How grounding and visual copresence affect reference reuse during dialogue

Dominique Knutsen (dominique.knutsen@univ-poitiers.fr), Christine Ros (christine.ros@univ-poitiers.fr) and Ludovic Le Bigot (ludovic.le.bigot@univ-poitiers.fr)
University of Poitiers & CNRS (CeRCA, UMR 7295), France
MSHS BÂT A5 – 5, rue Théodore Lefebvre – F86000 Poitiers, France

Abstract

The aim of this study was to examine reference reuse during dialogue. Dyads of participants discussed a route for an imaginary addressee. Half of the dyads actually navigated a real environment (i.e., their University campus) as they performed the task, which meant that they could rely on visual copresence as they interacted, whereas the other half could not, as only one of the participants found himself or herself in the task environment. The results show that reference reuse depends on who initially presented the reference, on whether it was accepted through repetition, and on the availability of visual copresence. These results are in line with the idea that among the references that belong to the common ground, some are more likely to be reused than others, due partly to the circumstances in which they were grounded.

Keywords: dialogue; collaboration; common ground; accessibility in memory; visual copresence

Introduction

During dialogue, at least two speakers interact in order to reach a common goal. As they do so, they increment their *common ground*, which includes the knowledge that they share and are aware of sharing (Clark, 1996; Clark & Marshall, 1981). The information contained in the common ground can then be reused by either speaker for coordination and adaptation purposes. The aim of the current study is to determine how the circumstances in which information is initially grounded – that is, added to the speakers' common ground – might affect its subsequent reuse by the speakers. In addition, because two speakers are not necessarily face-to-face as they interact, this study seeks to determine how visual copresence might also affect reference reuse.

Building up Common Ground in Dialogue

When two speakers interact, they are aware that they share a certain amount of knowledge. Part of this information constitutes their conversational background, or common ground. Before the interaction starts, the speakers already share a certain amount of knowledge due to past interactions, visual copresence or shared community membership (Clark & Marshall, 1981). Moreover, the common ground is incremented as the interaction unfolds, as the information produced by the speakers also becomes mutually shared (Clark & Schaefer, 1989).

According to Clark and Schaefer (1989), the common ground is incremented through a joint contribution process. During this process, one of the speakers produces, or presents, a piece of information with the intention that the other speaker will understand it. The second speaker then accepts this information by signaling his or her understanding either explicitly (e.g., by repeating the information presented or by saying "ok") or implicitly, by initiating the next relevant speech turn. This contribution process allows speakers to add to their common ground not only pieces of information, but also the references that they use to refer to things during the interaction (Clark & Wilkes-Gibbs, 1986).

Using Common Ground in Dialogue

Throughout the interaction, speakers consult their common ground in order to produce references that are adapted to their partner's level of knowledge. For instance, if Speaker A needs to refer an object for the second time during an interaction, he or she can consult the common ground to determine which reference was used the first time, reasoning that if Speaker B was capable of understanding this reference when it was produced for the first time, he or she should be capable of understanding it again (see Brennan & Clark, 1996).

However, among the references that belong to the common ground, all are not equally likely to be reused. For instance, Knutsen and Le Bigot (2012) found that references that had previously been added to the common ground (either implicitly or explicitly) were most likely to be reused when they corresponded to referents that were visible to the speakers at the time of reuse. Thus, although reference production during dialogue seems to be partner-adapted, such adaptation also depends on the state of mind of the speaker producing the references – in other words, adaptation is also partly egocentric (e.g., Keysar, 2007).

At this stage, one important question concerns how other features of the interaction might affect reference reuse as well. One first possibility is that the fact that both speakers take part in the contribution part has an influence on reference reuse. In a study conducted in the field of memory research, MacLeod (2011) found that individuals remember words better when they initially had the opportunity to produce these words aloud than when they heard these words being produced aloud by another person. Thus, whether speakers actually produce a reference during the grounding process could affect its subsequent accessibility in memory.

Visual Copresence in Dialogue

Another element which might also affect reference reuse is visual copresence at the time of the interaction. Indeed,

speakers do not always interact face-to-face: they sometimes need to rely on a communication medium. Using a medium to communicate has its advantages (such as allowing two speakers who are far away from each other to interact). However, it also constrains the interaction, insofar as the strategies available in face-to-face communication are not always available for the speakers to use in mediated communication (Clark & Brennan, 1991).

When two speakers interact over the phone, for instance, they cannot use visual cues such as their partners' facial expressions to monitor the latter's level of understanding. This can be problematic, as such cues help speakers to reach mutual understanding (Clark & Krych, 2004). Another problem could stem from the fact that one of the speakers might refer to objects that are not visible to the other speaker.

Importantly, speakers are capable of managing these additional constraints. For instance, in an experiment, Kraut, Fussell, and Siegel (2003) had dyads of speakers perform a collaborative task together. One speaker gave instructions to the other so that the latter could repair a bike. The authors manipulated visual copresence between speakers. One of the results was that when Speaker A did not have visual access to Speaker B's environment, the latter provided the former with descriptions about the objects that surrounded him or her so that A could nonetheless keep track of the task over time. Thus, one way for speakers to managing the constraints due to the media used is to adapt the content of their utterances, for instance by providing their partners with information that they would otherwise lack.

Current study

The aim of the current study is to offer a better understanding of reference reuse during dialogue. Accordingly with the points developed above, at least three features of the dialogue situation could affect the reuse that the speakers make of the references that belong to their common ground.

First, reference reuse might depend on who presented the reference in the first place. Indeed, because self-presented references are by definition self-produced, they should be more accessible in memory than partner-presented ones (see MacLeod, 2011).

Second, reference reuse might depend on how a presented reference was accepted. Indeed, as mentioned above, a reference can be accepted explicitly or implicitly. In the former case, acceptance sometimes involves repeating the presented reference verbatim (Clark & Brennan, 1991; Clark & Schaefer, 1989). In such cases, the speaker accepting the reference should benefit from the same self-production effect as the speaker presenting the reference. Thus, when a speaker accepts a reference by repeating it, he or she should be more likely to reuse it than if it has been accepted through other means.

Third, reference reuse might depend on whether or not speakers are visually copresent as they interact. Indeed, in the absence of visual copresence, speakers might reuse references to objects present in one of the speakers' environment more to compensate for the fact that these objects are not visible to the other speaker at the time of production. Thus, references should be more likely to be reused in situations where speakers are not visually copresent than in situations where they are.

To address these three points, an experiment was conducted where dyads of participants interacted in order to perform a collaborative route description task. As they performed the task, they had the opportunity to refer to a number of landmarks present on the route. Visual copresence was manipulated across dyads. Both experimental conditions involved navigating a real environment (i.e., the University campus). In one condition, both participants found themselves in this environment, whereas only one participant was in the environment in the other condition. The dialogues were transcribed and coded for reference presentation, acceptance and reuse. Thus, this experimental setting allowed examining reference reuse as a function of who initially presented the reference, how it had been accepted and whether the participants were visually copresent or not.

Method

Participants

Sixty students (eighteen male) from a French University initially volunteered to participate in the experiment. They all signed an informed consent form at the beginning of the experiment. The participants performed the task in dyads and knew each other prior to the experiment. However, the data from one dyad were discarded from the analysis because one of the participants in the dyad was not a native French speaker.

Apparatus

A microphone and a digital camera were used to record and to film the participants in the condition where the participants were visually copresent. In the condition where the participants were not visually copresent, a microphone was used to record the participants; one of the participants used a mobile phone and the other used a land phone.

Procedure and Task

The participants were told that their task was to interact in order to organize a visit of their University campus for foreign students who had never visited the campus before. They were also told that their route should include a fixed starting point (i.e., an amphitheater), a fixed intermediate stop (i.e., the University library) and a fixed arrival point (i.e., a dining hall).

In one of the conditions (the Visual Copresence Condition), a female experimenter met the participants at the laboratory and accompanied them to the amphitheater which served as starting point for the visit. There, she read out the instructions to them and told them that they should follow the route as they discussed it. She then filmed the

participants with the digital camera during the entire experiment.

In the other condition (the No Visual Copresence Condition), the experimenter also met the participants at the laboratory. Within the dyad, each participant was assigned one of two roles at random: the on-campus participant would walk the route during the conversation, whereas the remote participant would stay at the laboratory during the experiment. Both participants were informed that they would interact over the phone. The experimenter then read out the instructions and accompanied one of the participants to the starting point, while the other participant stayed at the laboratory with another experimenter.

In both conditions, there was no time limit to perform the task: the participants took as long as they needed to discuss the route.

Experimental Design

There were three independent variables in this study. The first one was the Condition (Visual Copresence Condition vs. No Visual Copresence Condition). This was a between-dyads IV. The second IV was Presentation (a reference could either be Self-presented or Partner-presented). This was a within-participants IV. The third IV was Acceptance (a reference had either been Accepted through repetition or Accepted through other means). This was also a within-participants IV.

Data Coding and Dependent Variables

The dialogues were transcribed. All of the landmarks (i.e., buildings and street names) introduced by the participants during the experiment were listed and numbered. In total, 79 different landmarks were mentioned by the participants.

Within each dyad and for each participant, each landmark was coded for presentation (which referred to the first time a reference was produced during the interaction), acceptance through repetition (which referred to verbatim repetition of a reference after its presentation) and reuse (which referred to all occurrences of reference production which could not be classified as presentations or acceptances). A dialogue sample and a coding example can be found in Table 1.

Table 1: Dialogue sample and coding example (the references being coded are italicized)

Dyad code	Condition	Speaker code	Utterance content	Coding
D07	Visual Copr.	B	so now we're arriving at *the laboratory*	Presented
D07	Visual Copr.	A	so *the laboratory*	Accepted through repetition
D26	No Vis. Copr.	A (remote)	on the left behind *the psychology building*	Presented
D26	No Vis. Copr.	B (on-campus)	we can walk behind *the psychology building*	Accepted through repetition
D14	A	Visual Copr.	this is *the laboratory*	Presented
D14	B	Visual Copr.	okay okay	Accepted through other means
D26	No Vis. Copr.	A (remote)	and then the hm *literature building*	Presented
D26	No Vis. Copr.	B (on-campus)	yes exactly	Accepted through other means

As regards presentation, when a reference to a landmark was produced for the first time during the interaction, it was coded 1 for presentation for its initiator and 0 for presentation for the other participant. This allowed determining, for each reference produced, whether it had initially been self- presented or partner-presented (from each participant's point of view).

To determine whether acceptance through repetition also caused references to benefit from a self-production benefit, the dialogues were also coded for such acceptance. A presented reference was coded A 1 for acceptance through repetition in the dyad when it was repeated by the other participant (i.e., the participant who did not present this reference) after its presentation and before its initiator produced another reference. In cases where a presented reference was accepted through other means than repetition, this reference was coded 0 for acceptance through repetition in the dyad.

Finally, reuse referred to all occurrences of reference production which could not be classified as presentations or acceptances. When a participant reused a reference, it was coded 1 for reuse for this participant. (How many times this reference was reused was not taken into account.) When a participant did not reuse a reference, it was coded 0 for reuse for this participant. In order to distinguish genuine reuse from mere repetition, a reference was only coded as reused if it was produced in a speech turn preceded by two speech turns which did not contain this reference. This level of coding was used at the DV in this experiment.

A reference could only be presented and potentially accepted through repetition once per dyad; however, any reference could be reused by both participants, by one participant only or by none of them.

Results

The task lasted 976 seconds on average in the Visual Copresence Condition (*SD* = 175 seconds) and 880 seconds on average in the No Visual Copresence Condition (*SD* =

253 seconds). The mean number of speech turns per dyad was 166.3 in the Visual Copresence Condition (*SD* = 83.57) and 166.9 in the No Visual Copresence Condition (*SD* = 78.45). In the Visual Copresence Condition, the participants presented a total of 307 references, 93 of which were accepted through repetition; the participants reused a total of 192 references in this condition. In the No Visual Copresence Condition, the participants presented a total of 278 references, 52 of which were accepted through repetition; the participants reused a total of 245 references in this condition.

A logistic mixed model including an identity variance covariance matrix was run in SPSS 20.0 to analyze reference reuse during the interaction (Baayen, Davidson, & Bates, 2008; Jaeger, 2008). This model allowed accounting for the fact that the participants were nested within dyads in this experiment. Because the number of landmarks produced differed between participants, the Satterthwaite correction was applied.

The random effects structure of the model included by-dyads and by-participants random intercepts to account for potential variability across dyads and across participants. By-dyads and by-participants random slopes corresponding to the main effects of Presentation and Acceptance were included to account for the dyads' and the participants' potentially different sensitivity to these IVs. By-random slopes corresponding to the main effect of Condition were not included in the model, as this was a between-dyads and between-participants IV (Barr, Levy, Scheepers, & Tily, 2013). Random slopes corresponding to the landmarks referred to during the experiment were included, as the participants might have differed in their sensitivity to these as well. The results are reported in Figure 1.

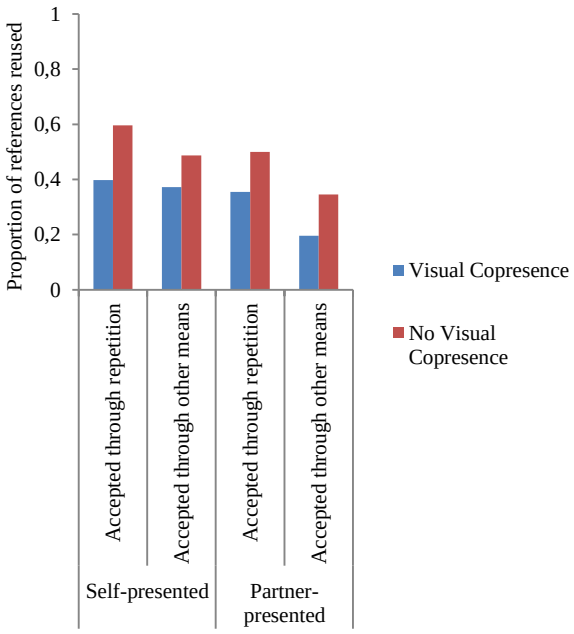

Figure 1: Proportion of references reused during the experiment.

The initial model included Presentation, Condition, Acceptance and all the interactions between these variables as fixed effects and reference reuse as the outcome variable. However, because none of the interactions reached significance, all interactions were removed from the final model, which is reported hereafter.

All three IV significantly predicted reference reuse. First, who initially presented a reference significantly predicted reference reuse, $F(1, 227) = 9.059$, $p = .003$. Self-presented references were more likely to be reused than partner-presented ones, $OR = 1.592$, $CI_{.95} = 1.174, 2.160$. Second, whether a reference was accepted through repetition significantly predicted reference reuse, $F(1, 255) = 7.258$, $p = .008$. References accepted through repetition were more likely to be reused than references accepted through other means, $OR = 1.575$, $CI_{.95} = 1.130, 2.194$. Finally, the condition in which the participants were in significantly predicted reference reuse, $F(1, 38) = 10.72$, $p = .002$. References were more likely to be reused in the No Visual Copresence Condition than in the Visual Copresence Condition, $OR = 1.902$, $CI_{.95} = 1.278, 2.831$.

Discussion

The aim of this study was to examine reference reuse during dialogue. The results obtained are in line with the three predictions formulated above. Indeed, they confirm that reference reuse depends on the circumstances in which to-be-reused references were initially grounded. First, self-presented references were more likely to be reused than partner-presented ones. This finding suggests that even though a reference, once presented, belongs to the partners' common ground, it nonetheless remains more likely to be reused by the speaker who produced it for the first time than by the other speaker later in the interaction.

However, self-presentation is not the only element that might affect reference reuse during dialogue. The results from this study also revealed that references accepted through repetition were more likely to be reused than references accepted through other means. So in a situation where Speaker A presents a reference for Speaker B to accept, not only is the latter less likely to reuse this reference than the former – this will be all the more true in cases where Speaker B accepts this reference by others means than verbatim repetition. In addition, the interaction between the presentation IV and the acceptance IV failed to reach statistical significance. This null finding does not allow concluding that repetition caused greater reuse of both self-presented references (which, if accepted, were necessarily accepted by the other participant) and partner-presented references (which, if accepted, were necessarily self-accepted); however, we cannot exclude the possibility that repetition benefited both the speaker repeating the reference and the speaker hearing his or her partner repeating it.

Moreover, the results confirmed that visual copresence also affects reference reuse. This is in line with Kraut et al.'s

(2003) suggestion that speakers adapt their speech depending on whether or not they can rely on visual copresence as they interact. In the current study, this led the participants to reuse references more often in the No Visual Copresence Condition than in the Visual Copresence Condition. We have suggested that this is due to dyads trying to compensate for the fact that one of the speakers could only rely on the dialogue history – and not on the memory of having seen the landmarks – to remember which references had been added to the common ground by reusing these references all the more. However, an alternative explanation might be that the participants in the Visual Copresence Condition needed to reuse references less, as they had others means to refer to landmarks at their disposal, such as pointing to landmarks, for instance. Note that both explanations are not necessarily incompatible with each other.

All in all, these results enrich the finding that among the references that belong to two speakers' common ground, all are not equally likely to be reused in the remainder of the interaction. This is consistent with the idea that reference production during dialogue is constrained by each speaker's own egocentrism (e.g., Keysar, 2007). Insofar as we have suggested – following MacLeod (2011) – that the influence of self-presentation and of repetition is due to the greater accessibility in memory of self-produced words, these results are also in line with the idea that adaptation is constrained by the accessibility in memory of the representations needed for adaptation (e.g., Horton, 2008). These results also shed further light on the idea that adaptation is shaped by whether the referents under discussion are mutually visible to the speakers or not during the interaction (Kraut et al., 2003). Moreover, in this experiment, each experimental condition involved a naturalistic interaction situation: in one case, the participants interacted face-to-face in a real environment, whereas in the other case, they interacted over the phone. This contributes to reinforcing the ecological validity of the findings reported above.

Acknowledgements

This work was conducted as part of Margot Chignac's Master degree. The authors thank her for collecting the data. The first author's PhD is supported by the Direction Générale de l'Armement and Région Poitou-Charentes.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390-412. doi:10.1016/j.jml.2007.12.005

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. doi:10.1016/j.jml.2012.11.001

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.). *Perspectives on socially shared cognition* (pp. 127-149). Washington: APA Books.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language, 50,* 62-81. doi:10.1016/j.jml.2003.08.004

Clark, H. H., & Marshall, C. R. (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: Cambridge University Press.

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science, 13,* 259-294. doi:10.1016/0364-0213(89)90008-6

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

Horton, W. S. (2008). A memory-based approach to common ground and audience design. In I. Kecskes (Ed.), *Intention, common ground, and the egocentric speaker-hearer* (pp. 189-222). Berlin/New York: Mouton de Gruyter.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59,* 434-446. doi:10.1016/j.jml.2007.11.007

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics, 4,* 71-84. doi:10.1515/IP.2007.004

Knutsen, D., & Le Bigot, L. (2012). Managing dialogue: How information availability affects reference production. *Journal of Memory and Language, 67,* 326-341. doi:10.1016/j.jml.2012.06.001

Kraut, R. E., Fussell, S. R., & Siegel, J. (2003). Visual information as a conversational resource in collaborative physical tasks. *Human-Computer Interaction, 18,* 13-49. doi:10.1207/S15327051HCI1812_2

MacLeod, C. M. (2011). I said, you said: The production effect gets personal. *Psychonomic Bulletin and Review, 18,* 1197-1202. doi:10.3758/s13423-011-0168-8