] We assume the alternative, S', is semantically stronger than S (i.e. S' entails S), in common with standard formal theories of implicatures (see Katzir, 2011, for a review of the constraints needed on alternatives). For example, *all* is an alternative to *some* because *all* entails some, but not *vice versa*.

Branigan (1998) found that double object syntactic constructions and prepositional object syntactic constructions can be primed in the absence of lexical overlap. They argued that since priming occurred in the absence of lexical overlap, syntactic representations ("syntactic frames") must exist that are abstract from lexical material. Branigan, Pickering, and Cleland (2000) demonstrated priming between a confederate and a participant in a dialogue-based communication task. Since the procedure for language production is reversed for comprehension, in order for priming to occur across modalities, there must be modality independent accessible representations. Our use of structural priming is underpinned by a similar logic to Branigan et al. Since implying and inferring are the reverse processes in a similar way to production and comprehension, we assume that in order to explain priming of enrichment between a speaker and a listener, there must be mechanisms that are shared across production and comprehension.

Participants in our task engaged in a dialogue-based, enrichment priming paradigm based on Branigan, Pickering & Clelland (2000). A participant and a confederate played a communication game in which players took turns describing cards to each other. In some conditions the confederate used an implication to refer to a card and in others an utterance that would cancel an implication. In each subsequent trial the participant had the choice between using an implication herself and using a literal expression. If implying and inferring involve shared mechanisms, rates of implication by the participant should be higher after she makes an inference (i.e. after the confederate implies) than after she does not.

Experiment 1

The game involved the speaker describing one of four cards to the listener. On each trial, cards were arranged according to the structure in Figure 1. The structure invited inference and literal referring expressions. Experimental trials involved reference to the [A] card (e.g. the card with the pencil) or the [AB] card (e.g. the card with the pencil and book). Consider how a speaker could refer to the [A] card. One option would be a simple noun phrase using the [A] object, such as, "the card with the pencil." However, note that according to the structure in Figure 1, the [A] object always occurred in two cards, the [A] and the [AB] card. Thus, to resolve ambiguity about the referent, the listener would be obliged to derive a Gricean inference; namely, that because the speaker did not say the card with the [A] and the [B] objects, they must mean the card with the [A] object only ("because the speaker did not say the card with the pencil and the book, they must mean the card with pencil only"). The other option would be to use a modified noun phrase using the [A] object, such as, "the card with the pencil only," and be explicit. Now consider reference to the [AB] card. Since there was an [AB] card but no [B]-only card, a simple noun phrase involving the (B) object, such as, "the card with the book", was sufficient to identify the [AB] card as the referent. However, by the same logic as the inference was derived for the [A] card, using a simple noun phrase involves cancelling (or suppressing) a Gricean inference that the speaker was referring to a [B]-only card ("because the speaker did not say the card with the pencil and the book, they must mean the card with the book only"). Alternatively, they could use a more explicit expression using a conjunction involving [A] and [B] objects, as in "The card with the pencil and the book". The dependent measure was the proportion of times the participant used unmodified single noun phrases to describe the card.

We tested two forms of priming. The first, *global priming,* was a between subject manipulation. For one group of participants (the *explicit* group), the confederate always used modified utterances to describe the [A] card ("The card with only a pencil"), but for the other group (the *implicit* group), the confederate always used unmodified utterances ("The card with a pencil"). Descriptions of [AB] cards were identical across conditions and used a single referent ("The card with the book."). The representations that govern the use of implicatures might become activated in alignment with the confederate. If so, there should be a higher rate of unmodified single object utterances in the [A] card trials ("The card with the pencil") in the implicit group than the explicit group (or equivalently, a lower rate of modified expressions). For the [AB] card trials, the predictions are more complicated. If implicature representations are highly activated, a single object utterance ("The card

with the [B]") would direct the listener to search for a [B]-only card. Since a [B]-only card would not exist, use of a single object utterance would cause confusion (albeit temporarily until the visual context forced the listener to reinterpret the utterance). A speaker with primed implicature representations might therefore choose to avoid unmodified single object utterances when referring to [AB] cards so that the listener is not disadvantaged. Thus there should be a lower rate of unmodified single object utterances to the [AB] card trials ("The card with the book") in the implicit group relative to the explicit group (or equivalently, a higher rate of conjunction expressions in the implicit group).

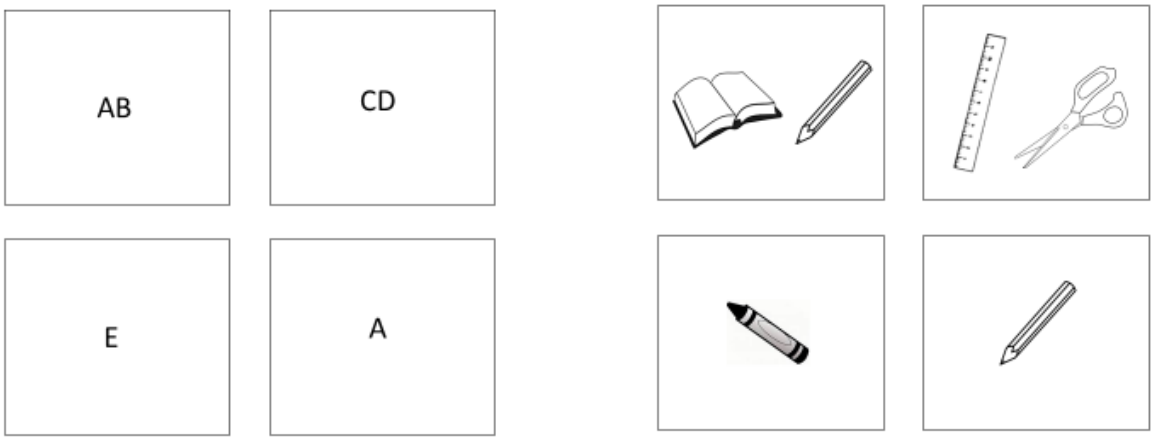

Figure 1. Left panel shows object configuration. Right panel shows an example trial.

Between subject priming manipulations show large priming effects when used with syntactic structures (Thothathiri & Snedeker, 2008) but their interpretation can be ambiguous. In particular, participants might tailor their utterances to suit the addressee by, for example, associating the use of quantity implicatures with a particular individual (as in audience design effects, e.g. Brown-Schmidt, Yoon & Ryskin, 2015; Keysar, Barr, Balin, Brauner, 2000; Garrod & Doherty, 1994; Horton & Gerrig, 2002, 2005). While this reflects priming of a sort, it may not reflect the same type of short-term priming of representations that are described in the structural priming literature. We therefore introduced a within-subject priming manipulation, *local* priming, that was based on the trial sequence, in addition to the global priming manipulation. The local priming manipulation involved the confederate using an implication on some trials but cancelling an implication on others. It therefore applied only to the implicit group because this was the only group where the confederate used an implicature.

For the local priming manipulation, there were four possible experimental confederate-participant trial sequences: [A]->[A], [AB]->[A], [A]->[AB], and [AB]->[AB]. The confederate always described the [A] card and the [AB] cards with simple noun phrases ("The card with a pencil"). Thus, after the confederate described [A] cards implicature representations should be activated immediately since comprehension required deriving an inference. Conversely, implicature

representations should be suppressed after the confederate described [AB] cards since comprehension might require cancelling an inference[3].

When the participant described subsequent [A] cards, there should be a higher rate of implications after [A] cards than [AB] cards, that is, a higher rate on [A]->[A] than [AB]->[A] target trials.

When the participant described subsequent [AB] cards, the effect should be reversed. Participant descriptions of [AB] cards should have lower rates of simple noun phrases following [A] cards than following [AB] cards. Since participant believes that the confederate will derive the inference in the [A]-> [AB] trials, and wishes to counter this view to avoid misleading the confederate, whereas they do not believe the confederate will derive the inference in the [AB]-> [AB] trials).

Consequently an interaction is expected between type of confederate trial and type of participant trial, with rate of unmodified noun phrases as the dependent measure.

Method

Participants

35 Cardiff University students (30 female) participated for payment or course credit.

[3] It is debatable whether there is activation of implicature representations after referring to [B] cards, or suppression after referring to [AB] cards, or both. A classical account of implicatures would require an implicature to identify the [B] card, leading to elevated activation, but the conventional meaning of the single object utterance is that it is consistent with a double object referent ("The card with a [B]" means *a card with a [B] and possibly something else*), and hence would not require suppression of the implicature. However, some authors assume that the implicature arises automatically (Horn, 1981; Rooth, 1985), since pragmatic principles and implicature reasoning themselves are applied by default and no blocking rules are specified. If the default is to derive the implicature with the single object utterance, the implicature must be cancelled (i.e. supressed) when referring to the [AB] card. Thus some accounts would predict suppression after referring to [AB] cards and others activation after [B] cards. All we need to assume for our hypothesis, however, is that one or other of these accounts are correct since both predict opposite effects on activation for identifying [B] and [AB] referent with single object utterances (one predicts activation after [B] but no effects after [AB], the other predicts no effects after [B] but suppression after [AB]).

Materials and counterbalancing

There were five objects in each display, organized according to Figure 1. The five objects were different on every trial. The locations of the card types [A, AB, C, DE] were rotated across trials so that each card type appeared equally often in the four possible display positions. Different objects were used on every trial.

Confederate and participant alternated speaking. There were four experimental prime-target sequences [A->A; AB->A; A->AB and AB->AB] and eight examples of each (32 trial pairs in total). Each experimental sequence was separated by a filler pair. Filler pairs tested C and DE cards. Presentation order was in one of two lists. One order was the reverse of the other.

Eight practice pairs were presented at the start of the experiment to acquaint participants with the procedure. These were a mixture of [A], [AB], [C], and [DE] trials and they were indistinguishable from the main experimental trials. Consequently there were 32 experimental pairs + 32 filler pairs + 8 practice pairs = 144 trials in total.

In the explicit condition confederates described the [A] card with a modified noun phrase using one of four modifiers: "only", "just", "on its own", and "by itself".

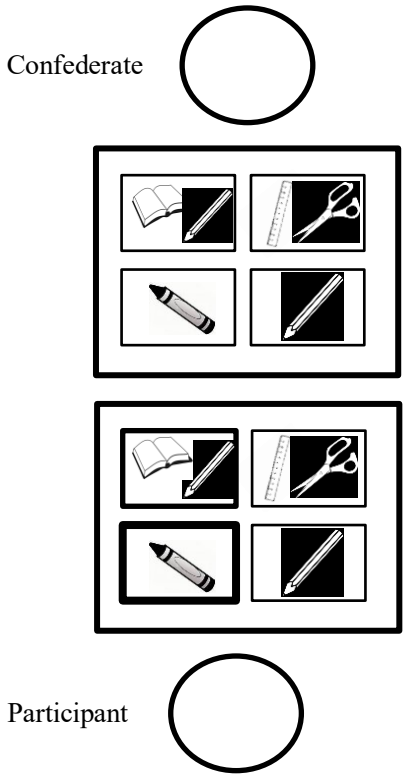

Figure 2. Experiment set up. Highlighted card shows participant which card to describe.

Procedure

Interlocutors sat at opposite sides of a table separated by two monitors (see Fig. 2). They were instructed to describe the card highlighted in bold when it was their turn to be speaker and to identify which card was being described when they were listener. Identification of the card was indicated by pressing one of four buttons on the keyboard. Interlocutors were treated identically by the experimenter throughout and the participant was not made aware that the confederate was not a genuine participant.

Results

Analysis procedure

Utterances were coded with respect to whether they were unmodified single object utterances, e.g. "The card with a pencil," or the alternative utterance type. For [A] card trials, the alternative was to use a modifier with the object, e.g. "The card with only a pencil" (i.e. to be explicit). For [AB] cards, the alternative was to describe both objects, e.g. "The card with a pencil and a ruler." The dependent variable throughout was the proportion of unmodified single object utterances. Data underwent a logit transformation and were analysed using ANOVA. For non-significant comparisons we report Bayes Factors (Dienes, 2011, 2014; Rouder et al. 2009) using the JZS prior (0.707).

We excluded 22 out of 1120 utterances because of an error in a picture. As listeners, participants selected the correct card 98% of the time.

Global priming

Participants in the implicit condition produced more implicit utterances than those in the explicit condition ($F(1,31) = 125.11$, $p < .001$, 95% CI = 3.88-5.61). Participants in the implicit condition produced a greater proportion of unmodified single object utterances (implicatures) on the A card trials than those in the explicit condition, $M = 0.91$ *vs* $M = 0.18$, ($F(1,31) = 108.50$, $p < .001$, 95% CI = 4.17-4.20). Thus the production choices of the participant were influenced by the confederate. Surprisingly, however, there were also differences on the AB cards (which were described with a single object utterance in both conditions). Participants in the implicit condition used significantly

more unmodified single object utterances than those in the explicit condition, $M = 0.82$ *vs* $M = 0.15$, ($F(1, 31) = 58.21$, $p < .001$, 95% CI = -.27 - .50) (participants who didn't use single object utterances used conjunctions involving both objects). Note that the latter effect is in the opposite direction to predictions of a shared implicature representation account. If implicature mechanisms were more strongly activated in the implicit condition, an utterance of the form, "The card with a [B]" would strongly (and misleadingly) imply a card with a [B] and nothing else. Participants in the implicit condition should therefore have used more conjunctions than those in the explicit condition so as to avoid directing the listener to search for a card that was not present. This effect therefore suggests that processes in addition to implicature mechanisms were being primed with the global manipulation. We discuss this further in the GD.

One explanation for the effect on the A card was that lexical material (modifiers) were primed in the explicit condition. For example, after hearing "The card with only an [A]", the modifier "only" could become particularly salient and influence the choice about whether to use an implicature or a modified expression. To test this, we examined the modifiers used by the participant as a function of the modifiers used by the confederate. If the priming effect were lexically based, participants should use the same modifier on trial N as the confederate used in trial N-1. Figure 3 shows the results for the explicit condition (modified responses were too low to be meaningful in the implicit condition). For each modifier, there were a large proportion of explicit responses that were not the same as the confederate. For example, for "only", 60% of the responses used a different modifier while 18% used the same modifier (the remainder used an unmodified expression). The proportion of trials in which the same modifier was used was sufficiently small that we were able to analyse the data after removing these trials (4.5%). When we did this, participants were still more likely to use unmodified utterances when the confederate was also implicit, ($F(1, 31) = 98.00$, $p < .001$, 95% CI = 4.01 – 6.09). Thus, the priming effect cannot be entirely due to lexical priming.

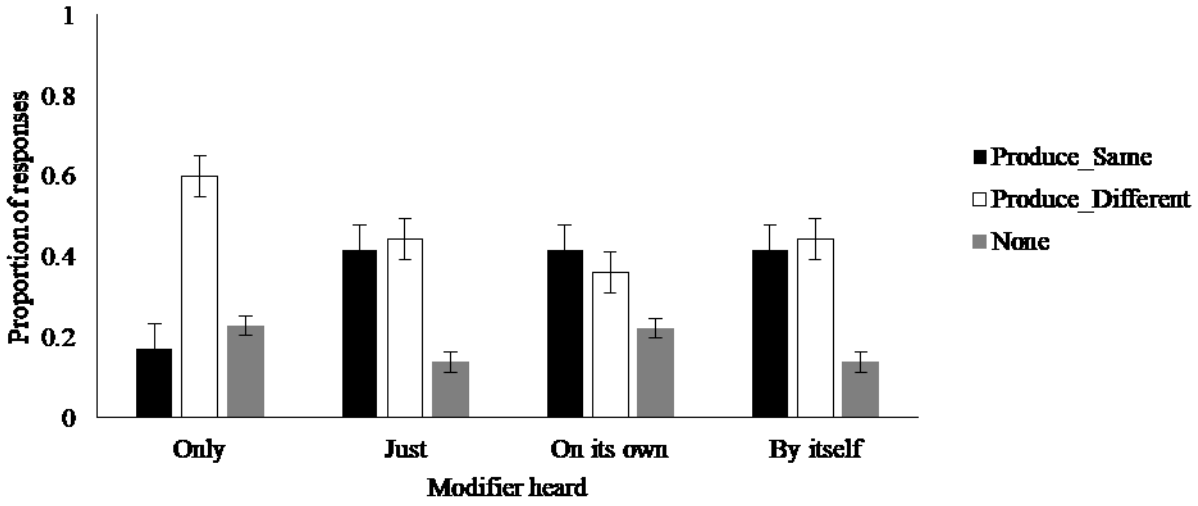

Figure 3. Proportion of responses using the same, different, or no modifier.

Local priming

We tested local priming effects by analysing participant utterances to target trials ([A] and [AB]) as a function of the preceding prime trials ([A] and [AB]). The dependent measure was the (transformed) proportion of unmodified single object utterances. As described above, there was no theoretical reason to expect effects in the explicit group and so we report the local analysis for the implicit group only (although the relevant comparisons are significant even when the explicit group are included).

We found no main effect of prime ($F(1, 16) = 1.43$, $p = .249$, BF = 0.30, 95% CI = -.52 - .15) nor target ($F(1, 16) = 1.76$, $p = .20$, BF = 1.31, 95% CI = -.32 – 1.40). However, there was an interaction between prime and target ($F(1, 16) = 9.99$, $p = .006$) whereby the unmodified single

object prime raised the rate of unmodified single object utterances to [A] cards but lowered them for [AB] cards (see Fig. 4), as predicted by an implicature representation account. Pairwise comparisons showed that this effect was significant on the AB trials but not the A trials ($t(16) = 2.53$, $p = .022$; $t(16) = 1.52$, $p = .148$, BF = 0.65).

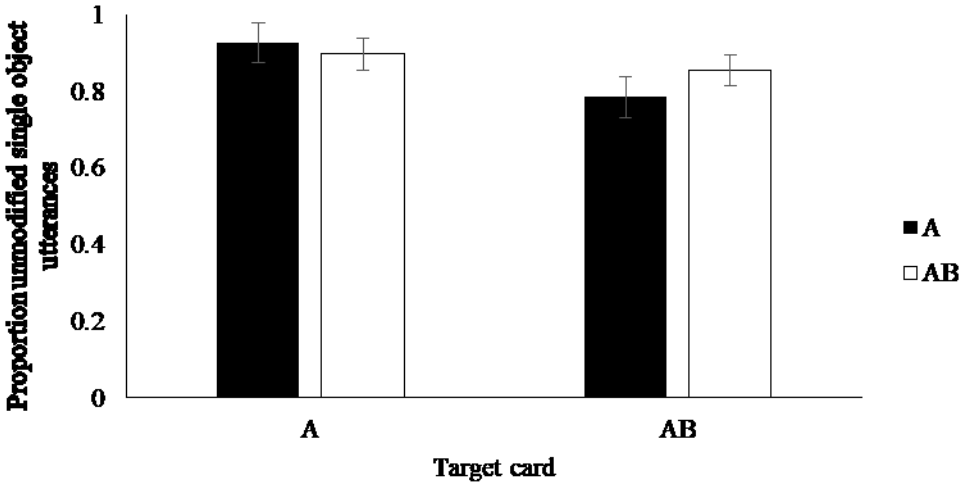

Figure 4. Proportion of implicit responses to targets for the implicit condition. Black bars correspond to A prime trials and white bars correspond to AB prime trials.

Experiment 2

The results of Experiment 1 showed significant priming of enrichment at the local and the global level. However, while the local priming results were consistent with shared implicature representations, the direction of the global priming effects were not. We therefore decided to replicate Experiment 1 (with one minor change).

The global and local priming manipulations of Experiment 2 were the same as those in Experiment 1. However, we varied whether the partner in the communication game was presented as another player or as the experimenter. One group of participants played the communication game against a confederate, with a distinct experimenter providing instructions etc. (just as in Experiment 1), and the other half played against the experimenter directly, without the presence of a confederate. Our motivation for including the partner manipulation was that we thought it possible that an individual's perception of their interlocutor might influence the degree of priming (e.g. van Baaren, Holland, Kawakami, & van Knippenberg, 2004; Bandura & Kupers, 1964; Branigan, Pickering, Pearson, & McLean, 2010; Lakin, Chartrand, & Arkin, 2008; McGuigan, 2013). For example, when the player was perceived as being in the participant's in-group, there might be more imitation compared to when they were not (see e.g. Bourgeois & Hess, 2008; Welkowitz, Feldstein, Finkelstein, & Aylesworth, 1972; Yabar, Johnston, Miles, & Peace 2006). In the event, however, we saw no effects of the player manipulation. We therefore present the experiment as a partial replication of Experiment 1 and leave the social components of enrichment to further research.

Participants

Forty participants from the Cardiff University were recruited and received either course credit or payment. Ten were assigned to each condition.

Design and Materials

The materials and design were the same as in Experiment 2 apart from the partner manipulation.

In the participant condition, the researcher took the role of the partner participant, just as in Experiment 1. There was third person who acted as experimenter. In the experimenter condition, the

researcher took the role of experimenter and partner. They instructed the participant about the experiment and then sat down opposite to play the role of partner.

Results

Confederate role

Numerically, participants produced more unmodified single object descriptions when they knew the confederate was the experimenter compared to when they thought the confederate was another participant (see Fig. 5). Despite the numerical difference this was not statistically significant (F (1, 36) = 1.13, p =.30, BF = 0.39, 95% CI = -1.28 – .40), nor was there was an interaction between confederate role and utterance form (F(1, 36) = .13, p = .73, BF = 0.3).

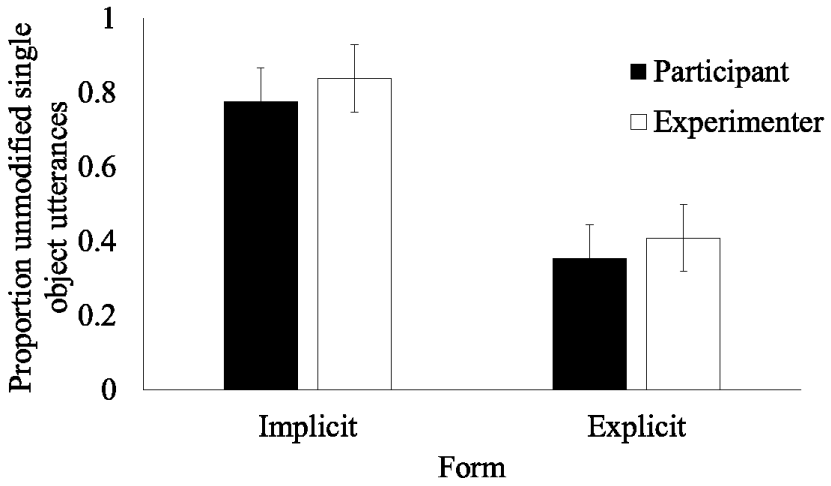

Figure 3. Proportion of unmodified single object utterances to targets. Black bars correspond to the participant condition and white bars correspond to the experimenter condition. Error bars show standard error.

Global priming

The general pattern of results replicated those of Experiment 1. Participants in the implicit condition produced more implicit utterances than those in the explicit condition (F (1, 36) = 45.72, p < .001, 95% CI = 1.97 – 3.65). On the [A] card trials participants in the implicit condition produced a greater proportion of unmodified single object utterances, M = 0.84 *vs* M = 0.21 (F (1, 36) =74.64, p < .001, 95% CI = 3.08 – 5.00). Thus the production choices of the participant were influenced by the confederate's descriptions. This pattern was also observed for the [AB] cards. Participants in the implicit condition used significantly fewer unmodified single object utterances than those in the explicit condition, M = 0.78 *vs* M = 0.56, (F (1, 36) = 5.06, p = .031, 95% CI = .15 – 3.01) (participants who did not use single object utterances used conjunctions involving both objects).

Participants' use of modifiers was analysed as in Experiment 1 to ensure that the global priming results couldn't be accounted for by the repetition of lexical material. If the priming effect was lexically based then participants should use the same modifier as the confederate used on the immediately preceding trial. As in Experiment 1, there were a large proportion of responses that did not use the same modifier as the confederate (see Fig. 6). The proportion of trials in which the same modifier was used was sufficiently small that we were able to analyse the data after removing these trials (4.6%). When we did this participants were still more likely to use unmodified conditions when the confederate was also implicit (F(1,32) = 62.31, p <.001, 95% CI = 2.81 – 4.76).

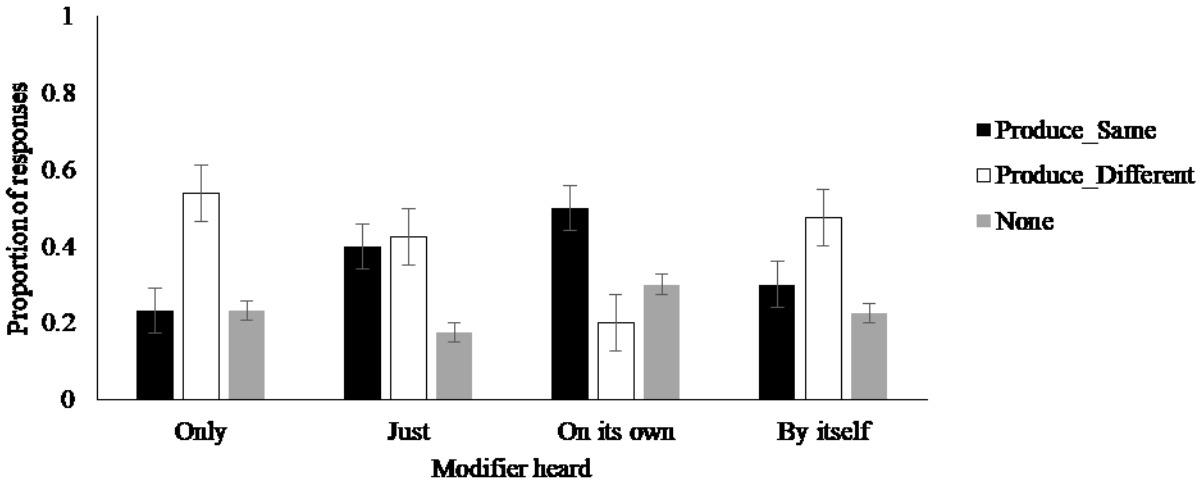

Figure 6. Proportion of responses using the same, different, or no modifier.

Local priming

The pattern of findings from Experiment 1 were replicated (see Fig. 7). There was no effect of prime ($F(1, 32) = .016$, $p = .90$, BF = 0.16) or target ($F(1, 32) = 3.58$, $p = .068$, BF = 6.77). However, there was an interaction between prime type and target ($F(1, 32) = 6.64$, $p = .015$). Following an [A] card prime the rate of unmodified single object utterances was increased for [A] cards but for [AB] cards the rate was lowered. We examined the interaction with pairwise comparisons and found significant effects on the AB targets but not the A targets ($t(32) = 2.06$, $p = .046$; $t(32) = 1.52$, $p = .163$, BF = 0.49).

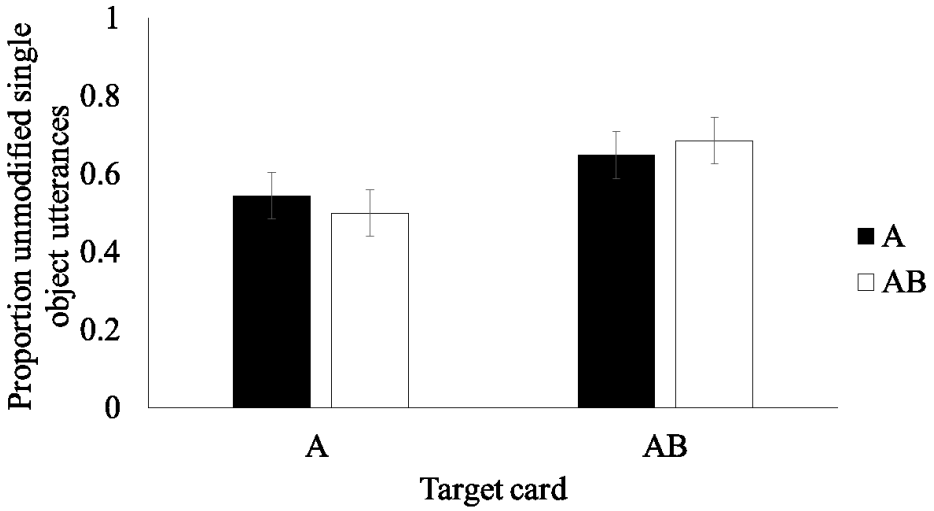

Figure 7. Proportion of unmodified single object utterances to targets. Black bars correspond to [A] prime cards and white bars correspond to [AB] prime cards. Error bars show standard error.

General discussion

In two studies we found that people can be primed to imply, at a global and a local level, by generating an inference. These results suggest that there are shared mechanisms across inference and implication. We now discuss the global and local priming results before turning to the production of implications more generally.

Global priming

The structures used by the confederate had a large effect on those used by the participant. Participants in the explicit condition, i.e. when the confederate used a modified noun phrase to describe the [A] cards ("the card with only the pencil"), were much more likely to use a modifier then those in the implicit condition, i.e. when the confederate did not use a modifier ("the card with the pencil"). Furthermore, even though the confederate referred to the [AB] cards in the same way in the explicit and implicit conditions (using a single, unmodified noun, "The card with a book"), participants were more likely to use a conjunction to describe the [AB] card ("The card with the book and the pencil") in the explicit than the implicit condition. There was clearly alignment between confederate and participant. However, we cannot say definitively what was being aligned.

Three possibilities can be eliminated by the data. The first is that the modifier primed participants e.g. hearing "only" primed participants to use "only" in their utterances. When we removed utterances in which the participant used the same modifier as the confederate, we still obtained significant priming effects, which argues against a modifier priming account. The second is that participants and confederate arrived at a "lexical pact", in which they agreed implicitly to use the same lexical expression on multiple trials. This explanation can be eliminated because different referents were used on every trial and so there was no opportunity for lexical entrainment. The third is that the mechanisms involved in comprehending and producing quantity implicatures were aligned. As we argued earlier, implicature alignment would be shown by a higher rate of conjunctions for the [AB] cards in the implicit condition than the explicit condition, contrary to the observed results.

Our (post-hoc) explanation is that participants in the implicit condition perceived the confederate to be *concise*, i.e. using as few words as possible while maintaining a minimum level of accuracy, and in the explicit condition, *precise*, i.e. using more words than strictly necessary but maximizing accuracy (in the form of minimizing the possibility of errors arising in communication). Participants aligned with the balance between the level of conciseness and preciseness of the confederate. That participants were able to do this provides empirical support for classical pragmatic claims that communicators strive to maintain a balance between two fundamental principles of communication, the cost to the speaker versus the cost to the listener, as in Zipf's "speaker and auditor economies" (1949), or similarly, Horn's (1984) Q and R pragmatic principles, "say as much as you can modulo truthfulness and R" and "say no more than you must, modulo Q", respectively (see also Rohde, Seyfarth, Clark, Jaeger & Kaufman, 2012). In our task, participants minimized the cost to the speaker in the implicit condition and the cost to the listener in the explicit condition. Exactly how they did this is a topic for future research but computationally the process is not trivial. Participants would have to first comprehend the confederate's utterance, then calculate the minimal amount of verbal material needed to describe the referent, note whether the amount of material was greater than this minimum, and then direct the production procedures to implement this conversational style.

A general explanation of global priming relates to the context in which the utterances are used. In the present study, the meaning of "The card with the [A]" is underspecified; it is an appropriate description for both the [AB] and the [A] card since it could mean "The card with the [A] or the [AB]." Participants then use the context to decide which of these it is. In the implicit condition, participants are primed to use the context to narrow down the meaning, that is, that context is a reliable cue. Whereas in the explicit condition, participants are primed that the enrichment cannot be derived from the context.

Local priming

As well as alignment at a global level, we also found alignment on a trial-by-trial basis. When describing [A] cards, participants were more likely to use an implication when the confederate had previously used an implication ([A] cards described using an unmodified noun phrase) than when the confederate had not ([AB] cards described using a single object noun phase), and conversely, when describing [AB] cards, participants were less likely to use a structure that

could give rise to an inference ([AB] cards described using a single object noun phase) when the confederate had previously used an implication ([A] cards described using an unmodified noun phrase) then when the confederate had not ([AB] cards described using a single object noun phase).

This pattern of data is consistent with shared representations across inference and implication. After making an inference, appropriate sentence level representations become activated, such as $[S \wedge \neg S']$. In subsequent production trials, the representation remains active and guides the formation of the utterance. In essence the language processor recognizes that the interlocutor is primed to derive an inference. For [A] targets, it is thus more likely that an implication will produced (an unmodified noun phrase), and for [AB] targets, it is thus more likely that the processor will take steps to cancel a potential inference (by using an explicit structure such as the conjunctive noun phrase).

Towards a model of enrichment in production

This study demonstrates that Gricean enrichment involves an interaction between the production system and the comprehension system, similar to other linguistic phenomena (see Pickering and Garrod, 2013, for a review). Unlike other phenomena, however, there are no mechanistic models of how enrichments are produced (only how they are comprehended, e.g. Bott & Noveck, 2004; Bott, Bailey, Grodner, 2014; Breheny, Katsos & Williams; Degen & Tanenhaus, 2015; Huang & Snedeker, 2009). We therefore sketch a model of enrichment production which incorporates our findings (see Figure 11).

The basic problem is how a conceptualized message e.g. *John ate some but not all of the cookies*, can be articulated in an implicit form, "John ate some of the cookies" rather than in an explicit form, "John ate some but not all of the cookies." The sentence must somehow undergo compression from initial conceptualisation to final articulation. We suggest that the compression process is mediated by the error monitoring system (e.g. Hartsuiker & Kolk, 2001), as follows. We assume that a message is conceptualized in preverbal form, e.g. $\exists x \mathrm{C}x[\mathrm{John\ ate}\ (x)] \wedge \neg\forall x \mathrm{C}x[\mathrm{John\ ate}\ (x)]$. The message is then transformed into a pre-articulated linguistic representation, F, involving lexical elements and syntax, "John ate some but not all of the cookies", by a formulator process. F is then passed through the error monitoring system to verify that the formulated message corresponds to the intended message, as with any other sentence. Crucially, the error checking process involves comprehension procedures that transform the sentence back to its original pre-verbal representation. We suggest that at this point, the processor recognizes that the representation of F, $\exists x \mathrm{C}x[\mathrm{John\ ate}\ (x)] \wedge \neg\forall x \mathrm{C}x[\mathrm{John\ ate}\ (x)]$, corresponds to a more abstract form$[S \wedge \neg S']$, with S as a shorter representation, $\exists x \mathrm{C}x[\mathrm{John\ ate}\ (x)]$, and S' as its alternative, $\forall x \mathrm{C}x[\mathrm{John\ ate}\ (x)]$, that is, a Gricean enrichment. Furthermore, the process of recognizing that F is in Gricean form is probabilistic, just as deriving implicatures is probabilistic in normal comprehension (since implicatures are not obligatory), with a probability proportional to the activation of the enrichment structure, $[S \wedge \neg S']$. The simplified sentence, S, can then passed back to the formulator for reformulation and subsequently articulated.

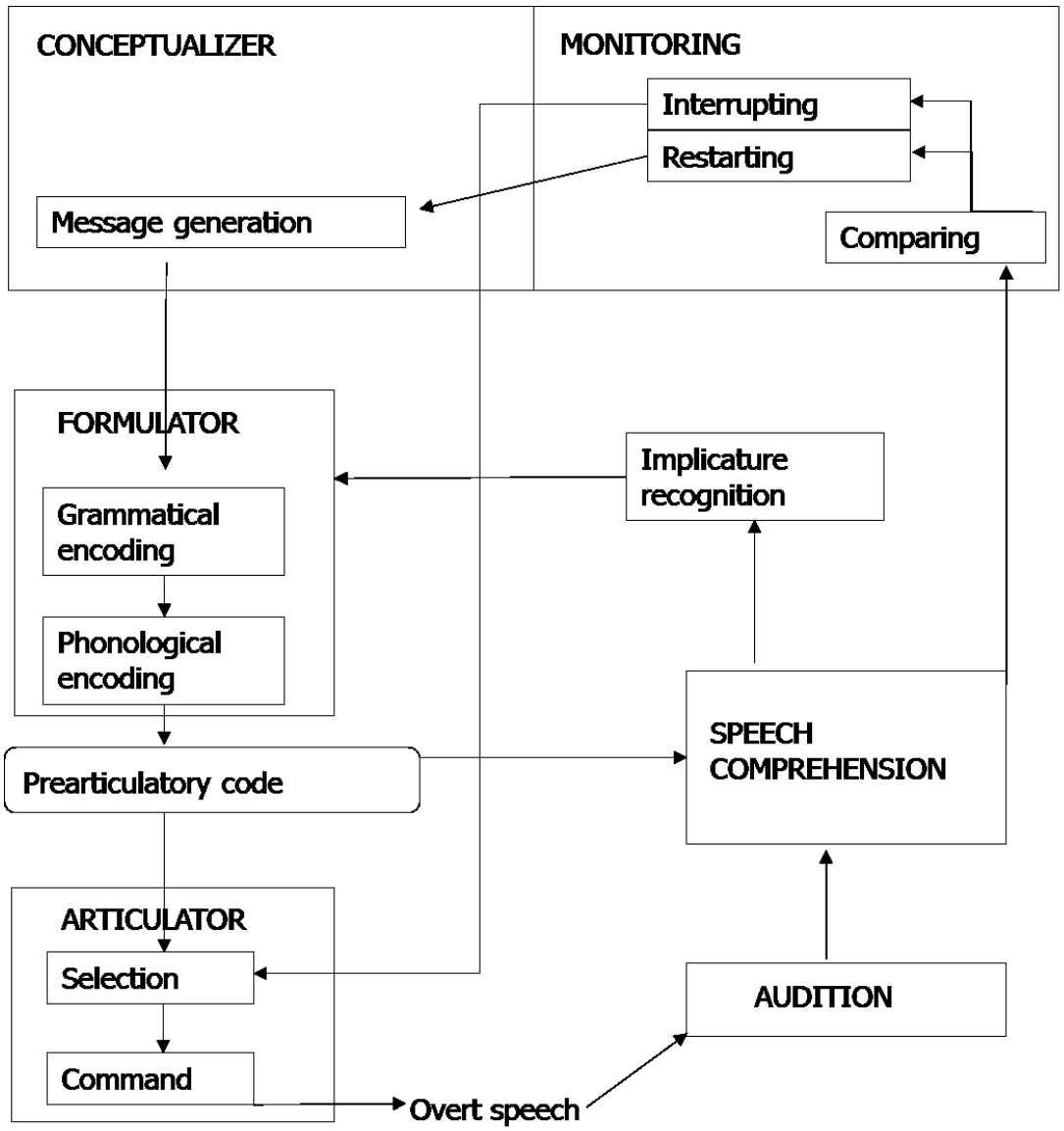

Figure 4. A hypothesized model for implicature production, based on Hartsuiker & Kolk (2001). A message to be articulated is generated at the conceptual level. This message is then transformed into a prearticulatory code via formulator processes (grammatical and phonological encoding). The prearticulatory code gets sent to the articulator for articulation (selection of phonological representations and programming of articulation) and to the comprehension system for error monitoring. During standard error monitoring if the prearticulatory code does not match with the conceptual message generated then speech is interrupted and formulation is restarted. For implicatures the process is slightly different. During the monitoring procedure the processor recognizes that the formulation of the utterance corresponds with a simpler representation; the implicature. At this point the message is reformulated using the implicature representation.

The local priming effects that we observed can therefore be explained as follows. Consider [A]->[A] trials compared with [AB]-> [A] trials. In comprehending [A], the participant is obliged to derive an inference, but not in comprehending [AB] (or the inference is cancelled if it is derived at all). This means that the representation $[S \wedge \neg S']$ becomes more active in [A] than in [AB] trials, and remains so during the subsequent production trials. When the participant formulates the initial sentence, F, in the [A] trial, and feeds F back through the monitoring process, the elevated activation of $[S \wedge \neg S']$ means that the processor is more likely to recognize that F has the $[S \wedge \neg S']$ structure in [A] trials than in [AB] trials. This in turn means that F is more likely to be reformulated as an implication. A similar explanation holds for the difference between [A]->[AB] and [AB]->[AB] trials. After [A] trials, the $[S \wedge \neg S']$ representation is more active than after [AB] trials, and the monitor is therefore more likely to recognize that an unmodified noun phrase

describing [AB] can be interpreted with an inference. Steps are therefore taken to correct the potential misinterpretation by describing the [AB] referent with a conjunction.

The model described above has many untested assumptions that may turn out to be incorrect or overly simplistic. Two in particular stand out. The first is that implications are not formed at the conceptual level but instead arise through a later compression process e.g. *John ate some but not all of the cookies* is the intentional message and this is later compressed into *John ate some of the cookies*. The alternative is that the implication is conceived at the conceptual level without the message having to undergo compression. Our assumption was motivated by the observation that enrichments considered here, quantity implicatures, are employed more for efficiency reasons than for speaker intentions (Levinson, 2000). Speakers are not trying to communicate more than the explicit message by using an implication, it is simply faster for the listener to derive the implication than for the speaker to say the words explicitly. It follows that the implication itself is part of a later, formulation process, rather than the intentional system. But we admit that there are other sorts of implications for which this reasoning may not apply. For example, relevance implicatures involve a speaker saying something irrelevant in order to communicate their message (e.g. A: Mrs X is an old bag; B: Isn't the weather lovely?). In these cases the intention behind the message is linked to the use of an implicature (B wishes to communicate a desire not to discuss the topic) and so cannot entirely be processed by the formulator or monitor.

The second assumption is that implications are mediated by an error monitoring process linked to the prearticulatory phase. An alternative is that compression occurs using a monitoring process, but at an earlier stage in processing. For example, monitoring could occur between conceptualizer and preverbal message, before the formulator, as in Levelt's (1989) conceptualizer-internal loop, which monitors for appropriateness and conceptual errors. The advantage of this model is that the conceptual, contextual and semantic information needed to make decisions about whether to make an implication would be readily available from its error-checking function. However, this account would need to explain why the internal conceptual loop would have access to sentential level representations, such as $[S \wedge \neg S']$, shared across comprehension and production, as needed to explain our data.

The different models described above are all possible accounts for a production model of enrichment but we lack data to distinguish between them. We look forward to future experiments that test these ideas. What we can say with certainty is that implication (production) and inference (comprehension) have overlapping representations, otherwise we would not have observed alignment effects between speaker and listener.

Conclusion

We presented two experiments that demonstrate interlocutors become aligned in their use of implications. That it was possible to prime the production of implicatures suggests that there are overlapping representations involved in implication and inference. More generally our study addresses how the comprehension and production of implicatures interact in the language system, an area of pragmatics that has so far been neglected. We also suggested a model of implicature production, which, while incomplete and untested, raises novel questions about the processing of Gricean implicatures.

Acknowledgements

We wish to thank Sophie Standen, Bianca Diaconou, and Lottie Hopkins for their assistance with data collection.

References

Bandura, A., & Kupers, C. J. (1964). Transmission of patterns of self-reinforcement through modeling. *The Journal of Abnormal and Social Psychology*, *69*, 1-55.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355–387.

Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99*, 150-171.

Bock, K. & Loebell, H. (1990). Framing sentences. *Cognition, 35,* 1-39.

Bott, L. & Noveck, I. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language, 51,* 437-457.

Bott, L., & Chemla, E. (2016). Shared and distinct mechanisms in deriving linguistic enrichment. *Journal of Memory and Language, 91,* 117-140.

Bott, L., Bailey, T., & Grodner, D. (2012). Distinguishing speed from accuracy in scalar implicatures. *Journal of Memory and Language, 66*, 123-142

Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology*, *77*, 343-352.

Branigan, H. & Pickering, M. (2017). 'An experimental approach to linguistic representation'. *Behavioral and Brain Sciences*, 1-73.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co- ordination in dialogue. *Cognition, 75*, 13–25.

Branigan, H. P., Pickering, M. J., Liversedge, S.P., Stewart, A. J., & Urbach, T. P. (1995). Syntactic priming: investigating the mental representation of language. *Journal of Psycholinguistic Research, 24,* 489-506.

Branigan, H. P., Pickering, M. J., Pearson, J., & McLean, J. F. (2010). Linguistic alignment between people and computers. *Journal of Pragmatics*, *42*(9), 2355-2368.

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition, 100*, 434–463.

Brown, P., Levinson, S. C., & Levinson, S. C. (1987). Politeness: Some universals in language usage (Vol. 4). Cambridge university press.

Chemla, E., & Bott, L. (2014). Processing inferences at the semantics/pragmatics frontier: Disjunctions and free choice. *Cognition, 130*, 380-396.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and Beyond.* Oxford UK: Oxford University Press.

Degen, J. & Tanenhaus, M. (2015). Processing scalar implicature: A constraint based approach. *Cognitive Science, 39*, 667-710.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? Perspectives on *Psychological Science, 6*, 274–290.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5,* 1–17.

Ferreira, F., Bailey, K., & Ferraro V. (2001). Good-enough representations in language comprehension. *Psychological Science, 11,* 11-15.

Fox, D., & Katzir, R. (2011). On the characterization of alternatives. Natural language semantics, 19(1), 87-107.

Gazdar, G. (1979). Pragmatics: Presupposition, implicature, and logical form.

Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.

Grice, H. P. (1975). Logic and conversation. 1975, 41-58.

Hartsuicker, R. & Kolk, H. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology, 42,* 113-157.

Horn, L. (1981). Exhaustiveness and the semantics of clefts. In *Proceedings of NELS* (Vol. 11, pp. 125-142).

Horn, L. R. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (ed.), Meaning, Form, and Use in Context: Linguistic Applications (GURT' 84), 11–42. Washington, DC: Georgetown University Press. Reprinted in Kasher (ed., 1998), vol. IV: 389–418.

Horton, W. S., & Gerrig, R. J. (2002). Speakers' experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language*, *47*(4), 589-606.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59*, 91-117.

Horton, W.S. & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. *Cognition, 96*, 127-142.

Huang, Y. & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology, 58,* 376-415.

Katsos, N., & Bishop, D. V. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition, 120*, 67-81.

Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy, 30*, 669–690.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science, 11,* 32-38.

Levelt, W. 1989. Speaking: From Intention to Articulation. Boston: MIT Press.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. MIT press.

McGuigan, N. (2013). The influence of model status on the tendency of young children to over-imitate. *Journal of experimental child psychology*, *116*, 962-969.

Pickering, M. & Branigan, H. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39,* 633-651.

Pickering, M. & Garrod, A. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*, 169-225.

Raffray, C. & Pickering, M.J. (2010). How do people construct logical form during language comprehension? *Psychological Science, 21*, 1090-197.

Rees, A., & Bott, L. (2018). The role of alternative salience in the derivation of scalar implicatures. Cognition, 176, 1-14.

Rohde, H., Seyfarth, S., Clark, B., Jaeger, G., & Kaufmann, S. (2012). Communicating with cost-based implicature: a game-theoretic approach to ambiguity. *Proceedings of the 16th Workshop on the Semantics and Pragmatics of Dialogue*, 107-116.

Rooth, M. (1985). Association with focus. Amherst, MA: University of Massachusetts Amherst PhD thesis. http://scholarworks.umass.edu/dissertations/AAI8509599.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225–237.

Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition*, *108*(1), 51-68.

van Baaren, R.B., Holland, R. W., Kawakami, K., & van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science*, *15*, 71-74.

van Tiel, B., & Schaeken, W. (2016). Processing conversational implicatures: Alternatives and counterfactual reasoning. *Cognitive Science*, *41,* 1119-1154.

Welkowitz, J., Feldstein, S., Finklestein, M., & Aylesworth, L. (1972). Changes in vocal intensity as a function of interspeaker influence. *Perceptual and Motor Skills*, *35*(3), 715-718.

Yarbar, Y., Johnston, L., Miles, L. & Peace, V. (2006). Implicit behavioral mimicry: Investigating the impact of group membership. *Journal of Nonverbal Behavior, 30,* 97-113.

Zipf, G.K. (1949). Human Behavior and the Principle of Least Effort.