Investigating the basis of memory-based effects on common ground

Xiaobei Zheng[1] Richard Breheny[2]

[1]Research Centre for Language and Cognition, Shenzhen University, Shenzhen, China
[2]Department of Linguistics, University College London, London, UK

Correspondence to:
Xiaobei Zheng
xiaobeizheng@szu.edu.cn

Abstract

Much previous research has investigated the effect of domain-general memory processes on production and comprehension in conversation. In this paper, we present a paradigm in which common ground targets are kept consistent between a participant and two different speakers, and demonstrate a speaker effect that draws participants' attention away from the common ground target in Experiment 1. Further, we hypothesise that one important factor that promotes speaker identity as a cue for memory processes is that speakers are in conversational partners' shared attention at the memory coding phase. In Experiment 2 and 3, we demonstrate a non-speaker cue can give rise to a memory-based interference effect when the relevant contextual information is in shared attention but not when it is not. Our results provide an important insight into the interplay of domain-general memory mechanisms and domain specific biases for shared attention.

Introduction

There is no doubt that previously shared information plays an important role in language exchanges and language processing. For example, definite reference can often seem ambiguous without the requisite background knowledge. When Max's father utters 'Max ate the apple' to Max's mother, the definite phrase is many ways ambiguous yet typically succeeds because speaker and hearer share some relevant background experience. Perhaps it refers to an apple that they placed in Max's lunchbox earlier in the day. By contrast, if Max's mother is addressed by a colleague, 'I ate the apple' an entirely different, specific set of shared experience would be required to make this act of reference successful. Also, even if both utterances occur in the course of the same day, there would be little sense that the act of reference is confusing or ambiguous. Studies consistently show an effect of past shared experience on how referential expressions are formed and understood (Clarke & Wilkes-Gibbs, 1986; Metzing & Brennan, 2003).

For a long time, it has been recognised that specific, episodic or relational/declarative memory must play a key role in language processing. Clark and colleagues (Clark & Marshall 1978; 1981, Clark, 1996) see it as laying the foundations of "common ground", which is seen as consisting of information which is commonly assumed, or 'common knowledge' (Stalnaker, 1978). Commonly assumed information can stem from previously shared attention to perceptually co-present objects and utterances, as well as from background cultural and semantic information (Lewis, 1969; Clark & Marshall, 1981). In the past decade or so, several studies have established that mechanisms for the encoding and cued retrieval of memory can serve to support language processes (Horton & Gerrig, 2005a; 2005b). But there remains a question as to what extent these have an effect independently of the perspective-taking processes required to take into account differences in knowledge among interlocutors in order to infer what can be commonly assumed.

A memory-based effect of speaker identities

Horton & Gerrig (2005a) demonstrate that participants' prior experience with two conversational partners influence the form of their subsequent referring expressions. When the partners could serve as a more reliable cue to previously discussed referents, participants were better able to use this information in subsequent production. Similar production studies demonstrate that prior experience of addressees influences speakers' choice of what linguistic form to use (Heller, Gorman & Tanenhaus, 2012; Gorman, Gregg-Harrison, Marsh & Tanenhaus, 2013; Knutsen & Le Bigot, 2012). In a study of memory-based effects on comprehension, Horton & Slaten (2012) demonstrate that participants hearing a temporarily ambiguous referential phrase form an earlier bias in gaze to target when the identity of the speaker is associated with only one of the targets, compared to a condition where speakers had previously referred to both targets. Similarly, Ryskin, Wang, & Brown-Schmidt (2016) examine on-line anticipation of a speaker's referent and demonstrate an effect of speaker-specific memory for spatial perspective. These results are consistent with on-line effects reported in studies of referential processing (Metzing & Brennan, 2003; Brown-Schmidt, 2009; van Berkum, van den Brink, Tesink et al., 2008; Barr, Jackson & Phillips, 2014). They are also in line with studies finding that memory for specific speakers has an effect on language tasks more generally (Creel & Tumlin, 2011; Kamide, 2012; Trude & Brown-Schmidt, 2012).

Although there is clear evidence that prior experience of specific speakers can have an effect on later language processes, there is evidence that effects can be variable. For instance, Creel & Tumlin (2011, Expt. 2) found no effect of speaker identity in a task where this information was implicit and not predictive for the participants' task; while in other experiments when the anticipated target was the object of a specific speaker's preferences, or where the speaker's identity was 100% predictive of target, then a clear effect of specific speaker identity emerged (Creel & Tumlin, 2011, Expts 3-4). Similarly, although Horton (2007) reports a memory-based effect even when the individual's presence at the memory coding phase was peripheral and incidental to the task, Brown-Schmidt & Horton (2014) failed to replicate this effect. As discussed by Brown-Schmidt & Horton (2014), this failure to replicate is not surprising since effects of context on memory are generally elusive (Smith & Vela, 2001), particularly when encoding is not explicitly part of the task (Eich, 1985; Mulligan, 2011).

Variable results have also been reported for studies where speakers constitute part of a correlated cue for retrieval. An effect was demonstrated in Horton & Gerrig (2005a) where each of two conversational partners referred to a distinct superordinate category with the participant. By contrast, Heller and colleagues (2012) failed to find much enhancement of memory-based effects when a correlated cue was available in the task, while Gorman et al. (2013) found only a weak effect.

Turning to theories of the role of memory processes in conversation, Horton & Gerrig (2005b) focus on the fact that memory-based processes are domain-general, operating in support of many processes other than those that may be employed in inferring speakers' intentions based on their beliefs about what is and is not common ground. However, there is little evidence that memory-based processes have an independent effect in the case of language use. According to proposals by Clark and others, mentioned above, memory-based processes would be involved in language use only to the extent that they support processes that are aimed at determining what information is commonly assumed and what is not. It is beyond question that inferences about interlocutors' mental states, particularly shared beliefs, may frequently draw on memory processes to access information about past experience with that person: what they may not know and so forth. The question is whether the mechanisms of memory independently affect language processing in a more direct way – simply making previous memories involving the conversational partner more available for use in language processing irrespective of whether the information contained in those memories support mental state inferences or not.

To date, very little research has been aimed at establishing whether this is the case. As mentioned, Horton (2007) does demonstrate an effect when the individual who serves as a cue is simply a bystander at the memory coding phase. This could be a case of memory having an effect on discourse independently of mental-state-inferential systems deployed in communication, since it is less likely that such processes would code the relevant information as commonly assumed with a bystander. However, as mentioned, Brown-Schmidt & Horton (2014) fail to replicate this effect. Similarly, in a study of 18 month-old children, Moll & Tomasello (2007) compared the response of a group of children who communicatively interact with an experimenter and some novel objects with the response of a group who experience the objects while the experimenter is merely a bystander. Moll & Tomasello found that the children who had richer joint experience were able to exploit shared information at test, while non-interactive group do not. In related research on entrainment effects (Metzing & Brennan, 2003; Kronmuller & Barr, 2007; Brown-Schmidt, 2009), whether memory associations between specific speakers and specific descriptions have an impact early in the time course of referential processing seems to be influenced by whether the speaker is live or not. The idea that the level of live interactivity enhances on-line access to shared assumptions is also established in Brown-Schmidt & Fraundorf (2015).

Thus, there is still something of an issue whether memory-based effects are truly independent of mental state processes for computing what is and is not common ground. This is a question we explore in Experiment 1. In this study, we manipulate whether, in filler items, a speaker communicates about the participant's privileged-ground objects, whose identity the speaker does not know. Because critical trials refer to objects whose identity the speaker should know, we test whether memory associations between the speaker and occluded objects distract participants from the process of focusing attention on the correct domain of objects.

Speaker-associated memory and shared attention

Furthermore, we also move the exploration forwards by analysing the nature of partner-associated memory. Focussing on memory-based effects themselves, Brown-Schmidt, Yoon & Ryskin (2015) discuss the possibility that the specific individual's relevance to the task might be a factor in whether their identity is coded along with target information. This proposal speaks to two separate potential facets of memory-based effects. One has to do with the fact that not all of the innumerable properties of a situation can be coded by memory. Some selection must take place. Second, individuals, as opposed to other contextual information, like location, could be seen as a potent sink for attention, particularly in social situations. This is demonstrated by Smith & Vela (2001) who show in a meta-analysis of context effects on memory retrieval that the use of the same vs. different experimenter had a significant impact on context effects on memory recall.

These two factors have the potential to combine in a virtuous way. In social encounters, interacting agents pay close attention to each other since they are typically rich sources of information relevant to the goals of that interaction (Richardson et al., 2007; 2009).

Thus specific individuals would be among the more effective cues to past situations, as long as their role in those situations were not too peripheral. This would account for the effects like in Clark & Wilkes-Gibbs (1986) as noted in Brown-Schmidt et al. (2015). Also, it would explain why live interlocutors give rise to better results than recorded ones, since live individuals provide a greater focus for attention on the speaker and a richer source of potential cues for future recall.

In this paper, we wish to advance this line of thinking by focusing on shared attention and considering it both as an analytical construct (Barwise, 1988; Breheny, 2006) and as a social-cognitive mechanism (Tomasello et al., 2005; Csibra, 2010; Knoblich, Butterfill & Sebanz, 2011). Formally, situations in which shared attention occurs involve an element of circularity: participants not only attend to the object of their shared attention but also to the situation in which they are attending to that (Barwise, 1988). One feature of this definition that is relevant to the role of memory in communication is that the agent with whom one shares attention is a necessary constituent of the object of one’s attention. It has also been proposed that shared attention, rather than common knowledge is the basis of communication (Barwise, 1988; Peacocke, 2005). This means that when communication occurs, participants share attention to the act of communication. Again, this means that one’s interlocutor is a necessary constituent of one’s attention in conversation, even if the topic of conversation is far removed from the utterance situation itself.

Turning to the cognitive perspective, a wide range of research has revealed mechanisms and biases that support shared social interaction (Richardson et al., 2007; Garrod & Pickering, 2004, Knoblich et al, 2011). In particular, it has been suggested that certain social/communicative stimuli trigger expectations of shared relevance; and primary among such cues are communicative utterances (Csibra, 2010; Tomasello et al., 2007). Long-standing results show that infants demonstrate selective memory for information when that is accompanied by communicative cues (Moll & Tomasello, 2007; Yoon et al., 2008). Beyond child development research, Marno et al. (2014; 2016) establish the same selective memory for relevant information given communicative/social signals with adult participants as with infants.

In summary, communication involves joint attention to the communicative act. This results in enhanced attention to both the speaker and the intended target of the communicative interaction (the message). This greater attention in turn results in greater integration in memory and thus a higher probability of cued recall. In experiment 1, we will have demonstrated that the partner-associated memory traces can have an effect on language processes that work against the task demand to attend only to information which the speaker can identify as common ground. In the following two experiments, we focus on the nature of such partner-associated memory. As aforementioned, the partner with whom one shares attention is a necessary constituent of the communicative act. In experiment 2 and 3, we intend to differentiate the effect of partner-shared memory representations of situations from general memory representations of situations. In these experiments, participants cooperate with only one speaker in the same game as experiment 1. By setting up a verbal sharing process on a particular situation cue (i.e., colour cue), we can find if the low-level memory based effect is mediated by shared attention. If it is the case, we would expect to only find the memory-based effect when the situational cue is verbally shared.

The current studies

In this paper we present a series of visual-world eye-tracking studies that address the issues raised above. In experiment 1, we establish a paradigm that can explore whether memory-based effects exist independently of mental state inferences key to establishing common ground. Our aim here is to explore whether memory-based effects can, in the right circumstances, work against the goals of the comprehender who should choose a common-ground referent. In experiments 2 and 3 we explore the link between speaker specific effects and shared attention.

Our visual-world experiment is based on the design in Keysar et al (2000) and many subsequent studies. Participants are asked to move objects around a 3*3 grid by an instructing speaker while being eye-tracked. Three of the nine grid positions are boxed off on one side so that objects in those positions are only visible to the participant. The other six positions are in common view. See Figure 1.

Three of the six common positions contain objects and the other three are empty. As in previous studies, critical trials involve an instruction like, “Move the apple to the bottom middle” where the target object (in this case, the apple) is in common view. In one condition, there is another apple (a competitor) among the three objects viewable only to the participant (in privileged view). In another condition, none of the three privileged-view objects is of the same type as the target (non-competitor distractors). Participants’ gaze is recorded with eye-tracking equipment and the question of interest concerns the extent to which the second object in private view (here a second apple) attracts attention away from the target object.

According to memory-based accounts mentioned above (Horton, 2007; Horton & Slaten, 2012) we should expect to find greater competitor distraction to the extent that participants associate privileged positions with a specific cue (for example, the speaker’s identity). To test this hypothesis, we use two ways of referring to objects in the grid. In previous studies, reference is often by way of type-based descriptions. I.e. descriptions make reference to the type of object ("the apple", "the dog", "the triangle"). Another way to refer to objects in the grid is by a location-based description. For example, "the object in the top right position". Using location-based descriptions, unlike type-based descriptions, a speaker can give an instruction about an object that they cannot see. For example, the speaker could refer to the banana in Figure 1 as, "the object in the top left position" while it is known that the speaker does not know what that object is. In our studies, participants heard filler instructions that used location-based instructions. In one condition, all instructions, including both type-based test items and location-based fillers refer only to objects in common ground. In another condition, a majority of location-based fillers make reference to privately viewable objects. We call the common-ground only condition the *homogeneous* condition. We call the condition where the speaker gives instructions about both common ground and privileged ground objects the *heterogeneous* condition. In both conditions, critical items are the same and they all involve type-based descriptions of objects in common ground, as in previous studies.

Experiment 1 below is set up to demonstrate an anticipation effect for our manipulation. Participants are being given instruction by two distinct speakers, one associated with a heterogenous condition and the other associated with a homogenous condition. The heterogeneous set of instructions includes filler items referring to privileged view positions. If we induce a speaker-specific memory effect, we expect to find more looks to these locations during heterogeneous trials than for the homogeneous set, particularly as the experimental session proceeds. This in turn ought to induce greater competitor interference in test trials, where the instruction refers to a common-view object by a simple description ("Move the apple..."). We include both the homogeneous and heterogeneous set of trials in one session but provide a consistent memory cue for instructions in each set – that being the identity of the speaker.

1. Experiment 1

1.1. Method

Participants

20 participants were recruited from UCL's participant pool. Two UCL students were recruited to play the speaker roles. One was female and the other was male. These students were blind to purpose of the study. All were paid or given course credit for their participation.

Stimuli and design

Participants were given instructions to move objects around a 3*3 grid as in Figure 1, where three objects were in common view between the instructing speaker and hearer and three objects were placed in grid positions so that they were only visible to the hearer (privileged view). Except for the training phase where participants were given a turn at being director, all instructions were given by the confederate speakers. One speaker gave instructions that referred to common ground objects only (homogeneous condition) and one speaker gave instructions that referred to both common and privileged ground objects (heterogeneous condition). The gender of the speaker in each condition was counterbalanced across the experiment. These instructions were pre-recorded in a single session. Thus each participant in a given set of conditions heard the same pre-recorded instructions. In order to enhance the authenticity of the recording, the speakers added some vocal sounds, such as humming and grunting, in the recoding. The participants were also told that only when the they click on a ring label on the screen would the microphone of the speaker be turned on.

Figure 1. Example display for the main game grids. Participants can see any objects in all the nine positions. The instructor sees the grid from the opposite side and the three objects in the private grey grids are not visible (here, the banana, the orange, and the strawberry).

For each trial, a new set of objects appeared on the display. Each such set contained six objects belonging to the same categories (i.e., fruit, animals or geometric shapes). On the non-competitor trials six pictures were different, while on competitor trials two pictures of the same subtype were presented alongside four randomly selected pictures from the same category. For instance, a competitor condition item might contain a picture of a banana, cherries, grapes, an orange, and two different apples, while for the matched non-competitor condition one of the twinned apples was replaced by a new object, for example, a strawberry (Figure 2).

A trial consisted of one instruction where the participant was asked to move an object to a location. Type-based instruction (i.e. "Move the apple to the bottom middle") and location-based instructions (i.e. "Move the one in the bottom left to the middle right") were used in critical trials and filler items respectively. There were 24 critical trials and 24 filler items in each association condition. 24 critical trials consisted of 12 competitor trials and 12 non-competitor distractor trials. In the competitor trials, the privileged objects included one that was the same type as (although not visually identical to) the target in the instruction (an apple). In the non-competitor distractor trials, no privileged object was of the same type as the target object. Figure 2 below shows initial set ups for a pair of trials in competitor and non-competitor conditions.

Figure 2. The initial display for the competitor condition (left) and non-competitor condition (right). In this example, the target green apple had competitor with the same subtype, e.g. a

red apple, (competitor condition) or an irrelevant non-competitor, e.g. a strawberry, (non-competitor) in participants' privileged grid.

The 24 filler items involved location-based instructions. In a homogeneous association condition, filler items and critical items were all targeted at objects in common view. In a heterogeneous condition, the 24 critical instructions and 8 of the 24 location-based filler instructions were targeted at the common view objects, but 16 of the 24 filler instructions were targeted at privileged view objects.

To sum up, the experiment had a 2(competitor) by 2(association) within-subject design. Altogether, each session with a participant involved 96 trials, 48 delivered by one of the two directors in the homogeneous and 48 delivered by the other director in the heterogeneous condition. The trials from each condition were intermixed. Two lists were used in the experiments. Half participants used the first list and the other used the second list.

Procedure

In each session, a participant was seated in front of a 17" screen fitted with Tobii X60 eye-tracking equipment in a room with only the experimenter. They were told that they would be interacting with two people located in another part of the college via a live video link. First, participants were given some initial instructions. All instructions were given verbally by one of the confederate speakers via prerecorded video appearing on the monitor in front of the participant. The confederate speaker told the participants that he and another director (also introduced) would play a game where they should follow the speaker's instructions to move objects around a grid. They were also shown how the screen would be set up in the speaker's view, and how the speaker would instruct them to move an objects from its original place to a specified position with either a type-based or a location-based instruction. Specifically, two locations were marked by two stars in different colours in the speaker's screen. Participants were shown that the speaker would ask them to move the object in the grid marked by a red start to the grid marked by a blue star. When the object in the privileged grid is not visible to the speaker, they would use a location-based description. Note that, the location-based instructions would be used in the participants' view.

Figure 3. The initial procedure for each trial. Participants were firstly required to move three objects from their private small grid to the three private grey frames in the big grids (fig.3a). After they click on the image of the speaker, three common objects were shown in the transparent frames (fig.3b.)

In the procedure, the 3*3 empty grid appeared on the left of the screen. On the right there was a shelf which contained three objects (Figure 3a). The speaker told the participants that these were their privileged objects, that he did not know what they were or care what they were. The speaker instructed the participants that they should move these objects to the privileged positions on the 3*3 grid as they like, and then clicked on the picture of the speaker in the top right hand corner of the screen when they were ready. When the participant had completed this task, three objects appeared in common view on the grid. The visual display now appeared as in Figure 3b. The speaker then gave one instruction to move an object to a new location. When this was completed by the participants, they clicked the "NEXT" button on the bottom right of the screen, and the whole procedure begins again for the next trial. Note that, on each new trial, the three privileged grid positions appeared in different, randomly assigned locations around the nine frame grid. Note also that, each instruction by an individual speaker was accompanied by a photo of that speaker (one male, one female) in the top right of the display. Participants always clicked on this photo to initiate the display of the common objects, after the first phase of each trial.

Note that if participants are not sensitive to the possible association between speaker and instruction type, the session will simply proceed as if they receive 48 type-based and 48 location-based instructions, with 16 of the latter targeted at private view objects. However, if participants make the association between speakers and targets of instructions, then only one speaker's instructions will give rise to the private-view bias effect.

1.2. Analysis and Results

Data Processing

Eye-movements that were initiated during the auditory instruction were processed according to the critical word ("apple") onsets for the purpose of aggregating the location and duration of each sample from the eye tracker. For analysis, we removed any sample that was deemed "invalid" due to blinks or head movements. The spatial coordinates of the eye movement samples (in pixels) were then mapped onto the appropriate object regions; if a fixation was located within the square surrounding an object, it was coded as belonging to that object, otherwise, it was coded as background.

Our analysis starts from 200ms time region after the critical word onset in accordance with standard assumptions about the time to program and launch an eye movement (Hallett, 1986). Following Hanna et al., (2003) the overall 600ms region of interest was chosen to correspond to the average critical word length (between five hundred and six hundred milliseconds). After this period participants' eye movements were often directed toward the mouse curser. The region was identified and synchronised for each participant on a trial-by-trial basis, relative to the onsets of the critical word.

For each participant (respectively item) and condition, we calculated the average proportion of looks to the competitor/non-competitor distractor over the 50ms time bins. These are plotted against time in Figure 4.

Figure 4. Proportion of looks to the distractor in four conditions in Experiment 1. The zero point on the time (x-) axis shows the absolute onset of the critical word ("apple").

Main Analyses

General linear mixed model regression was carried out using R (version 3.4.1) with *lmer* function in the package lmerTest to obtain parameter estimates. In order to explore whether the looking pattern changes as the language was incrementally processed and as the experiment processed, time and item order entered models as continuous variables. Time, item order, competitor type (competitor vs. non-competitor), association type (homogeneous vs. heterogeneous) were entered into the model as fixed effects, participants were entered into the model as a random effect. The looks to the distractor for each 50ms bin were used as the dependent variable. The two levels of competitor type and association type were coded using deviation contrasts (-.5 vs. +.5). The continuous variable time and item order was centred. Random effects were fitted using a "maximal" random effects structure supported by the data (Barr, Levy, Scheepers, & Tily, 2013). If the model failed to converge, the random slope accounted for the least variance were removed. The significant interactions were then followed up using new models with different reference levels. Mixed-model estimates, standard errors and t-values are shown in Table 1.

Table 1. Parameter estimates for the full model in Exp. 1.

	Estimated parameter	SE of estimate	t-value	Pr
Fixed effects				
intercept	0.189	0.022	8.510	0.000 ***
association	0.021	0.032	0.651	0.523
competitor	0.143	0.032	4.337	0.000 ***
time	0.233	0.036	6.431	0.000 ***
item order	0.001	0.002	0.593	0.553
association : competitor	0.073	0.057	1.282	0.215
association : time	-0.096	0.073	-1.320	0.187
competitor : time	0.400	0.073	5.511	0.000 ***
association : item order	0.011	0.004	2.967	0.003 **
competitor : item order	0.005	0.004	1.392	0.164
time : item order	-0.031	0.011	-2.931	0.003 **
association : competitor : time	-0.044	0.145	-0.303	0.762
association : competitor : item order	0.037	0.007	5.125	0.000 ***
association : time : item order	0.013	0.021	0.606	0.544
competitor : time : item order	-0.073	0.021	-3.492	0.000 ***
association : competitor : time : item order	-0.047	0.042	-1.114	0.265
Random effects #				
subject (intercept)	0.009	0.095		
association	0.018	0.133		
competitor	0.018	0.136		
association: competitor	0.053	0.230		

Notes: * = *p*<.05 ** = *p*<.01 *** = *p*<.001; # random effects show variance and standard deviations.

There was an effect of time, ES= 0.2332, t=6.431, p<0.001, indicating that participants increased their looks to the distractor under all conditions over after the critical noun in the instruction.

There was an effect of competitor type due to a longer look to the distractor in the competitor condition than the non-competitor condition, ES= 0.1425, t= 4.337, p<0.001. This effect of competitor types also increased with time, ES= 0.3996, t= 5.511, p<0.001, indicating that the effect of competitor increased as the instruction was incrementally processed. There was an interaction among competitor type, time and item order, ES= -0.07336, t=-3.492, p<0.001, implying that the participants may be more skilled to identify the target and distractor as the experiment continued, so the interactive effect with incremental language processing reduced.

Although there was no main effect of association type, the interaction between association type and item order was found, ES= 0.01076, t= 2.967, $p<0.01$, indicating that the participants had a longer look to the distractor in the heterogeneous condition than in the homogeneous condition as the experiment continued.

There was an interaction between time and item order, ES= -0.03078, t= -2.931, $p<0.01$, and an interaction among the association type, the competitor type, and the item order, ES= 0.03716, t=5.125, $p<0.001$.

The following analysis explored the effects of association in two competitor conditions. The effect of association in both competitor and non-competitor conditions were explored by two new models with participants as random effects, and item order, time and association type (homogeneous vs. heterogeneous) as fixed effects. **In the non-competitor condition**, no main effect of association type and interaction was found. **In the competitor condition**, the visual bias to the distractor increased as time passed, ES= 0.4330, t=7.266, $p<0.001$. There was no different visual bias to the distractor between the heterogeneous and homogeneous condition, but there was an interaction between association type and item order, ES= -0.02934, t=-4.924, $p<0.001$, showing that the participants had a longer look to the distractor in the heterogeneous condition than in the homogeneous condition as the experiment continued. There was also an interaction between time and item order, ES= -0.06746, t=-3.908, $p<0.001$, suggesting that participants' looking pattern were more influenced by the learning strategy, but less influenced by their language processing.

The effect of competitor type in both homogeneous and heterogeneous conditions was explored by two new models with participants as random effects, and time, item order and association type (homogeneous vs. heterogeneous) as fixed effects. The main effects of competitor type were found in both the **homogeneous** and the **heterogeneous** conditions. However, the interactions between competitor type and item order were also found in both conditions, showing that the competitor effect reduced as experiment continued in the **homogeneous** condition, ES= 0.01353, t=2.7, $p<0.01$, but increased in **heterogeneous** condition, ES= -0.02363, t=-4.509, $p<0.001$.

1.3. Discussion

Participants in this study received 96 instructions to move objects around a grid. 48 instructions used type-based descriptions ("Move the apple...") while 48 used location based descriptions ("Move the object in the top right..."). 16 of those 48 location-based instructions targeted private-view objects. Instructions were delivered by two directors. Each delivered 24 critical, type-based instructions. One delivered the 16 location-based filler instructions targeting the private-view objects (Heterogeneous). The other delivered instructions only targeting common view objects (Homogeneous). If an association between individual speaker and instruction set (Heterogeneous vs. Homogeneous) occurs in the experimental session, we would expect participants to dwell more on private-view objects during critical trials for the Heterogeneous speaker. We find that participants significantly increased bias to the competitor object in the Heterogeneous condition as predicted. Moreover, the attraction to the competitor object increased in the course of the session in heterogeneous condition, consistent with this attraction to the privileged domain being a result of memory associations building up over time.

Our design dissociates any effect of memory that may be recruited by social-cognitive systems to establish common ground targets, from an apparently more domain general memory effect of the kind postulated by Horton & Gerrig, because the common ground was consistent between two conditions, but the domain-general memory processes were influencing participants' expectations about what the speaker would refer to, including privileged items. Given the equivocal results for previous studies seeking memory-based effects in on-line production (Brown-Schmidt & Horton, 2014), our results provide important support for the idea that general memory-based processes have an effect on reference assignment.

Experiment 1 also confirms that individual speaker identity provides a potent source of memory-based effects. We discussed above the idea that speakers make potent cues for memory-based effects due to their relevance to prior social events (Brown-Schmidt et al., 2015). In our next two experiments, our interest is in exploring the basis of this speaker-identity effect.

We begin by trying to replicate the memory-based effect without a speaker-identity cue. In experiment 2, only one speaker interacts with participants. That speaker delivers both the homogeneous set of instructions and the heterogeneous set of instructions. However, when the speaker delivers each type of instruction, a different colour cue is provided and the procedure obliges the participants to attend to that cue.

2. Experiment 2

In our second experiment, each trial is accompanied by one of two salient potential cues, being the colour of the privileged areas of the grid (blue or red). One colour is paired with the set of homogeneous instructions and the other with the heterogeneous instructions. In each trial, the participant has to attend to the colour by answering a question regarding the colour (Figure 6a). Any associations in memory between colour and association type should result in more competitor interference in the heterogeneous condition than the homogeneous over time, since the colour associated with the heterogeneous instructions should result in more attention to privileged areas of the grid.

2.1. Method

Participants

20 participants were recruited from UCL's participant pool. One UCL student from the Department of Linguistics was recruited to play the speaker role. The speaker was blind to purpose of the study. All were paid for their participation.

Stimuli and design

The set of stimuli were chosen from the same set of pictures as in Experiment 1. There were 20 critical trials and 20 filler items in each of two association conditions. As in the first experiment, 20 critical trials consisted of 10 competitor trials and 10 non-competitor distractor trials, and they all involved type-based instructions. The 20 filler items involved location-based instructions. In a homogeneous association condition, all 40 instructions were targeted at objects in common ground. In a heterogeneous condition, the 20 critical instructions and 5 of the 20 location-based filler instructions were targeted at the common ground objects, but 15 of the 20 filler instructions were targeted at privileged ground objects.

All trials, except the non-competitor condition, had two objects of the same kind (e.g. two apples) in the display. All the 10 competitor experimental items and 20 filler items displays included a pair of objects of the same kind (e.g., two apples) and four other different objects. For competitor items the paired-kind object in common ground (the shared one of the objects belonging to the same subcategory, e.g., the apple in common ground) was the target. For fillers none of the paired-kind objects was referred to. So there was one chance in three that either paired-kind object was target, equal to chance.

In order to distinguish the association conditions, the participants' side of the shelves could show one of two colours, as shown in Figure 5 below. For half of the instructions, the participants' side of the shelves were red, while the other half had the side in blue. The speaker gave 80 instructions, 40 of which either referred to common ground objects only (homogeneous condition) and 40 referred to both common and privileged ground objects (heterogeneous condition). One colour accompanied homogeneous instructions while the other accompanied the heterogeneous instructions. The match between the colour and the instruction type was counterbalanced. Half of the participants received the red-homogeneous and blue-heterogeneous instructions pair, and the other half received the blue-homogeneous and red-heterogeneous instructions pair.

Figure 5. The initial display for each trial in the experiment 2. One colour accompanies the Homogeneous instructions and one, the Heterogeneous instructions.

The procedure was the same as Experiment 1, except there was an added question phase before the speaker gave instructions. Participants were asked a question about the colour in the background of their privileged grid locations. This question appeared on the participant's screen, as in Figure 6a. They responded by clicking on the answer, given a choice on the screen (see Figure 6a). The Director neither saw the question on the participant's screen nor knew of the answer. This fact was made clear to the participant in the instruction and training phase.

To sum up, the experiment had a 2(competitor) by 2(association) within-subject design. Altogether, each session with a participant involved 80 trials, 40 delivered in the homogeneous and 40 in the heterogeneous condition. In each association condition, 10 trials referred to targets that had no competitor in privileged ground and 10 referred to targets which had competitors in privileged ground. Two types of critical instructions were matched and used the same recordings. The trials from each condition were intermixed. Two lists were used in the experiment, so under each colour-instruction pair there were two list versions. Half of the participants received the first list, and the other half received the second list.

Figure 6. The question phases in experiments 2(see a) and 3(see b). In 2, the question appears on the screen and is visible to the participant only. The participant answers the question silently by clicking on the correct colour. The director does not see the question or the answer. In experiment 3, the director asks the participant the same question ("What colour is your side?"), and participants were only presented the answers (see b). In experiment 3's procedure the director repeated the answer in confirmation to emphasise shared knowledge of the colour.

2.2 Results

Eye-movement analyses were the same as Experiment 1. The best fit model including time, item order, competitor type (competitor vs. non-competitor), association type (homogeneous vs. heterogeneous) as fixed effects, participants were entered into the model as a random effect. The looks to the distractor for each 50ms bin were used as the dependent variable. Random effects were fitted using a "maximal" random effects structure supported by the data. Mixed-model estimates, standard errors and t-values are shown in Table 2.

Table 2. Parameter estimates for the full model in Exp. 2.

	Estimated parameter	SE of estimate	t-value	Pr
Fixed effects				
intercept	0.189	0.018	10.552	0.000 ***
association	0.011	0.029	0.382	0.707
competitor	0.107	0.034	3.158	0.005 **
time	0.144	0.040	3.584	0.000 ***
item order	0.008	0.002	3.470	0.000 ***
association : competitor	-0.092	0.067	-1.374	0.185
association : time	-0.002	0.080	-0.022	0.983
competitor : time	0.272	0.080	3.388	0.001 ***
association : item order	0.008	0.005	1.628	0.104
competitor : item order	-0.007	0.005	-1.350	0.177
time : item order	-0.035	0.014	-2.480	0.013 *
association : competitor : time	-0.088	0.161	-0.549	0.583
association : competitor : item order	0.006	0.010	0.659	0.510
association : time : item order	0.034	0.030	1.230	0.219
competitor : time : item order	-0.107	0.030	-3.821	0.000 ***
association : competitor : time : item order	0.321	0.056	0.575	0.566
Random effects #				
subject (intercept)	0.005	0.074		
association	0.014	0.116		
competitor	0.019	0.138		
association: competitor	0.074	0.271		

Notes: * = p<.05 ** = p<.01 *** = p<.001; # random effects show variance and standard deviations.

There was a main effect of time, ES= 0.1439, t= 3.584, p<0.001, and a main effect of the item order, ES= 0.008371, t= 3.470, p<0.001, suggesting that participants increased their look to the distractors as the language was unfolded, and as the experiment continued.

There was a main effect of competitor type, ES= 0.1067, t= 3.158, p<0.01. The effect of competitor type also increased with time, ES= 0.272, t= 3.388, p<0.001, indicating that the effect of competitor increased as time passed. There was also an interaction among the competitor type, time, and the item order, ES= -0.1068, t= -3.821, p<0.001, suggesting the as the experiment continued, the incremental processing of language has less influence on the effect of competitor type.

There was no main effect of and interaction with association type.

Planned analysis explored the effects of competitor type in two association conditions. Two new models were established with participants as random effects, and time, item order and association type (homogeneous vs. heterogeneous) as fixed effects. In both the **homogeneous** and the **heterogeneous** condition, there were main effects of time (*p*s<0.05), showing that participants increased their look to the distractor overall as the language was unfolded. There were interactions between the competitor type and the time (*p*s<0.05), suggesting that participants increased their look to the distractor as language was incrementally revealed. There were interactions among competitor type, item order and time (*p*s<0.05), showing that as experiment passed, the language processing effect on the competitor type reduced.

The effect of association in both competitor and non-competitor conditions were explored by two new models with participants as random effects, and item order, time and association type (homogeneous vs. heterogeneous) as fixed effects. In both the **non-competitor and competitor conditions,** no effect and interaction with the association type was found (*p*s>0.152).

For each participant (respectively item) and condition, we calculated the average proportion of looks to the competitor/non-competitor distractor over the 50ms time bins. These are plotted against time in Figure 7.

Figure 7. Proportion of looks to the distractor in four conditions in Experiment 2. The zero point on the time (x-) axis shows the absolute onset of the critical word ("apple").

2.3 Discussion

As in the first experiment, the majority of instructions that participants received in Experiment 2 are directed at common-view objects, but a small proportion are directed at privileged view objects, via location-based instructions. If the participant is sensitive to cues to the homogeneous and heterogeneous sets of instructions, we should see more competitor interference in the heterogeneous condition. Unlike in our first experiment, where speaker identity was the cue, in this experiment a prominent colour in the display was the cue. Two features of our design gave this cue the best chance possible to be effective: We required participants to attend to it (by answering a question about it) just before each trial's instruction was given. In addition, the visual salience of the colour cue was greater than the visual salience of the image of the different speakers in Experiment 1 (see Figure 3); also the colour filled the privileged grid positions. In spite of the prominence of the colour cue and the fact that participants were required to attend to it just prior to each instruction, we failed to find the same kind of memory based effect as in the first experiment.

This contrast in results is consistent with previous findings discussed above that memory effects are variable and that speaker identity is a potent cue for memory effects (Smith & Vela, 2001). The results are also in line with Heller et al. (2012) and Gorman et al. (2013), which were unable to detect much impact of categories (such as phonological onset of the name or object category) as cues to the ground status of some information. One difference between these latter studies and Experiment 2 however is that the potential cue in Experiment 2 is directly relevant to the participants' preliminary task and not simply an incidental source of association available in the experimental context.

In Experiment 3, we explore why we were unable to replicate the memory based effect of the first experiment using salient and task-relevant colour cues. One observation we can make about the previous two experiments is that the stimuli that serve as memory cue may be available at different points to the participant. In particular, we are analyzing the interference effect when participants hear the critical word ("apple"). At this point in Experiment 1, we can be more sure that the participant is attending to a potential cue (being the voice of the individual speaker). By contrast, in Experiment 2, even though the colour cue is highly salient on the screen, we can be less sure that the participant attends to that while they hear the critical word. This line of thought points to one reason why speaker identity is liable to be more relevant (in some sense) than other potential cues, at least when conversation is involved (Brown-Schmidt et al., 2015).

Our proposal is that, at least in part, what makes the speaker's identity both a source of attention and a potent cue in this situation is that the speaker is part of what is the current target of shared attention and was previously part of the target of shared attention at the coding phase for any relevant associations. Our specific proposal is that shared attention in communication creates a bias to code the objects of shared attention more deeply in memory due to its likely relevance. Since interlocutors are necessary parts of what is shared attention, it makes sense that they would serve as strong cues for memory-based effects. From this perspective, there are two potentially important differences between the cues in experiment 1 and 2: on the one hand, they involve specific speaker identity vs. physical context cue; on the other, one is in shared attention while one is not. In previous research where memory effects involving specific individuals have been weak, these are more clearly cases where that individual may not have been jointly engaged with the participant at the coding phase. For example, the procedure in Horton (2007) involves an individual only peripherally involved with the on-screen task that the participant undertakes in the preliminary phase of the procedure.

In Experiment 3, we explore whether speaker identity is the deciding factor here, or sharedness of attention to the cue. This experiment has an identical procedure to Experiment 2 except for one critical change. In Experiment 3, the participant is asked the question about the background colour of the private-view grids by the Director, instead of that question appearing on-screen only for the participant. In the procedure, the Director also confirms they have heard the answer. Thus, the colour of the background comes into shared attention at this point in the procedure, whereas in the previous experiment, the colour was not shared. If speaker identity is the deciding factor, we expect no effect of association condition (Homogeneous vs. Heterogeneous) when it comes to interference by the distractor. However, if shared attention to potential memory cues enhances memory-based processes, we should expect more competitor interference in the competitor condition vs. the non-competitor condition over time, in the Heterogeneous condition.

3. Experiment 3

3.1. Method

Participants

20 participants were recruited from UCL's participant pool. One UCL student from the Department of Linguistics were recruited to play the speaker role. The speaker was blind to purpose of the study. All were paid for their participation.

Stimuli and procedure

The set of stimuli were chosen from the same set of pictures as in previous studies. The procedure was the same as Experiment 2 with one difference: the director asks a question ("What colour is your side?") about the coloured background that the participant can see - see Figure 6b. As illustrated in Figure 6b, the participant's screen at this point shows the choice between the two colours (Blue and Red). As in experiment 2, participants click on the correct answer on the screen. At this point, the Director says the name of the colour by way of confirmation. The confirmation part of the procedure was included to highlight uptake of the answer by the Director. The use of confirmation has been shown to better establish the answer to a question being shared between interlocutors (Brown-Schmidt, 2009).

3.2. Analysis and Results

Eye-movement analyses were the same as the previous experiments. Time, item order, competitor type (competitor vs. non-competitor), association type (homogeneous vs. heterogeneous) entered the model as fixed effects. Participants were entered into the model as a random effect. Random effects were fitted using a "maximal" random effects structure supported by the data. The significant interactions were then followed up using new models with different reference levels. Mixed-model estimates, standard errors and t-values are shown in Table 3.

Table 3. Parameter estimates for the full model in Exp. 3.

	Estimated parameter	SE of estimate	t-value	Pr
Fixed effects				
intercept	0.257	0.025	10.481	0.000 ***
association	0.034	0.050	0.665	0.514

competitor	0.062	0.040	1.571	0.132
time	0.198	0.045	4.408	0.000 ***
item order	-0.009	0.007	-1.284	0.214
association : competitor	0.151	0.059	2.549	0.019 *
association : time	0.302	0.090	3.363	0.001 ***
competitor : time	0.333	0.090	3.718	0.000 ***
association : item order	0.047	0.013	3.625	0.002 **
competitor : item order	0.042	0.016	2.632	0.016 *
time : item order	-0.029	0.016	-1.867	0.061 .
association : competitor : time	-0.091	0.179	-0.509	0.611
association : competitor : item order	0.075	0.023	3.335	0.003 **
association : time : item order	0.069	0.031	2.221	0.026 *
competitor : time : item order	0.032	0.031	1.017	0.309
association : competitor : time : item order	-0.013	0.062	-0.209	0.835
Random effects #				
subject (intercept)	0.011	0.104		
association	0.046	0.215		
competitor	0.027	0.163		
item order	0.001	0.030		
association : competitor	0.051	0.226		
association : item order	0.003	0.053		
competitor : item order	0.004	0.066		
association : competitor : item order	0.008	0.089		

Notes: * = $p<.05$ ** = $p<.01$ *** = $p<.001$; # random effects show variance and standard deviations.

There was an effect of time, ES= 0.1976, t=4.408, p<0.001, indicating that participants increased their looks to the distractor under all conditions over after the critical noun in the instruction.

Although there was no main effect of competitor types, the effect of competitor types increased with time, ES=0.3334, t= 3.718, p<0.001, and the item order, ES=0.04163, t= 2.632, p<0.05, indicating that the participants looked longer to the distractors as the language was unfolded, and as the experiment continued.

Although there was no main effect of association types, the effect of association types increased with time, ES=0.3016, t=3.363, p<0.001, and item order, ES=0.04726, t= 3.625,

$p<0.01$, indicating that participants looked longer to the heterogeneous distractor as the language was unfolded, and as the experiment continued.

There was an interaction between association type and competitor type, ES= 0.1512, t=2.549, $p<0.05$, and an interaction among association type, competitor type, and item order, ES= 0.07548, t=3.335, $p<0.01$.

Following analysis explored the effects of competitor type in two association conditions. The effect of competitor types in both homogeneous and heterogeneous conditions were explored by two new models with participants as random effects, and time, item order and association types (homogeneous vs. heterogeneous) as fixed effects. In both the **homogeneous** condition and the **heterogeneous** condition, there were interactions between the competitor type and time (*p*s<0.05), suggesting that participants increased their look to the distractor in the competitor condition as the language was unfolded. However, a main effect of competitor type (ES= -0.1379, t= -2.949, $p<0.01$) and an interaction between competitor type and item order (ES= -0.07937, t= -4.389, $p<0.001$) were reported only in the **heterogeneous** condition, showing that participants have a longer look to the competitor items, and the pattern was also increased as the experiment continued.

The effect of association in both competitor and non-competitor conditions were explored by two new models with participants as random effects, and item order, time and association types (homogeneous vs. heterogeneous) as fixed effects. **In the non-competitor condition**, no main effect of association types was found, ES= 0.04208, t= 0.770, p=0.45006. An unpredicted interaction between association types and time (ES= -0.34720, t= -2.893, $p<0.01$) shows that participants increased their look to the distractor in the heterogeneous condition than in the homogeneous condition as the instruction unfolded. It is possible that because the privileged grids had different colours, this constantly highlighted the their being in the potential referential domain in the heterogeneous condition, even when there was no competitor object in those grids. But on top of the bias, participants had stronger competitor preference **in the competitor condition**, because there was a marginally significant main effect of the association type, ES= -0.1092, t= -1.862, p=0.0774, and an interaction between association types and the item order, ES= -0.085, t= -5.230, p<0.001, suggesting that participants looked longer to the heterogeneous competitors than the homogeneous competitors, and they also increased the looking pattern as the experiment continued.

For each participant (respectively item) and condition, we calculated the average proportion of looks to the competitor/non-competitor distractor over the 50ms time bins. These are plotted against time in Figure 8.

Figure 8. Proportion of looks to the distractor in four conditions in Experiment 3. The zero point on the time (x-) axis shows the absolute onset of the critical word ("apple").

3.3. Discussion

The difference in procedure between experiments 2 and 3 lay solely in the phase before the instruction when the participant answered a question about the background colour on the screen. In Experiment 2 we attempted to ensure that information about the colour in each trial was only known to the participant (privileged information). In experiment 3, we attempted to ensure that the information about the colour was shared between the participant and the Director. We find a marked difference between these experiments. In experiment 3, when the cue to memory-based effects was shared, we found more interference from the distractor in the heterogeneous condition – the same result as in Experiment 1, when different speakers' identities served as the cue. When the cue was not shared, we found no such effect.

4. General Discussion

We presented three experiments in which participants are instructed to move objects around a 3*3 grid where some objects are in common view with the director and others are seen only by the participant. As in many previous experiments, we designed the study so that we could compare eye gaze to a privileged distractor position when participants hear, "move the apple...". That distractor position may contain a competitor (an apple) or a non-competitor (a non-apple). Filler instructions in our study make use of location-based instructions ("move the object in the top right..."). A smaller proportion of those filler trials target privileged-view positions, meaning that participants occasionally are asked to attend to these positions in the course of an instruction. This manipulation means that participants may be attracted to privileged positions when being given instructions. In the Competitor condition, an object of the mentioned type ("apple") is in private view. Thus we should then expect to see greater attention to this distractor compared to the Non-competitor distractor. In our first experiment, the private-target filler instructions are given by one of two speakers (Heterogeneous condition); the other speaker's instructions only target common-view objects in both type-based experimental items and location-based filler items (Homogeneous condition). If participants associate only the one speaker with private-view instructions, then we should expect to see a greater attention to competitor distractors in the heterogeneous condition than the homogeneous condition. In Experiment 1, we found this effect of association between the Heterogeneous speaker and private view objects. This is a demonstration of the operation of domain-general memory systems on language processing since the effect of memory runs contrary to the successful operation of language processing, which is aimed to recovering a definite referent from common ground. Since there has been little previous research which convincingly shows a memory-based effect independent of any potential mental-state inferential processes for inferring referents, the results of Experiment 1 provide important support for the idea that domain-general memory associations can independently influence processing definite reference.

In our first experiment, the participants alternately received instructions from two speakers, one of whom occasionally referred to the participants' private-view objects. The situation is, to some extent, similar to a three-party conversation except no communication happening between the two confederate speakers. Yoon & Brown-Schmidt (2018)'s study on multiparty conversation demonstrated that speakers produced distinct expressions for different partners when the partners had different degrees of experience with the same shared knowledge. These results indicate that the speaker-specific memory of the shared knowledge influences audience design. The present study further showed that the interference from the privileged ground varied between two partners when they had the same shared knowledge but different degrees of involvement with the partner's privileged ground. The effect of the speaker-specific memory is applied into the domain of privileged ground, in addition to the common ground.

In our studies, the memory-based effect does not manifest an association between a cue (e.g. a particular speaker) and a referential expression or particular referent, rather, the association is between a cue and a whole set of objects in privileged-ground. We believe that the effectiveness of our design in bringing out these memory-based effects lies in the fact that it is not unusual for objects whose identities are unknown to the speaker, or partially known, to nevertheless be a subject of conversation. One particular example of this is when someone asks a question about an object. Brown-Schmidt et al. (2008) illustrates that when it comes to questions, participants do readily anticipate non-common-ground referents. As in Brown-

Schmidt et al., what is not common ground about our occluded objects is not their existence, but their specific identities (being an apple, for example). Thus, we can make sense of the effectiveness of our design in terms of a more general state of affairs where conversation may target multiple domains of reference. Heller et al. (2016) successfully model more widely discussed perspective taking results in a Bayesian framework where different domains of reference are weighted as a result of factors present in the conversation. If we consider our results from the perspective of Heller et al.'s model, we can say that memory associations which are unrelated to inferences about common ground can affect these weightings. This is a line of thinking that would need to be pursued in future research.

In experiments 2 and 3 we further explored the basis of this memory effect: Is it only speaker identity that can be the basis of domain-general memory-based effects? In these experiments, only one speaker gives both the Homogeneous and Heterogeneous sets of instructions. The different sets of instructions are accompanied by distinct prominent colours in the background of the privileged positions. Participants are asked to identify the colour prior to each trial, to ensure they attend to the potential cues to the different instruction sets. In Experiment 2, the colour cue is only known to the participants. In Experiment 3 the colour cue is shared between participant and director via a question/answer interaction. The idea that specific speaker identity is critical to memory-based effects would suggest that both Experiments 2 and 3 would be less likely to replicate the effects of Experiment 1. Our alternative idea is that speaker identity is a potent memory cue, at least in part, because the speaker is part of what is jointly attended to in communicative interactions. We propose that objects of previously shared attention are more liable to cued retrieval. Thus, any information that has previously been in shared attention can be the source of memory-based effects. The prediction then was that we see the memory-based effect in Experiment 3, and this is what we found.

We wish to reiterate that the effects demonstrated in experiments 1 and 3 must result from a domain-general memory mechanism and not one that is absolutely constrained by processes that might determine what is common ground. Our hypothesis is, however, that shared, or joint, attention creates a strong bias to code associations among information that is the focus of shared attention, and this includes the agents of shared attention itself. The result is that what has been previously shared is more susceptible to retrieval and use in later processing, but the processes that underpin this retrieval are domain-general. If our hypothesis is on the right track, it means that domain-general memory processes are modulated by domain-specific social processes for shared attention and that common ground information becomes more available for use in conversation as a result of this interplay of memory mechanisms and social-cognitive processes for shared attention.

Finally, it has been pointed out to us that the increased bias found to privileged objects in experiments 1 and 3 might not be the result of normal, more or less unconscious, processes for resolving reference, but may be due to an awareness on the part of participants that the heterogeneous cues signal some kind of different task. We are sceptical that that these biases reflect much other than normal referential processes but cannot rule this possibility out. Nevertheless, the key point is that any such awareness triggered by cues, must be based on memory representation of previous trials. This point is borne out by the effect of item order in experiments 1 and 3. Thus, the main conclusions about the nature of memory effects would still stand.

References

Barr, D. J., Jackson, L., & Phillips, I. (2014). Using a voice to put a name to a face: the psycholinguistics of proper name comprehension. *Journal of Experimental Psychology: General,* 143(1), 404.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278. http://dx.doi.org/10.1016/j.jml.2012.11.001

Barwise, J. (1988). Three Views of Common Knowledge, in *Proceedings of the Second Conference on Theoretical Aspects of Reasoning About Knowledge*, M.Y. Vardi (ed.), San Francisco: Morgan Kaufman, pp. 365–379. (reproduced in Barwise, (1989).

Barwise, J. (1989). *The Situation in Logic*. Stanford, CA: CSLI.

Breheny, R. (2006). Communication and folk psychology. *Mind and Language* 21(1): 74–107.

Brown-Schmidt, S. (2009). Partner-specific interpretation of maintained referential precedents during interactive dialog. *Journal of Memory and Language*, 61(2), 171–190. doi:10.1016/j.jml.2009.04.003.

Brown-Schmidt, S., & Fraundorf, S.H. (2015). Interpretation of informational questions modulated by joint knowledge and intonational contours. *Journal of Memory and Language*. 84: 49-74.

Brown-Schmidt, S., & Horton, W. S. (2014). The influence of partner-specific memory associations on picture naming: a failure to replicate Horton (2007). *PloS One*, 9(10), e109035. http://dx.doi.org/10.1371/journal.pone.0109035.

Brown-Schmidt, S., Yoon, S., and Ryskin,R.A. (2015). People as contexts in conversation. In *Psychology of Learning and Motivation*, Vol.62, ed. B.H. Ross San Diego, CA: Elsevier Academic Press, 59–99.

Clark, H. H. (1996). *Using language*. Cambridge university press.

Clark, H. H., & Marshall, C. (1978). Reference diaries. In D. L. Waltz (Ed.), *Theoretical issues in natural language processing*, TINLAP-2 (pp. 57–63). New York, NY: ACM. doi:10.3115/980262.980272

Clark, H. H., & Marshall, C. M. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge, England: Cambridge University Press.

Clark,H.H., & Wilkes-Gibbs,D. (1986). Referring as a collaborative process. *Cognition* 22, 1–39.doi:10.1016/0010-0277(86)90010-7

Creel, S. C., & Tumlin, M. A. (2011). On-line acoustic and semantic interpretation of talker information. *Journal of Memory and Language*, 65, 264–285.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language* 25: 141-168.

Eich, E. (1985). Context, memory, and integrated item/context imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 764–770.

Garrod, S. & Pickering, M. (2004). Why is conversation so easy? *Trends in Cognitive Science*. 8(1): 8-11.

Gorman, K. S., Gegg-Harrison, W., Marsh, C. R., & Tanenhaus, M. K. (2013). What's learned together stays together: Speakers' choice of referring expression reflects shared experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 843.

Hallett, P.E. (1986). Eye movements. In K.B. Boff, L. Kaufman & J.P. Thomas (Eds.), *Handbook of Perception and Human Performance I: Sensory Processes and Perception*. pp. 10-102. New York: Wiley.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49, 43-61.

Heller, D., Gorman, K. S., & Tanenhaus, M. K. (2012). To name or to describe: Shared knowledge affects referential form. *Topics in Cognitive Science*, 4, 290–305. doi:10.1111/j.1756-765.2012.01182.x

Heller, D., Parisien, C., & Stevenson, S. (2016). Perspective-taking behavior as the probabilistic weighing of multiple domains. *Cognition*, 149, 104-120.

Horton, W. S. (2007). The influence of partner-specific memory associations on language production: Evidence from picture naming. *Language and Cognitive Processes*, 22, 11141139.

Horton, W. S., & Gerrig, R. J. (2005a). The impact of memory demands on audience design during language production. *Cognition*, 96, 127142.

Horton, W. S., & Gerrig, R. J. (2005b). Conversational common ground and memory processes in language production. *Discourse Processes*, 40, 135.

Horton, W. S., & Slaten, D. G. (2012). Anticipating who will say what: The influence of speaker-specific memory associations on reference resolution. *Memory & Cognition*, 40, 113-126.

Kamide, Y. (2012). Learning individual talkers' structural preferences. *Cognition*, 124, 66–71.

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, 3238.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: theory and data. In The Psychology of Learning and Motivation, edited by Brian H. Ross, 59-101. *The psychology of learning and theory: Advances in research and theory*. San Diego: Elsevier.

Knutsen, D. & Le Bigot, L. (2012). Managing dialogue: How information availability affects collaborative reference production. *Journal of Memory and Language*, 67: 326-341.

Kronmüller, E., & Barr, D. J. (2007). Perspective-free pragmatics: broken precedents and the recovery-from-preemption hypothesis. *Journal of Memory and Language*, 56, 436–455.

Lewis, D. (1969). Convention: *A philosophical study*. Cambridge, MA: Harvard University Press.

Marno, H., Davelaar, E. & Csibra, G. (2014). Nonverbal Communicative Signals Modulate Attention to Object Properties. *Journal of Experimental Psychology: Human Perception & Performance,* 40(2): 752-62.

Marno, H., Davelaar, E. & Csibra, G. (2016). An object memory bias induced by communicative reference. *Acta Psychologica*, 163, 88-96.

Metzing, C., & Brennan, S. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, 49(2), 201–213. doi:10.1016/S0749-596X(03)00028-7.

Moll, H., & Tomasello, M. (2007). How 14- and 18-month-olds know what others have experienced. *Developmental Psychology*, 43(2), 309-317.

Mulligan, N. W. (2011). Conceptual implicit memory and environmental context. *Consciousness and Cognition*, 20, 737–744.

Peacocke, C. (2005). Joint attention: Its nature, reflexivity, and relation to common knowledge. In, N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds* (pp. 298–324). Oxford: Clarendon Press.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination. *Psychological Science*, 18, 407–413.

Richardson, D. C., Dale, R., & Tomlinson, J. M. (2009). Conversation, gaze coordination, and beliefs about visual context. *Cognitive Science*, 33, 1468–1482.

Ryskin, R.A., Wang, R.F., & Brown-Schmidt, S. (2016). Listeners use speaker identity to access representations of spatial perspective during online language comprehension. *Cognition*. 147: 75-84.

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: a review and meta-analysis. *Psychonomic Bulletin & Review*, 8, 203–220.

Stalnaker, R. C. (1978) Assertion. In: *Syntax and semantics, vol. 9: Pragmatics*, ed.P. Cole, pp. 315–32. Academic Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675–735.

Tomasello, M., Carpenter, M., Liszkowski, U. (2007). A new look at infant pointing. *Child development*. 78:705-32.

Trude, A. M., & Brown-Schmidt, S. (2012). Talker-specific perceptual adaptation during online speech perception. *Language and Cognitive Processes*, 27, 979–1001.

Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, 20, 580–591.

Yoon, S. O., & Brown-Schmidt, S. (2018). Aim low: Mechanisms of audience design in multiparty conversation. *Discourse Processes*, 55(7), 566-592.

Yoon, J. M. D., Johnson, M. H., & Csibra, G. (2008). Communication induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 13690–13695. doi:10.1073/pnas.0804388105.

Figures

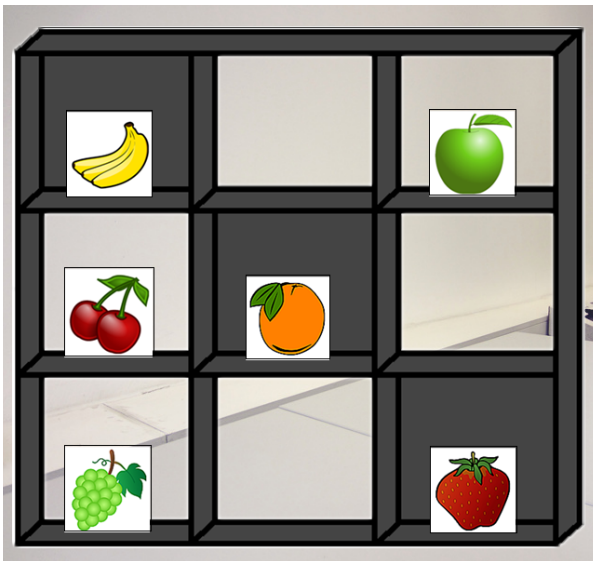
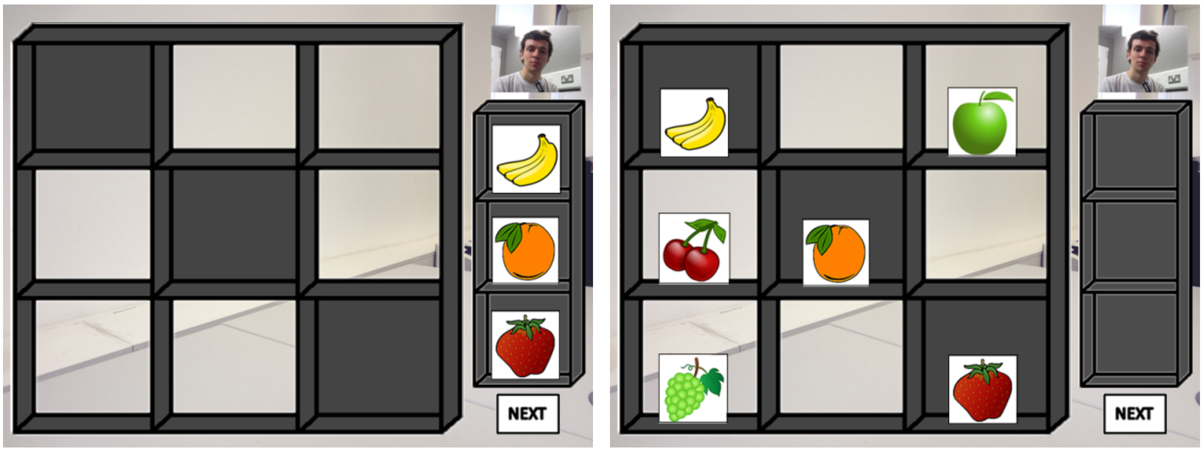


Figure 1.

Figure 2.



a. b.

Figure 3.

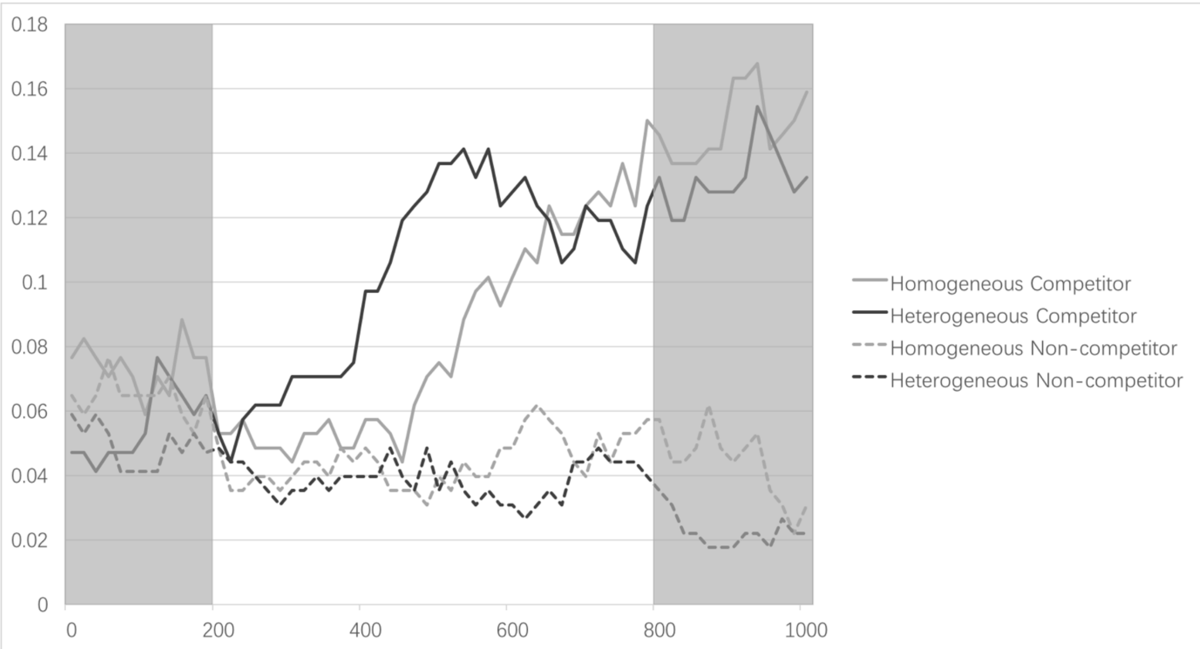


Figure 4.

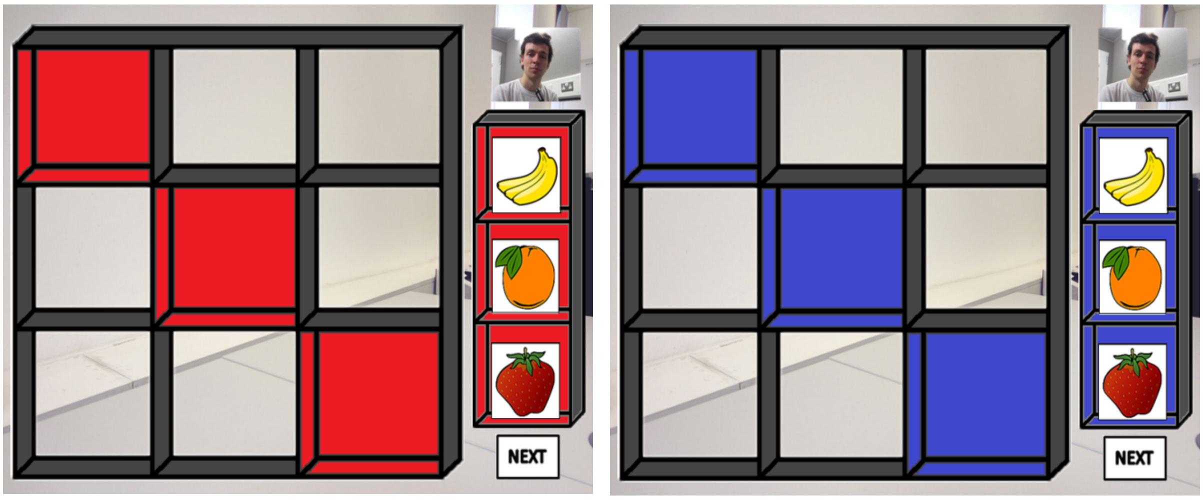
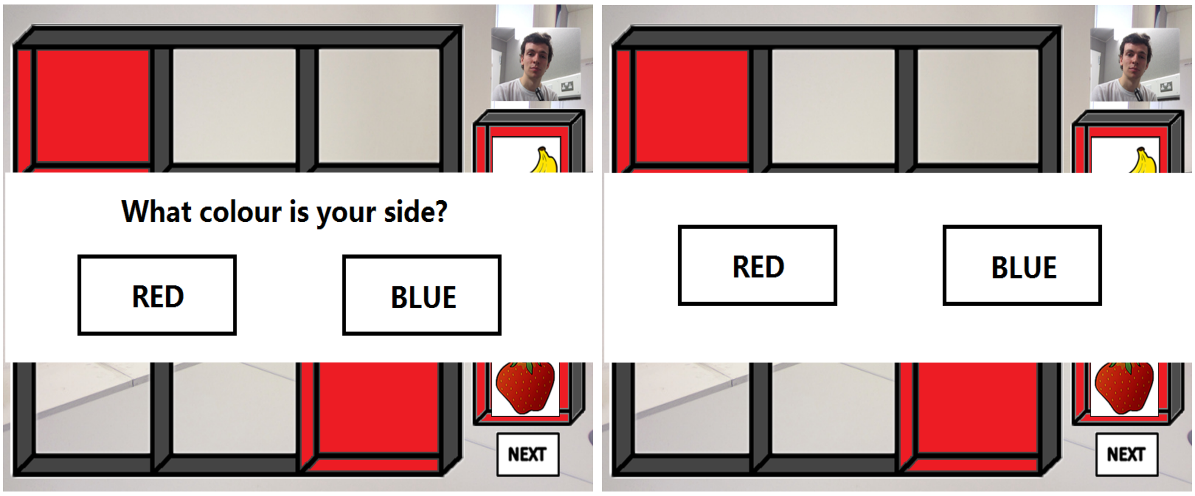


Figure 5.



a. Experiment 2

b. Experiment 3

Figure 6.

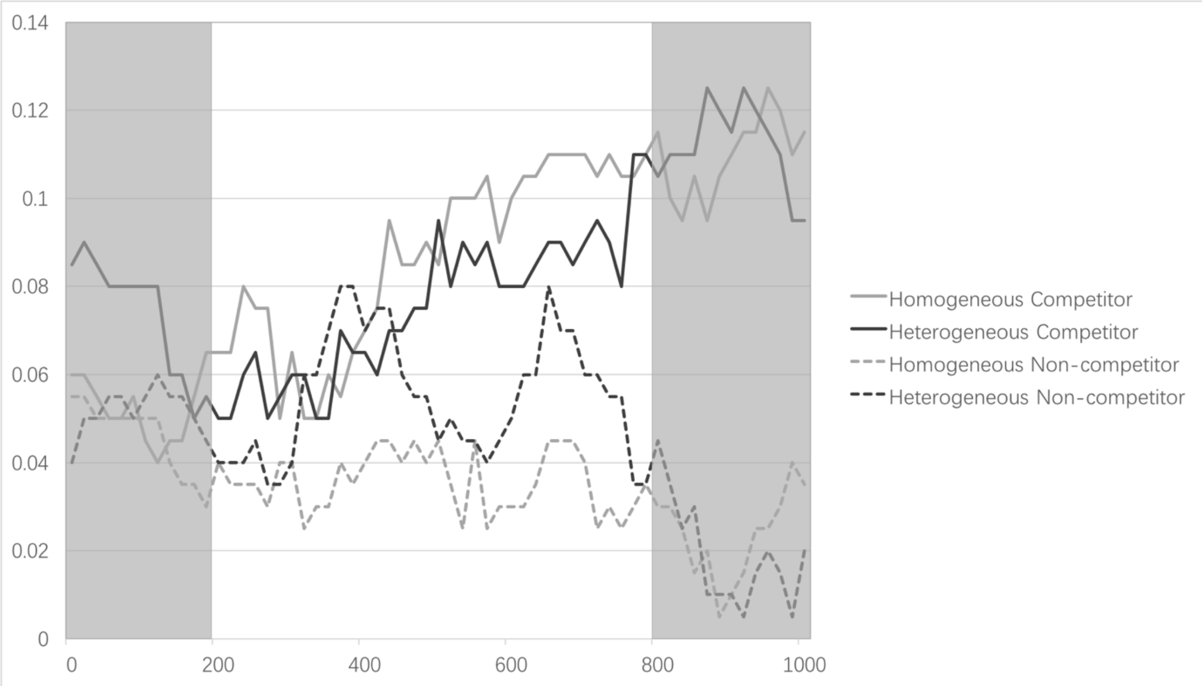


Figure 7.

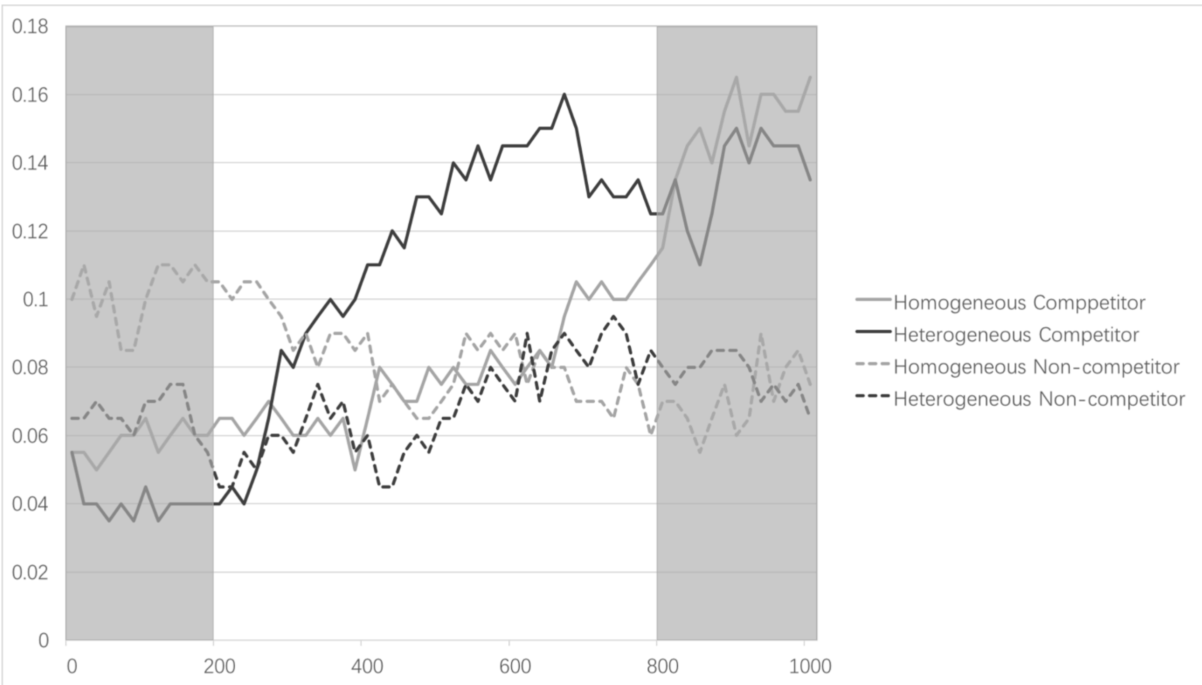


Figure 8.