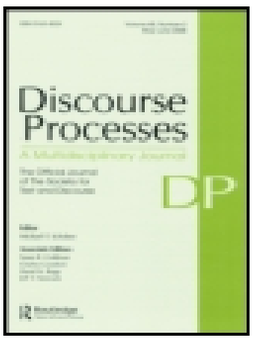

Discourse Processes

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hdsp20

Audience Design in Collaborative Dialogue between Teachers and Students

Christina Turner & Dominique Knutsen

To cite this article: Christina Turner & Dominique Knutsen (2021): Audience Design in Collaborative Dialogue between Teachers and Students, Discourse Processes, DOI: 10.1080/0163853X.2021.1904768

To link to this article: https://doi.org/10.1080/0163853X.2021.1904768

DISCOURSE PROCESSES
https://doi.org/10.1080/0163853X.2021.1904768

Routledge
Taylor & Francis Group

Audience Design in Collaborative Dialogue between Teachers and Students

Christina Turner[a] and Dominique Knutsen[b]

[a]Department of Psychology , University of Essex, UK; [b]CNRS,UMR 9193, Sciences Cognitives et Sciences Affectives (SCALab), University of Lille, Lille, France

ABSTRACT

When two people interact, reference presentation is shaped with the intention of supporting addressee understanding, allowing for ease of acceptance, thus minimizing overall collaborative effort. To date, analysis of such *audience design* has focused largely on adult–adult or adult–child interaction but seldom on adult–teenager interaction, including teacher–student interaction. An experiment was conducted in a British school in which teachers and students interacted to establish a reference for abstract tangram figures. Teachers were able to account for the students' increased ability to behave in a more adult-like collaborative way with dialogue features similar to those in adult–adult contexts. Set apart was dialogue with young students, where teachers continued to guide the interaction by producing lengthier descriptions and by encouraging participation. Dialogue with young students differs from that with other teachers in terms of the amount of effort put into the interaction and in how this effort is distributed and shared among dialogue partners.

Introduction

For a conversation to succeed it must be built on knowledge that dialogue partners share (see Clark, 1996). Indeed, communication issues may occur if the speaker makes an ambiguous reference choice such as "the canvas bag" or "that film showing at the cinema tonight." Both choices are open to misinterpretation; for instance, one person may use "the canvas bag" to indicate the shopping bag used for food, while another person may use it to describe his or her favorite bag used for personal belongings. Hence, dialogue partners must ensure that they share a similar interpretation of the references they use as they interact (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986; Fukumura, 2015). This background knowledge is referred to as *common ground*, which is defined as the knowledge that two dialogue partners (or more) share and are aware of sharing (Clark, 1996; Clark & Marshall, 1981; Stalnaker, 1978). According to Clark and Marshall (1981), speakers make assumptions about common ground based on linguistic co-presence (i.e., information mentioned during past interactions between both partners is deemed part of their common ground; e.g., Brenana & Clark, 1996; Clark & Wilkes-Gibbs, 1986), on non-inguistic features such as physical co-presence (i.e., information relative to the partners' environment is deemed part of their common ground; e.g., Clark & Krych, 2004), and on community co-membership (i.e., information that is usually shared within communities both partners belong to is also deemed part of their common ground; e.g., Fussell & Krauss, 1992; Isaacs & Clark, 1987; Lau et al., 2001).

The accumulation of linguistic common ground is a fundamentally collaborative activity (Clark & Brennan, 1991; Clark & Schaefer, 1989; Clark & Wilkes-Gibbs, 1986), whereby one speaker starts by

CONTACT Dominique Knutsen dominique.knutsen@univ-lille.fr; dom.knutsen@gmail.com Laboratoire SCALab UMR CNRS 9193, Université de Lille, Rue du barreau, BP 60149, Villeneuve d'Ascq Cedex 59653, France.

presenting a piece of information (such as a reference). At this point, the speaker may use strategies such as the use of the indefinite article to signal uncertainty (e.g., "*a* cat"), adding a try-marker to indicate a provisional reference (e.g., "*sort of* cat"), and, as dialogue progresses toward an agreed perspective regarding the referent, the use of less descriptive language alongside definiteness (Clark & Wilkes-Gibbs, 1986; Sacks & Schegloff, 1979). In return, the addressee *accepts* the information presented by producing evidence that it was understood well enough for current purposes (Clark & Brennan, 1991). Strong evidence is provided when the addressee paraphrases what he or she understands, presents an acknowledgment of understanding such as "yep," or when he or she repeats the phrase verbatim (e.g., by restating steps in a recipe or repeating the digits of a phone number; e.g., McInnes & Attwater, 2004). The addressee might also signal understanding by initiating the next relevant speech turn demonstrating his or her readiness to move on (Clark & Schaefer, 1989). If the addressee is having difficulty in understanding the speaker, the acceptance phase will begin with the communication of this and will only end once the speaker and the addressee have established that they have finally managed to understand each other. In sum, a piece of information must have been accepted to be deemed part of the partners' common ground, thus highlighting the central importance of the feedback provided by the addressee in the establishment of mutual comprehension.

When both partners are engaged in these processes of presentation and acceptance, it is expected that collaborative effort is minimized, that is, that the total amount of effort put into the presentation and acceptance stages by the dyad remains as low as possible (this is referred to as the least collaborative effort principle; Clark & Brennan, 1991; Clark & Schaefer, 1989; Clark & Wilkes-Gibbs, 1986). This prompts people to present references that can be understood easily by their current partners based on the common ground they have accumulated previously, a process known as audience design (Clark, 1996; Clark & Murphy, 1982). Indeed, references that are easy to understand can be accepted immediately (or at least without having to go through a costly, extensive repair process), thus reducing both partners' overall efforts to reach mutual comprehension. However, the way in which collaborative effort is divided among dialogue participants depends on a number of factors, among which the perceived ability to collaborate. In a study on spatial dialogue conducted by Schober (2009), participants whose spatial abilities were high or low talked about the location of various objects in a display. The analysis of the descriptions produced revealed that participants with high spatial abilities tended to produce and interpret utterances based on their low spatial ability partner's perspective, thus increasing the cognitive costs associated with presentation and acceptance for themselves while reducing these costs for their partner. In other words, the effort required to reach mutual comprehension is not always taken on by the person presenting the information; this mainly depends on the characteristics of the people engaged in the interaction.

Presentation and acceptance were illustrated in an influential study by Clark and Wilkes-Gibbs (1986) where dialogue partners completed a *matching* task using abstract tangram figures (see also Bortfeld & Brennan, 1997; De Ruiter et al., 2012; Horton & Gerrig, 2002; Isaacs & Clark, 1987; Knutsen et al., 2019; Krauss & Weinheimer, 1966; Schober & Clark, 1989; Wilkes-Gibbs & Clark, 1992). In this experiment, one participant (the matcher) placed tangrams in a predefined order that only the second participant (the director) had knowledge of. The director and matcher interacted freely to complete this activity and repeated it six times (the same pictures were used all six times, but the order in which they were presented was different each time). The authors found that on trial 1 the directors presented an extended description of the figures such as "looks like a person who's ice-skating, except they're sticking two arms out in front"; trial 1 descriptions also included a high proportion of indefinite references and try markers to guide the addressee's interpretation of the reference. While descriptions over trials simplified, this initial acceptance of the presentation became the foundation for further references, with the description on trial 2 being "the person ice skating that has two arms" and the description on trial 6 being "the ice skater." This highlights that once a *conceptual pact* (i.e., a partner-specific, temporary agreement as to how to refer to something; see (Brennan & Clark, 1996; Metzing & Brennan, 2003) is established, speakers will rely on this in later conversation to present references which can be accepted easily, that is, to engage in subsequent audience design.

The nature and timing of audience design has been the focus of intense debate. Whereas initial models of audience design suggested that this process requires speakers to explicitly represent their current partner's state of mind, more recent approaches have attempted to understand how the potentially high cognitive costs incurred by audience design may be balanced with the demand for rapid speech production (e.g., Brennan & Hanna, 2009; Horton & Keysar, 1996; Pickering & Garrod, 2004). For instance, research conducted within the framework of the memory-based approach to dialogue has shown that the mere presence of a given dialogue partner is sufficient to boost the accessibility in memory of information mentioned during previous interactions with that partner (and that both partners are thus likely to be able to understand easily; Horton & Gerrig, 2002, 2005a, 2005b, 2016). However, Ferreira (2019) has recently suggested that whereas some aspects of audience design such as syntactic or lexical priming (i.e., the repetition of syntactic structures and/or words during dialogue, which potentially makes utterances easier to understand for one's dialogue partner) are based on automatic priming processes, other aspects of audience design such as speaking differently to children and adults requires cognitive control. Some situations may even require speaker to predict in advance the communicational consequences of the utterances they plan on producing, thereby also increasing the cognitive costs associated with audience design.

Adult and young child dialogue pairs

It is important to highlight that although the studies listed above focus mainly on audience design in dialogue among adult speakers, numerous studies have focused on the features of audience design in adult–child dialogue (e.g., Bates & Silvern, 1977; Clark, 2010, 2015; Clark & Bernicot, 2008; Clark & Estigarribia, 2011; Epley et al., 2004; Nadig & Sedivy, 2002; Ntsame-Mba & Caron, 1999; O'Neill, 1996; Sachs & Devin, 1976; Shatz & Gelman, 1973; Warden, 1976). When adults are establishing common ground with young children, the status of the addressee (child) is immediately available, allowing for audience design. One strategy, widely used by adults, is the use of repeats to support younger children to be better understood and older children to recognize common ground (e.g., (Clark & Bernicot, 2008; Clark & Estigarribia, 2011). Adults adapt other introductory techniques for new vocabulary, depending on the age of their addressee. For younger children, adults consistently use devices such as final-position placing for the new learning (e.g., What's the boy doing? Is the boy *dancing*?) to elevate importance. Adults are less consistent in using this technique when talking to older children. One reason cited for this is that by age 4 years children have increased influence over the exchange, contributing more and sometimes offering alternative references (Clark, 2010). Audience design features are therefore affected by the age of the child and their ability to provide feedback to their adult partner.

For audience design to be successful, adults and children must be able to jointly collaborate to establish a reference through presentation and acceptance, with the addressee's response incrementally influencing the development of mutual understanding (Clark & Schaefer, 1989). The age at which children are able to take the perspective of another is key to this. Golinkoff (1986) has suggested that preverbal infants appear to learn how to establish a shared reference by initiating negotiation episodes using communicative signals rather than signals pertaining to emotion. Further research concludes that children as young as 2 years old are able to appreciate what is and is not shared knowledge (Clark, 1997, 2007; Ganea & Saylor, 2007; O'Neill, 1996). They are able to track prior conversation with an adult, making inferences about what is shared, and can use this to resolve ambiguity (Ganea & Saylor, 2007).

Children then become increasingly better at taking other people's perspective as they interact, even though a number of differences remain between children and adults. For instance, Matthews et al. (2010) found that whereas adults understand that conceptual pacts are partner-specific (i.e., that an old dialogue partner should not use a new reference when referring to an object that has already been mentioned before but that a new reference might be used by a new dialogue partner), 3- and 5-year-old children do not. Focusing more on language production, Warden (1976) suggested that children under 5 years old are unable to take into account shared knowledge and can only inconsistently do so between the ages of 5 and 9. A study by Branigan et al. (2016) provides similar evidence, showing that

children are able to use shared knowledge to tailor their utterances to a specific partner but not in a fully adult-like manner (e.g., they did not seem to use definiteness to indicate common ground, as adults usually do). In contrast, Grigoroglou and Papafragou (2016) reported that 4- and 5-year-old children are capable of engaging in audience design but only when their partner's dialogic needs are particularly salient (e.g., when the partner expresses his or her incomprehension explicitly).

Teacher and student dialogue pairs: overview and goal of the current study

In sum, the manner in which adults engage in audience design during dialogue is well documented. However, although many studies have focused on audience design when adults address young children (and when young children address adults), dialogue between adults and teenagers has seldom been investigated. More precisely, two studies have examined cases where teenagers address adults. Fukumura (2016) found that when producing descriptions for adult addressees, 11- to 16-year-old teenagers are able to produce more disambiguating information than 6- to 10-year-old children. However, teenagers and children produced a similar amount of privileged information (i.e., information unknown to their adult addressee). Thus, teenagers may well have adult-like skills in ambiguity avoidance but have yet to fully master audience design. A similar conclusion was drawn by Ntsame-Mba and Caron (1999), who examined audience design in children aged between 7 and 14 years old. They found that whereas 7-year-old children find it difficult to build common ground and to subsequently engage in audience design, 14-year-old teenagers are almost as efficient as adults when it comes to establishing mutual comprehension. For instance, 14-year-old teenagers manage definite references (which may be used to signal common ground) better than 10- and 7-year-old children, although they remain less efficient than an adult control group. Taken together, these two studies suggest that teenagers interact with adults more efficiently than do children. However, to our knowledge, the opposite situation (i.e., a setting in which in adult addresses a teenager) has not been examined yet.

These findings suggest that teenagers progressively acquire the ability to manage an interaction in a collaborative way. But are adults aware of this when they interact with teenagers? The current study seeks to shed some light on this question by focusing specifically on teacher–student dialogue settings. From a theoretical perspective, this question is particularly important because such awareness could affect how collaborative effort is divided in the adult–teenager pair. In an adult–teenager interaction, the adult believing that the teenager is unable to collaborate could lead the adult to systematically take on most of the dyad's collaborative effort (for a similar rationale, see Schober, 2009), even when this is no longer necessary. The ability to recognize and acknowledge teenagers' increasing collaborative dialogic skills and to encourage the teenager to play an active role in collaboration is particularly important in real-life situations such as teacher–student interactions. Indeed, studies have already highlighted the role of dialogue in teaching (see Calcagni & Lago, 2018) and, more specifically, the importance of establishing and maintaining common ground in the classroom. For instance, Alibali et al. (2013) demonstrated that mathematics teachers spontaneously gesture more when they believe they have reached a "trouble spot," that is, that they have "lost" common ground with their students and that it must be reestablished before the dialogue may resume. However, the current study is one of the first to date to specifically examine audience design in a teacher–student dialogue setting.

The first goal of the current work is thus to answer the question of whether teachers interacting with students take into account the fact that their conversational partner is capable of behaving in an increasingly collaborative way by offering a more systematic description of the features of dialogue in situations where a teacher interacts with a student. The second goal is to compare audience design in dialogues with 11- to 12-year-old students on the one hand and with 16- to 17-year-old students on the other hand to capture the teachers' potential ability to take into account the fact that collaboration becomes more efficient during adolescence. In the control group, teachers interacted with other teachers.

Finally, the third goal of this work is to examine audience design outside the laboratory in a more ecological environment than that offered by controlled laboratory settings. We thus conducted our experiment in a British high school in which teachers and students performed a matching game together (in the control condition, the experiment was performed by pairs of teachers). In this kind of experiment, dialogue partners have the opportunity to build common ground around the pictures used in the task; they can then resort to this common ground to make dialogue easier in subsequent trials (Clark & Wilkes-Gibbs, 1986).

The general hypothesis tested in this study was that teachers use information about the community membership (age and school year group) of the addressee to inform audience design. Central to this is the choice of school year group, with (in the British system) year 7 students being 11- to 12-years-old and the youngest children in the secondary school and year 12 students being 16- to 17-year olds transitioning to adulthood.[1] To test the hypothesis, participants worked in pairs to complete a matching task involving abstract pictures. These pairs always involved one teacher and either another teacher (staff–staff condition [ST/ST]), a year 12 student (staff–year 12 condition [ST/Y12]), or a year 7 student (staff–year 7 condition [ST/Y7]). Participants interacted freely, stimulating collaboration on reference identification for each abstract picture. It was expected that teachers will perceive year 7 students to need the most support, year 12 students less support, and other teachers very little support when establishing references. This should lead teachers to produce more indefinite references (e.g., *a* cat rather than *the* cat), more try-markers (*a kind* of cat rather than a cat), and to provide lengthier descriptions in the first trial of the matching task when interacting with year 7 students than when interacting with year 12 students and when interacting with year 12 students than when interacting with adults (i.e., other teachers). What is more, the decrease in the number of indefinite references, try-markers, and words produced usually observed over trials in this kind of task should be slower when interacting with year 7 students than when interacting with year 12 student, and when interacting with year 12 students than when interacting with other teachers.

Although the main hypothesis of this study focused on the features of the teachers' speech, we also examined the features of the students' speech. In particular, we examined the contribution of the students to the interaction. The hypothesis was that year 7 students would contribute less to the establishment of shared references (i.e., that they would produce fewer words) than year 12 students and that year 12 students would contribute less than teachers. This expected effect could also be explained in terms of the perceived power and social distance of the teacher, perhaps making younger students less likely to influence the conceptual pact (e.g., by accepting the references presented faster, without asking additional questions, for instance), thus contributing less to the exchange (Brown & Levinson, 1987).

Methods

Participants

The experimental protocol described hereafter was approved by the Psychology Ethics Committee of the University of Essex. Participants (all native English speakers) were teachers or students at a large secondary school in England. Prior to the experiment, parents or guardians of year 7 and year 12 students received a letter providing them with general information about the study. Also included in the letter was specific information so parents knew prior to consent that the experiment entailed their child working alongside a teacher in a pair to complete a matching activity. It was also made clear that this activity would last for no longer than 45 minutes. Attached to the letter was a consent form giving further information regarding confidentiality arrangements. Consequently, parents were aware that although the interaction between the child and teacher would be recorded (audio only), the child's name would not be mentioned and the child's data (age and gender) would be stored anonymously. Parents or guardians signed the form to indicate they understood the general purpose of the experiment, that their child's voice would be recorded, that their child was free to leave the study at any time,

that their child would be debriefed at the end of the experiment, and that after the experiment any further questions could be asked. All year 7 and year 12 participants returned a signed consent form. The teachers who participated in the experiment were provided with the same information (also in a letter with attached consent form), although this referred to their own involvement as a participant. All teacher participants signed this consent form.

Forty teachers (26 women; average time spent in teaching 13.68 years, *SD* = 8.90 years) were randomly selected to participate in the study. Ten year 7 students (5 girls; average age 11.73 years, *SD* = 0.47 years) were randomly selected from all those who had parental consent to participate in the study (one additional ST/Y7 pair was initially recruited, but their results were not included in the data analysis, as one of the participants was not a native English speaker). Ten students in year 12 (8 girls; average age 16.80 years, *SD* = 0.42) were mainly recruited face-to-face at the start of the data collection period through advertisement in lessons and parental consent forms gathered on arrival for the experiment. In addition, we made sure that the students recruited had no history of spoken language impairment. All participants were offered a full debrief on completion of the experiment, although not all accepted this. Where a debrief took place, the participant was informed of the general hypotheses; they understood that the study was exploring the features of adult–teenager interaction and whether differences between year groups were found.

Each pair of teachers reported the regularity of their conversations with each other over a school term using a 4-point scale (0–3). A higher score represented increased frequency of conversation. On average, they reported rarely conversing with each other (*average* = 0.80, *SD* = 0.79). To ascertain whether the teacher knew the student in their pair, they reported whether they had taught the student before and for how long. Fifty percent of the ST/Y12 dyads reported they had never taught their student-partner. Only one remaining year 12 student had been taught for longer than 1 year by their teacher-partner. In the ST/Y7 dyads only one year 7 student had been taught by the teacher in their pair, but this was for less than a year. In general, then, the participants within the dyads did not regularly converse or know each other.[2]

Materials and apparatus

A pool of 24 tangram pictures were randomly selected for use in the experiment, and these were randomly divided into two sets (sets A and B; see Figure 1 for an example). Within each set, 12 pictures were randomly arranged in each of six grids (a different order was used in each grid), comprising four rows and three columns, and printed onto single sheets of A4. Each position in the grid was numbered in the bottom right-hand corner from 1 to 12. The same grid without the tangram pictures and

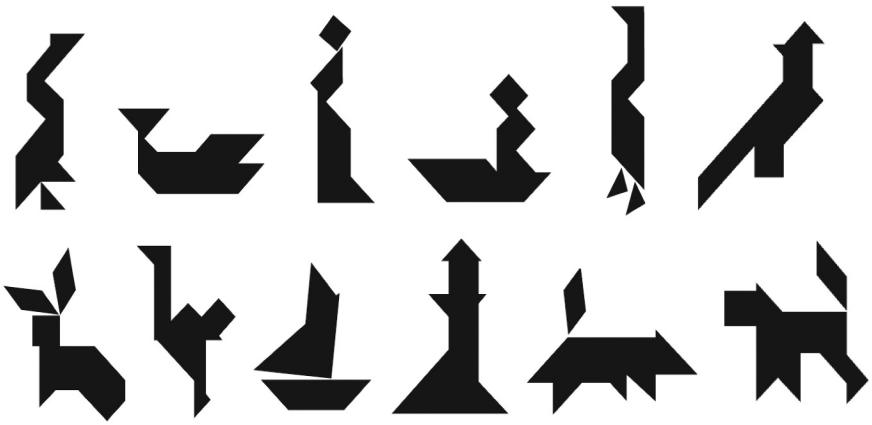

Figure 1. Tangram figures, set A.

numbers was also printed, providing a blank version. The 12 tangrams used were made into separate cards, and a dot sticker was placed in the top right-hand corner of each card to indicate the orientation of the picture. In each of the three dyad conditions, five pairs used set A and five pairs used set B to complete the experiment.

Procedure

Participants played a referential communication task (Clark & Wilkes-Gibbs, 1986; Krauss & Weinheimer, 1966). The experiment was undertaken in a quiet room in the secondary school to which all participants belonged. Before the beginning of the experiment, participants were informed that the study sought to investigate human dialogue and so they would have an opportunity to talk with one another during the experiment. Within each dyad, one person was allocated the role of director and the other person the role of matcher. In the ST/ST dyads the roles were randomly allocated. In the ST/Y12 and ST/Y7 dyads the teacher was always the director and the student was always the matcher. Each pair sat facing away from each other so they were unable to see each other's materials or actions. The participants were requested to remain in these positions throughout the experiment to ensure that they did not look at each other's grids or at each other during or between trials.

The director was given the first grid containing the 12 tangram figures. The matcher was given separate cards of the same 12 tangram figures along with the blank grid. The goal of the task was explained, so that both participants knew that the matcher's aim was to place the tangram figures in the grid so that it matched the grid of the director. The participants were told that the director should give the matcher instructions to allow him or her to place each card in the grid to achieve the goal. Before starting the experiment, the participants were told that to solve each puzzle, they were free to interact as they liked. They were also told that after each trial, they would be informed of the number of mistakes they had made but any further information (e.g., information about which cards had not been placed correctly) would not be provided. Once each trial was complete, the director was given the next grid that contained the same pictures but in different places. This was repeated over six trials.

Experimental design

The first between-participants independent variable was dyad condition split into three levels: ST/ST, ST/Y7, and ST/Y12. The second within-participants independent variable was trial number comprising six levels from trial 1 to trial 6. Trial number was treated as a continuous variable and was centered for the purpose of the statistical analyses.

A first set of dependent variables (DVs) quantified the amount of effort put into the task by directors and matchers: the number of words produced by the director, the number of words produced by the matcher, the number of speech turns produced by the dyad, the time taken to complete the task (in seconds), and the number of mistakes made by the matcher. Moreover, as explained below, we found while analyzing the data that the director often encouraged the matcher to describe the picture him- or herself. A second set of (binary) DVs was thus defined (ad-hoc) to account for this: the probability of the reference used to describe the figure being introduced by the director rather than the matcher and the probability of the director encouraging the matcher to describe the picture. A third set of (binary) DVs was used to characterize the director's speech: the probability that the reference produced by the director would contain a try-marker and the probability of the director's reference being indefinite. A final (binary) DV was used to characterize the matcher's speech: the probability that the matcher would ask the director a clarification question. More detail about the coding scheme used is provided in the following section.

Data coding

The interactions between participants were transcribed and then coded following the procedure described hereafter.

Amount of effort put into the task. During the coding process, the number of words produced by the director, the number of words produced by the matcher, and the number of speech turns produced by the dyad were noted. All words and speech turns produced by the participants while performing the task were included in this count. The only words and speech turns that were not included corresponded to occurrences in which the participants addressed the experimenter (e.g., to ask a question about the experiment). We also noted the time length of each trial along with the number of mistakes made in each trial. A mistake was made each time one of the pictures was not placed correctly in the matcher's grid. Mistakes were counted after the end of each trial, such that temporary mistakes (i.e., mistakes made by the matcher during the trial which were corrected before the end of the trial) were not counted.

Additional variables: reference introduction. In some cases the reference used to describe the picture was introduced by the matcher rather than the director. Furthermore, when this was the case, the matcher had sometimes been encouraged to do so by the director, as shown in the following examples (in Example 1, the matcher was encouraged by the director to produce a reference but not in Example 2):

Example 1 (Dyad 2, Trial 1 – ST/Y7 condition)

Director: and then what does your final one look like
Matcher: erm it looks kind of like a bird
Director: like a bird has it got one foot
Matcher: yep
Director: good

Example 2 (Dyad 30, Trial 1 – ST/ST condition)

Matcher: right yeah got one that looks like a dog
Director: erm that is position 12 so bottom row furthest right

Two DVs were coded to account for this phenomenon: whether the reference has been introduced by the director or not (this variable was coded 1 when the director produced the reference and 0 when the matcher produced the reference) and whether the director encouraged the matcher to produce a reference or not (this variable was coded 1 when the director encouraged the matcher and 0 when he or she did not).

Importantly, in some cases more than one reference was produced (by one of the participants or by both participants) to describe the picture, as in Example 3 below. In such cases, only the first reference mentioned was taken into account in the coding of the current trial. For instance, in Example 3, the director referred to the picture as a tower, whereas the matcher referred to it as a lighthouse. In this case, only the first reference (i.e., the big tower) was taken into account in the coding: This reference would have been coded as presented by the director, and the matcher was not encouraged to present a reference in this trial.

Example 3 (Dyad 18, Trial 1)

Director: because the one that is below the rabbit looks a bit like I don't know it there's a triangle a really big base of a triangle erm and erm it looks like a big tower

Matcher: like a lighthouse
Director: yes
Matcher: yeah

Characterization of the director's speech. In each trial, the reference produced by the director was coded as indefinite or as "other." Indefinites were references that were preceded by the determiner *a* or *an*, such as "*a* person sitting down." The "other" category included all other references: definite (e.g., "*the* person sitting down"), possessive (e.g., "*my* person sitting down"), or demonstrative (e.g., "*this* person sitting down"). This category also included cases where no determiner was produced by the participants (e.g., "person sitting down"). This variable was coded 1 when the director used an indefinite reference and 0 when he or she did not use an indefinite reference. The reference produced by the director was also coded as including a try-marker such as "maybe," "a kind of," or "a bit like" (e.g., "*a bit like* a Siamese cat"). This variable was coded 1 when the director used a try-marker and 0 when he or she did not use a try-marker. As with the DVs presented below, only the first reference presented by the director was coded by reference indefiniteness and for the presence of try-markers.

Characterization of the matcher's speech. The content of the feedback produced by the matcher was examined to identify cases where he or she asked the director a clarification question. This coding was used when the matcher requested further information before agreeing the reference later in the exchange, as shown in Example 4.

Example 4 (Dyad 25, Trial 4)

Director: 7 is the sharky fish
Matcher: the one with the nose lower
Director: that's it

Only the feedback following the first presentation of a reference in each trial was included in this coding. This variable was coded 1 when the matcher requested further information and 0 when he or she did not.

Double coding. The data from six dyads (one-fifth of the dataset) were double-coded. The inter-rater agreement was high: Both coders agreed on the coding 95.24% of the time. Specifically, Cohen's kappa was calculated for five of the DVs presented below: the probability of the director introducing the reference ($\kappa = 0.97$, almost perfect agreement), the probability of the director encouraging the matcher to introduce the reference ($\kappa = 0.87$, almost perfect agreement), the probability of the director producing a try-marker ($\kappa = 0.76$, substantial agreement), the probability of the director producing an indefinite reference ($\kappa = 0.79$, substantial agreement), and the probability of the matcher asking a clarification question ($\kappa = 0.71$, substantial agreement).[3] All disagreements were solved through discussion. Because the agreement between coders was at least substantial in all cases, all remaining data were single-coded.

Results

The corpus generated in this study included 7,758 speech turns. A total of 51,087 words were produced (40,838 words were produced by the directors and 10,249 words were produced by the matchers). The participants successfully completed the task in 95.56% of trials (96.67% of trials in the ST/ST condition, 96.67% of trials in the ST/Y12 condition, and 93.33% in the ST/Y7 condition).

The data were analyzed in SAS University Edition (GLIMMIX procedure). Mixed models (see Barr et al., 2013) were used to analyze the data. Specifically, linear mixed models were used in cases where

the DV was continuous, and generalized linear mixed models (logistic models; Jaeger, 2008) were used in cases where the DV was binary.

Mixed models were used to account for by-unit variability through the inclusion of random intercepts and random slopes (Barr et al., 2013). The units considered in the analyses reported below were the participants and the items (i.e., the pictures used in the matching task). In the current analysis, random intercepts and slopes were only included in the analyses if their contribution to the models was significant. If they were not, they were removed from the model (doing so does not significantly affect the outcome of the model; the identification of random intercepts and slopes that do not significantly contribute to the model is performed automatically in SAS; Kiernan et al., 2012). What is more, in some cases the participants did not discuss the final image in the display, which meant that the number of observations could be slightly different across dyads. The Satterthwaite correction was used to account for this by correcting the degrees of freedom (Keselman et al., 1999).

The analyses on the first two sets of DVs (amount of effort put into the task and reference introduction) were conducted on the entire dataset. The analyses on the last two sets of DVs (characterization of the director's speech and of the matcher's speech) were only conducted on trials where the reference was presented by the director. Importantly, although we had initially planned on including an analysis on the number of mistakes made by the matcher in each trial, the number of mistakes made was too small for us to conduct the analysis we had planned.

Amount of effort put into the task

Effect of condition and trial number on the number of words produced per picture by the director

The data corresponding to this analysis are shown in Figure 2. The random-effects structure of the model used to analyze the data (as well as the random-effects structure of all models used in the analyses reported hereafter) is shown in Table 1. The parameters of the model are shown in Table 2. A significant effect of condition was found, $F(2, 27) = 3.80$, $p = .035$. An inspection of the b coefficients revealed that directors in the ST/Y7 produced more words than directors in the ST/ST condition. There was no significant difference between the ST/Y12 condition and the ST/ST condition. An additional pairwise comparison (Bonferroni-corrected) revealed no significant difference between the ST/Y12 condition and the ST/Y7 condition, adjusted $p = .216$. A significant effect of trial number was also found, $F(1, 29) = 96.98$, $p < .001$. An inspection of the b coefficient revealed that directors produced fewer words over trials. Finally, there was a significant condition × trial number interaction, $F(2, 27) = 4.38$, $p = .023$. An inspection of the b coefficients revealed that the decrease in the number of

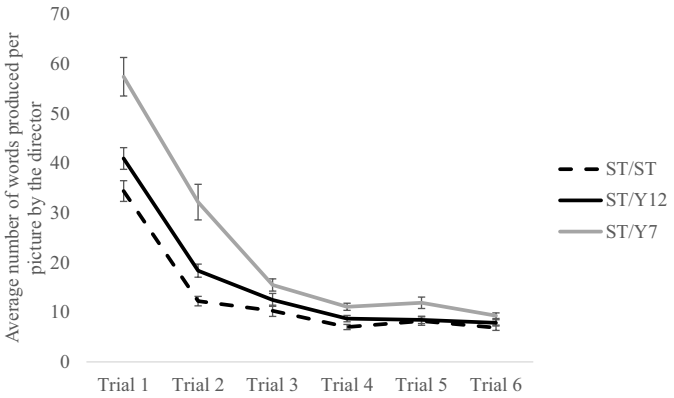

Figure 2. Average number of words produced per picture by the director as a function of condition and trial number. The bars represent the standard error.

Table 1. Random effects structure of the models used to analyze the data

Analysis conducted	Random slopes included	Random intercepts included
Effect of condition and trial number on the number of words produced per picture by the director	By-dyad and by-item slopes corresponding to trial number	By-dyad and by-item intercepts
Effect of condition and trial number on the number of words produced per picture by the matcher	By-dyad and by-item slopes corresponding to trial number	By-dyad and by-item intercepts
Effect of condition and trial number on the number of speech turns produced per picture by the dyad	By-dyad slopes corresponding to trial number	By-dyad and by-item intercepts
Effect of condition and trial number of the time taken (in seconds) to complete each trial	By-dyad slopes corresponding to trial number	By-dyad intercepts
Effect of condition and trial number on the probability of the director presenting the reference	By-dyad slopes corresponding to trial number	By-dyad intercepts
Effect of condition and trial number on the probability of the director encouraging the matcher to present the reference him- or herself	By-dyad slopes corresponding to trial number	By-dyad intercepts
Effect of condition and trial number on the director's production of indefinite references	By-dyad slopes corresponding to trial number and by-item slopes corresponding to condition	By-dyad intercepts
Effect of condition and trial number on the director's production of try-markers	By-dyad slopes corresponding to trial number and by-item slopes corresponding to condition	By-dyad intercepts
Effect of condition and trial number on the matcher's requests for clarification	By-dyad slopes corresponding to trial number and by-item slopes corresponding to condition	By-dyad intercepts

Table 2. Results—analysis of the number of words produced by the director

Effect	*b*	*SE*	*P*
Condition: ST/Y12	2.96	3.62	.421
Condition: ST/Y7	9.73	3.62	.021
Condition: ST/ST	0		
Trial	−4.37	1.08	<.001
Trial × condition: ST/Y12	−1.31	1.52	.395
Trial × condition: ST/Y7	−4.37	1.52	.008
Trial × condition: ST/ST	0		

words produced was stronger in the ST/Y7 condition than in the ST/ST condition. The amplitude of the decrease was not significantly different in the ST/Y12 condition and the ST/ST condition.

Effect of condition and trial number on the number of words produced per picture by the matcher

The data corresponding to this analysis are shown in Figure 3. The parameters of the model are shown in Table 3. A significant effect of trial number was found, $F(1, 29) = 37.91$, $p < .001$. An inspection of the *b* coefficient revealed that matchers produced fewer words over trials. Neither the main effect of condition nor the condition × trial number interaction reached statistical significance, respectively $F(2, 26) = 1.66$, $p = .209$ and $F(2, 27) = 1.34$, $p = .280$.

Effect of condition and trial number on the number of speech turns produced per picture by the dyad

The data corresponding to this analysis are shown in Figure 4. The parameters of the model are shown in Table 4. A significant effect of trial number was found, $F(1, 27) = 70.31$, $p < .001$. An inspection of the *b* coefficient revealed that dyads produced fewer speech turns over trials. Neither the main effect of condition nor the condition × trial number interaction reached statistical significance, respectively $F(2, 27) = 0.58$, $p = .566$ and $F(2, 27) = 0.73$, $p = .493$.

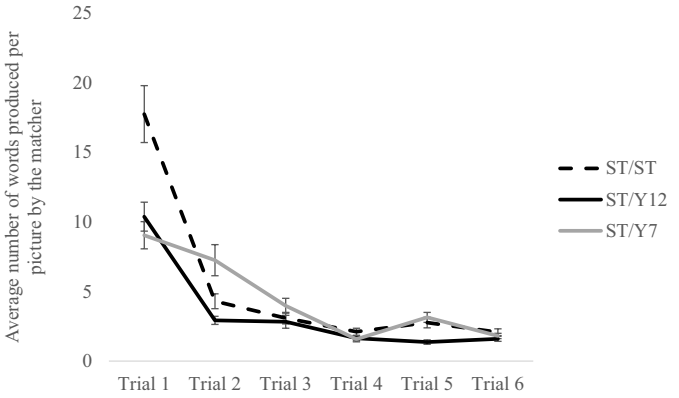

Figure 3. Average number of words produced per picture by the matcher as a function of condition and trial number. The bars represent the standard error.

Table 3. Results—analysis of the number of words produced by the matcher

Effect	*b*	*SE*	*P*
Condition: ST/Y12	−1.89	1.04	.080
Condition: ST/Y7	−0.88	1.04	.405
Condition: ST/ST	0		
Trial	−2.40		
Trial × condition: ST/Y12	0.98	0.68	.161
Trial × condition: ST/Y7	0.94	0.68	.177
Trial × condition: ST/ST	0		

Figure 4. Average number of speech turns produced per picture by the dyad as a function of condition and trial number. The bars represent the standard error.

Table 4. Results—analysis of the number of speech turns produced by the dyad

Effect	*b*	*SE*	*P*
Condition: ST/Y12	−0.07	0.48	.886
Condition: ST/Y7	0.41	0.48	.402
Condition: ST/ST	0		
Trial	−0.78	0.16	<.001
Trial × condition: ST/Y12	0.13	0.23	.580
Trial × condition: ST/Y7	−0.15	0.23	.526
Trial × condition: ST/ST	0		

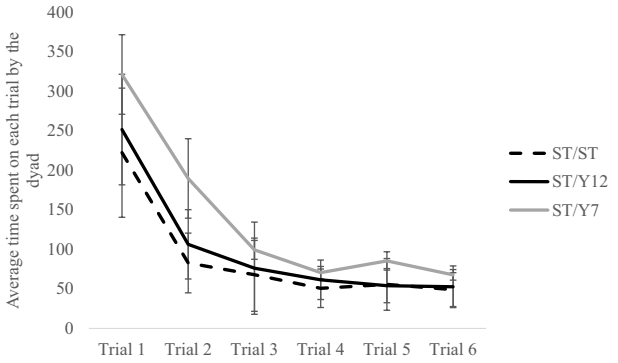

Figure 5. Average time (in seconds) spent on each trial by the dyad as a function of condition and trial number. The bars represent the standard error.

Table 5. Results—analysis of the time taken to complete each trial

Effect	*b*	*SE*	*P*
Condition: ST/Y12	12.30	17.94	.494
Condition: ST/Y7	50.90	17.94	.005
Condition: ST/ST	0		
Trial	−27.56	6.37	<.001
Trial × condition: ST/Y12	−5.76	9.00	.524
Trial × condition: ST/Y7	−18.41	9.00	.043
Trial × condition: ST/ST	0		

Effect of condition and trial number of the time taken (in seconds) to complete each trial

The data corresponding to this analysis are shown in Figure 5. The parameters of the model are shown in Table 5. A significant effect of condition was found, $F(2, 120) = 4.38$, $p = .015$. An inspection of the *b* coefficients revealed that pairs in the ST/Y7 condition took longer to complete each trial than pairs in the ST/ST condition. There was no significant difference between the ST/Y12 condition and the ST/ST condition. An additional pairwise comparison (Bonferroni-corrected) revealed no significant difference between the ST/Y12 condition and the ST/Y7 condition, adjusted $p = .100$. A significant effect of trial number was also found, $F(1, 27) = 93.88$, $p < .001$. The condition × trial number interaction failed to reach statistical significance, $F(2, 120) = 2.19$, $p = .117$.

Reference introduction

Effect of condition and trial number on the probability of the director presenting the reference

The data corresponding to this analysis are shown in Figure 6. The parameters of the model are shown in Table 6. A significant effect of trial number was found, $F(1, 14) = 6.10$, $p = .018$. An inspection of the *b* coefficient revealed that the probability of the director introducing the reference increased over trials. There was a significant condition × trial number interaction, $F(2, 38) = 3.54$, $p = .039$. An inspection of the *b* coefficients revealed that the increase in the probability of the director introducing the reference over trials was weaker in the ST/Y7 condition than in the ST/ST condition. The amplitude of the increase was not significantly different in the ST/Y12 condition and the ST/ST condition. The main effect of condition failed to reach statistical significance, $F(2, 35) = 2.47$, $p = .099$.

Effect of condition and trial number on the probability of the director encouraging the matcher to present the reference him- or herself

The data corresponding to this analysis are shown in Figure 7. Note that only the data from the ST/ST condition and the ST/Y7 condition were included in this analysis. Indeed, there was no variability in

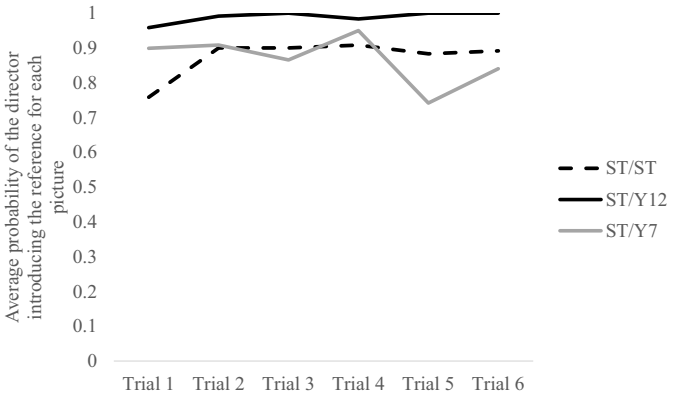

Figure 6. Average probability of the director introducing the reference to describe each picture as a function of condition and trial number. Standard errors are not provided as the DV is binary; dispersion is reflected by the odd ratio (provided below).

Table 6. Results—analysis of the probability of the director presenting the reference

Effect	*b*	*SE*	*P*
Condition: ST/Y12	1.95	1.26	.129
Condition: ST/Y7	−0.76	1.05	.479
Condition: ST/ST	0		
Trial	0.80	0.29	.009
Trial × condition: ST/Y12	−0.06	0.50	.902
Trial × condition: ST/Y7	−0.94	0.38	.021
Trial × condition: ST/ST	0		

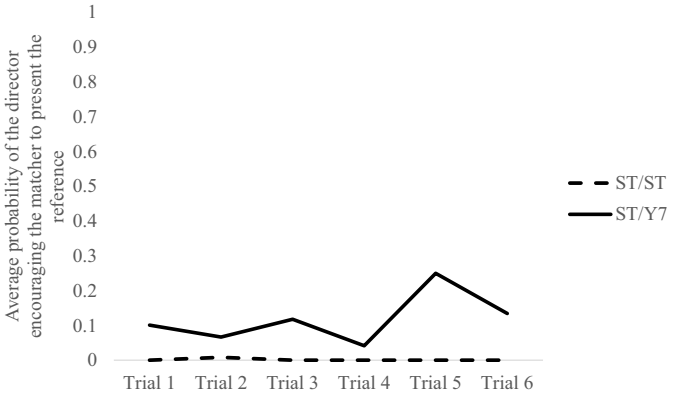

Figure 7. Average probability of the director introducing the reference to describe each picture as a function of condition and trial number. Standard errors are not provided as the DV is binary; dispersion is reflected by the odd ratio (provided below).

the ST/Y12 condition (i.e., the director never encouraged the matcher to present the reference him- or herself in this condition), so including these data would have prevented the model from converging. The parameters of the model are shown in Table 7. A significant effect of condition was found, $F(1, 18) = 4.48$, $p = .049$. An inspection of the *b* coefficient revealed that directors were more likely to encourage the matcher to present the reference in the ST/Y7 condition than in the ST/ST condition. Neither the effect of trial number nor the condition × trial number interaction reached statistical significance, respectively $F(1, 18) = 0.98$, $p = .335$ and $F(1, 18) = 1.14$, $p = .299$.

Table 7. Results—analysis of the probability of the director encouraging the matcher to present the reference

Effect	*b*	*SE*	*P*
Condition: ST/Y7	0.12	0.06	.049
Condition: ST/ST	0		
Trial	−0.01	0.01	.957
Trial × condition: ST/Y7	0.02	0.02	.299
Trial × condition: ST/ST	0		

Characterization of the director's speech

Effect of condition and trial number on the director's production of indefinite references

The data corresponding to this analysis are shown in Figure 8. The parameters of the model are shown in Table 8. A significant effect of trial number was found, $F(1, 27) = 149.18$, $p < .001$. An inspection of the *b* coefficient revealed that directors became less likely to produce indefinite references over trials. Neither the main effect of condition nor the condition × trial number interaction reached statistical significance, respectively $F(2, 26) = 1.35$, $p = .277$ and $F(2, 26) = 0.49$, $p = .617$.

Effect of condition and trial number on the director's production of try-markers

The data corresponding to this analysis are shown in Figure 9. The parameters of the model are shown in Table 9. A significant effect of trial number was found, $F(1, 27) = 146.38$, $p < .001$. An inspection of the *b* coefficient revealed that directors produced fewer try-markers over trials. Neither the main effect of condition nor the condition × trial number interaction reached statistical significance, respectively $F(2, 26) = 2.36$, $p = .114$ and $F(2, 25) = 1.26$, $p = .300$.

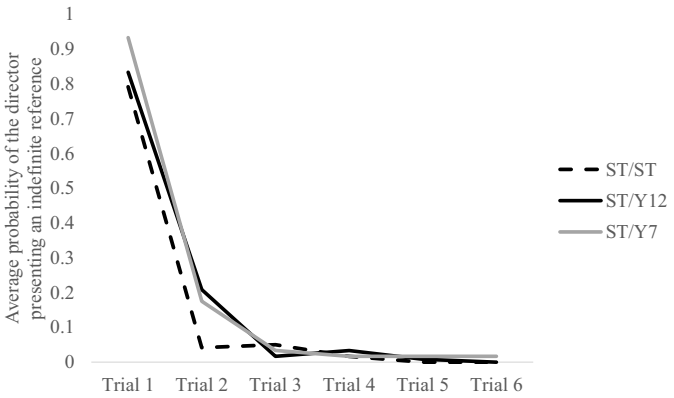

Figure 8. Average probability of the director presenting an indefinite reference as a function of condition and trial number. Standard errors are not provided as the DV is binary; dispersion is reflected by the odd ratio (provided below).

Table 8. Results—analysis of the director's production of indefinite references

Effect	*b*	*SE*	*P*
Condition: ST/Y12	1.80	1.26	.163
Condition: ST/Y7	1.91	1.26	.141
Condition: ST/ST	0		
Trial	−3.18	0.45	<.001
Trial × condition: ST/Y12	0.52	0.59	.383
Trial × condition: ST/Y7	0.52	0.59	.382
Trial × condition: ST/ST	0		

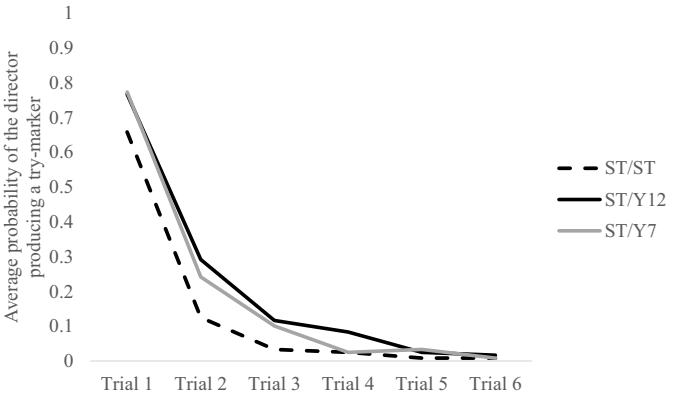

Figure 9. Average probability of the director producing a try-marker as a function of condition and trial number. Standard errors are not provided as the DV is binary; dispersion is reflected by the odd ratio (provided below).

Table 9. Results—analysis of the director's production of try-markers

Effect	*b*	*SE*	*P*
Condition: ST/Y12	2.06	1.08	.066
Condition: ST/Y7	2.15	1.08	.055
Condition: ST/ST	0		
Trial	−2.29	0.33	< .001
Trial × condition: ST/Y12	0.39	0.42	.352
Trial × condition: ST/Y7	0.65	0.41	.124
Trial × condition: ST/ST	0		

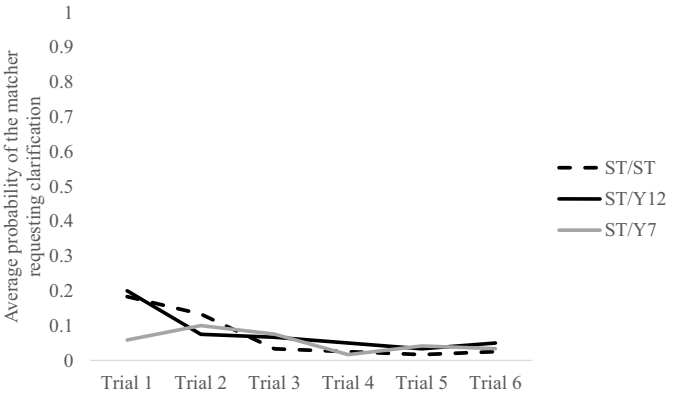

Figure 10. Average probability of the matcher requesting clarification as a function of condition and trial number. Standard errors are not provided as the DV is binary; dispersion is reflected by the odd ratio (provided below).

Characterization of the matcher's speech

Effect of condition and trial number on the matcher's requests for clarification

The data corresponding to this analysis are shown in Figure 10. The parameters of the model are shown in Table 10. A significant effect of trial number was found, $F(1, 33) = 14.70$, $p < .001$. An inspection of the b coefficient revealed that matchers produced fewer requests for clarification over trials. Neither the main effect of condition nor the condition × trial number interaction reached statistical significance, respectively $F(2, 30) = 0.25$, $p = .782$ and $F(2, 32) = 2.45$, $p = .102$.

Table 10. Results—analysis of the matcher's requests for clarification

Effect	*b*	*SE*	*P*
Condition: ST/Y12	0.22	0.46	.636
Condition: ST/Y7	−0.09	0.48	.854
Condition: ST/ST	0		
Trial	−0.56	0.17	.002
Trial × condition: ST/Y12	0.10	0.23	.660
Trial × condition: ST/Y7	0.49	0.24	.045
Trial × condition: ST/ST	0		

Summary of the results

Amount of effort put into the task

Directors produced more words in the ST/Y7 condition than in the ST/ST condition. They also produced fewer words over trials; however, this decrease was stronger in the ST/Y7 condition than in the ST/ST condition. Over trials, matchers produced fewer words, and dyads produced fewer speech turns. Participants in the ST/Y7 condition took longer to complete each trial than participants in the ST/ST condition, and all pairs completed the task quicker across trials.

Reference introduction

In general, the director introduced the reference; this tendency increased over trials, especially in the ST/ST condition compared with the ST/Y7 comparison (where it actually seemed to decrease, as suggested by Figure 6). On occasion, matchers were encouraged to introduce the reference. The likelihood of this was greater in the ST/Y7 condition than in any other conditions.

Characterization of the director's speech

The director produced fewer indefinite references and try-markers over trials. There were no significant differences between conditions.

Characterization of the matcher's speech

The matchers asked for less clarification over the number of trials independent of condition.

Discussion

This study extends previous research by exploring the extent to which teachers engage in audience design when the conversational partner is a student and illustrating how collaborative effort is divided in the dyad in such a setting. The study was conducted in a school, offering a more ecological environment than a laboratory setting. It sought to compare whether teachers adapt their speech differently depending on whether the student is transitioning into their teenage years (year 7 students, 11- to 12-year-olds) or into adulthood (year 12 students, 16- to 17-year-olds).

Amount of effort put into the task

The results of the current study replicated and extended previous findings on the amount of effort put into the matching task by directors and matchers (e.g., Clark & Wilkes-Gibbs, 1986). Indeed, the results of the current study revealed that regardless of condition, the number of words and speech turns produced by the participants decreased over trials, as did the amount of time necessary to complete the task. This confirms that dialogues between teachers and students exhibit similar features to dialogues between adults, at least to a certain extent (see also Fukumura, 2016; Ntsame-Mba & Caron, 1999).

Condition *did* affect the amount of effort put into the task by the teacher playing the role of director in the dyad. The results showed that directors applied significantly more effort, by producing more

words, when introducing a reference to a year 7 student than to another teacher. This can be interpreted in light of the least collaborative effort principle (Clark & Brennan, 1991; Clark & Schaefer, 1989; Clark & Wilkes-Gibbs, 1986), according to which speakers try to present references that can be understood easily by their current partners. In the current study, teachers might have assumed that more detail was necessary to enable year 7 students to adopt their perspective (i.e., the teacher's perspective) and hence select the correct card, and so this extra effort was deemed worthwhile. In contrast, when working with another adult, teachers would have perceived that a short description would suffice, and so minimized effort by providing just this. According to the literature, one consequence of such audience design is that the information presented may be accepted straightforwardly, but we found no evidence of this in the current study, as the nature of the feedback produced by the matcher did not differ significantly across conditions. However, one (other) direct consequence of directors producing lengthier descriptions in the ST/Y7 condition was that the participants took longer to complete the task in this condition, compared to the ST/ST condition.

Importantly, the decrease in the number of words produced by the director over trials was stronger when conversing with a year 7 student than when conversing with an adult. These results suggest that when conversing with year 7 students, directors applied a greater amount of effort in initially establishing the reference. This is in line with our assumption that teachers believe that they need to put extra effort into interacting with younger students, possibly because the latter are perceived as less collaborative (i.e., as less capable of collaborating). However, once the grounding criterion was reached, word production was rapidly minimized, quickly reducing effort. Thus, ST/Y7 pairs were just as capable as other pairs at managing common ground once it had been established.

Audience design such as this requires partner needs to be inferred almost instantly. Presumably, teachers make assumptions about the ability of their partner to perspective-take not solely based on age but also based on prior experience of teaching children with the same community membership (Clark & Marshall, 1981). This community membership is immediately available ("is the student in year 7 or not?"), and hence teacher knowledge of their addressee's perspective is salient, providing the simplicity required for audience design to be computationally feasible (e.g., Galati & Brennan, 2010). Furthermore, teachers and students share community co-membership; both have experience of belonging to a school system underpinned by mutual expectations and this contributes to common ground (e.g., Fussell & Krauss, 1992; Isaacs & Clark, 1987; Lau et al., 2001). Year 7 students are the youngest in a British secondary school and so expect more guidance, and adults in the role of teacher expect to provide this. It is likely that this impacted on the lengthier descriptions produced by adults when working with year 7 students. In sum, through experience of teaching year 7 students, teachers have *knowledge* of their addressee's perspective, and then are *able* to create mental models of their needs, adapting reference presentation accordingly.

The results discussed up until this point thus suggest that teachers put extra effort into interacting with young students. One obvious limitation of our work is that we did not examine how useful such extra effort actually was. We interpret this finding as reflecting the teachers' attempts to reduce the collaborative effort in ST/Y7 pairs. However, we do not know what would have happened if the teachers had not "helped" their students in this situation. Would the students' performance have been worse (potentially leading to an increase in the number of words and speech turns produced by the dyad) or would it have stayed the same (in which case the teachers producing lengthier description would in fact reflect an erroneous representation of their students' ability to perform the task)? These questions should be addressed in future experiments in which teachers' contributions will be scripted, allowing us to control the length of the descriptions produced.

Reference introduction

In this kind of task, although participants are told that they may interact freely to complete the task, references are generally introduced by the director rather than the matcher (e.g., Clark & Wilkes-Gibbs, 1986). This was also the case in the current study, but the results revealed that there were

differences between conditions. Specifically, the probability of the directors introducing the reference themselves in each trial increased over trials, but this increase was weaker in the ST/Y7 condition than in the ST/ST condition. In other words, matchers remained more likely to introduce references over trials in the ST/Y7 condition than in the ST/ST condition. The first implication of this is that dialogue between teachers and students differs from adult–adult dialogue not only in terms of the *amount* of effort put into the interaction (as highlighted in the previous section), but also in terms of how this effort is *distributed* and *shared* among dialogue partners. Results regarding whether the matchers were encouraged by the directors to introduce references help shed further light on this phenomenon. These results revealed that directors were more likely to encourage the matcher to present the reference in the ST/Y7 condition than in the ST/ST condition. Thus, year 7 matchers did not take the lead in the task spontaneously but were encouraged by their teachers to do so. One possible interpretation is that teachers feel a heightened responsibility for encouraging year 7 students. They are more likely to expend extra effort in empowering the students (by encouraging them to introduce a reference) in a purposeful attempt to reduce status imbalance, enriching the quality of adaptation. Through classroom experience, teachers are also likely to be aware that self-generation of a reference will assist memory (Knutsen & Le Bigot, 2014; MacLeod, 2011; Rosner et al., 2013; Slamecka & Graf, 1978). Therefore, in an effort to ensure that references are *firmly* placed in common ground teachers adapt to year 7 students by encouraging them to initially produce the reference.

Characteristics of director's speech and matcher's speech

Previous research on adult–child interaction suggests audience design is characterized by an adjustment of *what* is said, so that adults make lexical choices dependent on whether the addressee is another adult or a child (e.g., Bates & Silvern, 1977; Clark, 2010, 2015; Clark & Bernicot, 2008; Clark & Estigarribia, 2011; Epley et al., 2004; Nadig & Sedivy, 2002; Ntsame-Mba & Caron, 1999; O'Neill, 1996; Sachs & Devin, 1976; Shatz & Gelman, 1973; Warden, 1976). In addition, studies show that while older teenagers manage definite references and the process of establishing common ground better than younger children, they have yet to master this in a fully adult-like manner (Fukumura, 2016; Ntsame-Mba & Caron, 1999). Consequently, in this study it was expected that directors would adapt reference presentation by including more try-markers and indefinite references when conversing with year 7 students than with year 12 students and when conversing with year 12 students than with adults. The data did not support these hypotheses: although the use of try-markers and indefinite references decreased over trials, this was not significantly dependent on dyad condition.

This finding thus suggests that once the grounding criterion was met, the reference was deemed uniquely identifiable. Once the conceptualization was accepted, try-markers became redundant (Clark & Brennan, 1991) and there are were no gains in continuing to use indefinites (Clark & Marshall, 1981).

Summary and directions for future research

The current study provides an insight into how teachers adapt speech production for students transitioning into teenage years (11- to 12-year-olds). The lack of a significant difference between the ST/Y12 condition and the other two conditions prevents us from drawing conclusions regarding how teachers adapt speech for students transitioning into adulthood (16- to 17-year-olds).

Two main conclusions can be drawn from the results of this study. Firstly, student–teacher dialogue exhibits features that are at least in part similar to those of adult–adult dialogue. In particular, just like in adult–adult dialogue, teachers and students produce fewer words, speech turns, indefinite references, try-markers, and clarification requests as they gather common ground (e.g., Clark & Wilkes-Gibbs, 1986). It thus seems that teachers take into account the fact that their teenage dialogue partner is becoming increasingly able to behave in an adult-like, collaborative way (e.g., Fukumura, 2016; Ntsame-Mba & Caron, 1999). Second, it is noteworthy that teachers continue to guide the interaction (as adults do with younger children; Clark & Bernicot, 2008; Clark & Estigarribia, 2011), at least when

interacting with younger students. This suggests that audience design when talking to students is not characterized by what is said but mainly by an adjustment of the amount of effort put into the interaction by the teacher. Teachers expend more effort by producing more words when interacting with a younger student compared to another teacher. This contributes further to previous research by proposing that audience design for younger students is set apart by the teachers' effort in producing lengthier descriptions and in attempts to encourage participation in the dialogue.

The current study also sought to examine dialogue outside the laboratory. Although replicating the main findings of the matching task in teacher–student interaction settings is of central interest to the dialogue research community, this also raises a number of methodological questions regarding the generalizability of the results. In particular, it is not necessarily true that all adults engage in audience design in the same way as teachers. Teachers interact with children daily, and so this wealth of experience informs mental models based on age group needs, allowing for audience design. Furthermore, in the school where the study was conducted, stark visual prompts such as year 7 students being in uniform and year 12 students not, adds to the scope of the teacher to preplan adaptations: The uniform worn by year 7 students may have exaggerated their differing status from adults. Whether adaptations are made in the same way for children out of uniform, where the adult may have to rely more heavily on using feedback to then adapt, is a goal for future research. This suggestion is in line with Ferreira's (2019) recent suggestion that some (automatic) aspects of audience design may be triggered by salient cues in the speaker's environment, whereas other aspects of audience design rely more heavily on cognitive control.

Research also suggests that motivation affects the extent to which people engage in audience design (Keysar et al., 1998; Rossnagel, 2000). Teachers have a vested interest in encouraging their students so they can build relationships that support academic success. This might make them all the more likely to engage in audience design when interacting with their students. Other adults may not have the same motivation and so whether they would encourage the year 7 students more than other adults when engaging in dialogue is open to further research. Future research will thus need to determine whether adult–teenager interaction displays the same characteristics as those found in the current study in dialogues among teachers and students.

Finally, the current study raises the question of whether the results obtained were due to the students differing from the teachers in terms of age, in terms of status, or both. We hope that future work in which one of these factors will be manipulated while the other one will remain constant (e.g., using settings in which teachers interact with mature students and/or settings in which adults who are not teachers interact with teenagers) will help shed light on this question.

Conclusion

Dialogues between teachers and students differ from dialogues among adults in terms of the way in which the collaborative effort toward reaching mutual understanding is shared in the dyad. This mainly concerns the way in which information is introduced by the teacher. Specifically, teachers adapt to the needs of younger students by providing additional information (e.g., lengthier descriptions) and by encouraging active participation in the interaction. Evidence of adaptation is thus located in the amount of effort a teacher expends in establishing common ground with younger students rather than in how they frame a reference (e.g., use of indefinite pronouns and try-markers).

Notes

1. It is important to highlight here that whereas year 7 students are required to wear a uniform, year 12 students are not. This might have made the difference between both age groups more salient to the teachers playing the role of directors in the current study. We return to this point in the discussion.
2. As highlighted here, we tried to make sure that the participants did not know each other well before the experiment. Nonetheless, this was not always possible, as the participants worked or studied in the same school.

Thus, in the ST/ST condition, 6 pairs knew each other before the experiment; in the ST/Y12 condition, 5 pairs knew each other before the experiment; in the ST/Y7 condition, one pair knew each other before the experiment. To check that this did not affect the results of the experiment, a series of additional analyses was conducted in parallel to the main analyses reported hereafter, with prior knowledge as an additional independent variable. Only the data from the ST/ST and ST/Y12 conditions were included in the analyses. Indeed, there were not enough pairs who knew each other prior to the experiment in the ST/Y7 condition to include the data from this condition in the analysis. We found that only the probability of the director introducing the reference was influenced by prior knowledge: in the ST/ST condition, the director was less likely to introduce the reference when he or she knew the matcher beforehand. However, this cannot explain the results obtained in this study. Indeed, as detailed below, directors were less likely to introduce the reference in the ST/Y7 condition (where most pairs did not know each other prior to the experiment) than in the ST/ST condition, especially in later trials. Thus, these results cannot be attributed to the fact that more participants knew each other in the ST/ST condition, as this would have led the directors in this condition to introduce the references less often than in the ST/Y7 condition.

3. No double-coding was performed on the other DVs included in this study, as they were computed automatically (i.e., number of words and speech turns produced and time taken to complete the task).

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Alibali, M. W., Nathan, M. J., Church, R. B., Wolfgram, M. S., Kim, S., & Knuth, E. J. (2013). Teachers' gestures and speech in mathematics lessons : Forging common ground by resolving trouble spots. *ZDM - Mathematics Education*, *45*(3), 425–440. https://doi.org/10.1007/s11858-012-0476-0

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*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, E., & Silvern, L. (1977). Adjustment and politeness in preschoolers. *Journal of Communication*, *27*(2), 104–111. https://doi.org/10.1111/j.1460-2466.1977.tb01834.x

Bortfeld, H., & Brennan, S. E. (1997). Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes*, *23*(2), 119–147. https://doi.org/10.1080/01638537709544986

Branigan, H. P., Bell, J., & McLean, J. F. (2016). Do you know what I know? The impact of participant role in children's referential communication. *Frontiers in Psychology*, *7*, 213. https://doi.org/10.3389/fpsyg.2016.00213

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*(6), 1482–1493. https://doi.org/10.1037/0278-7393.22.6.1482

Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialog. *Topics in Cognitive Science*, *1*(2), 274–291. https://doi.org/10.1111/j.1756-8765.2009.01019.x

Brown, P., & Levinson, S. C. (1987). *Politeness. Some universals in language use*. Cambridge University Press.

Calcagni, E., & Lago, L. (2018). The three domains for dialogue : A framework for analysing dialogic approaches to teaching and learning. *Learning, Culture and Social Interaction*, *18*, 1–12. https://doi.org/10.1016/j.lcsi.2018.03.001

Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, *64*(1), 1–37. https://doi.org/10.1016/S0010-0277(97)00010-3

Clark, E. V. (2007). Young children's uptake of new words in conversation. *Language in Society*, *36*(2), 157–182. https://doi.org/10.1017/S0047404507070091

Clark, E. V. (2010). Adult offer, word-class, and child uptake in early lexical acquisition. *First Language*, *30*(3–4), 250–269. https://doi.org/10.1177/0142723710370537

Clark, E. V. (2015). Common ground. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence* (pp. 328–353). John Wiley & Sons.

Clark, E. V., & Bernicot, J. (2008). Repetition as ratification: How parents and children place information in common ground. *Journal of Child Language*, *35*(2), 349–371. https://doi.org/10.1017/S0305000907008537

Clark, E. V., & Estigarribia, B. (2011). Using speech and gesture to introduce new objects to young children. *Gesture*, *11* (1), 1–23. https://doi.org/10.1075/gest.11.1.01cla

Clark, H. H. (1996). *Using language*. 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). APA Books.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*(1), 62–81. https://doi.org/0.1016/j.jml.2003.08.004

Clark, H. H., & Marshall, C. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (Vol. 2, pp. 10–63). Cambridge University Press.
Clark, H. H., & Murphy, G. L. (1982). La visée vers l'auditoire dans la signification et la référence. *Bulletin de Psychologie*, *35*(11-16), 767–776.
Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, *13*(2), 259–294. https://doi.org/10.1207/s15516709cog1302_7
Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39. https://doi.org/10.1016/0010-0277(86)90010-7
De Ruiter, J. P., Bangerter, A., & Dings, P. (2012). The interplay between gesture and speech in the production of referring expressions: Investigating the tradeoff hypothesis. *Topics in Cognitive Science*, *4*(2), 232–248. https://doi.org/10.1111/j.1756-8765.2012.01183.x
Epley, N., Morewedge, C., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*(6), 760–768. https://doi.org/10.1016/j.jesp.2004.02.002
Ferreira, V. S. (2019). A mechanistic framework for explaining audience design in language production. *Annual Review of Psychology*, *70*(1), 29–51. https://doi.org/10.1146/annurev-psych-122216-011653
Fukumura, K. (2015). Interface of linguistic and visual information during audience design. *Cognitive Science*, *39*(6), 1419–1433. https://doi.org/10.1111/cogs.12207
Fukumura, K. (2016). Development of audience design in children with and without ASD. *Developmental Psychology*, *52*(1), 71–87. https://doi.org/10.1037/dev0000064
Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology*, *62*(3), 378–391. https://doi.org/10.1037/0022-3514.62.3.378
Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, *62*(1), 35–51. https://doi.org/10.1016/j.jml.2009.09.002
Ganea, P. A., & Saylor, M. M. (2007). Infants' use of shared linguistic information to clarify ambiguous requests. *Child Development*, *78*(2), 493–502. https://doi.org/10.1111/j.1467-8624.2007.01011.x
Golinkoff, R. (1986). 'I beg your pardon?' : The preverbal negotiation of failed messages. *Journal of Child Language*, *13*(3), 455–476. https://doi.org/10.1017/S0305000900006826
Grigoroglou, M., & Papafragou, A. (2016). Are children flexible speakers? Effects of typicality and listener needs in children's event descriptions. In A. Papafragou, D. J. Grodner, D. Mirman, & J. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 782–787).
Horton, W. S., & Gerrig, R. J. (2002). Speakers' experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language*, *47*(4), 589–606. https://doi.org/16/S0749-596X(02)00019-0
Horton, W. S., & Gerrig, R. J. (2005a). Conversational common ground and memory processes in language production. *Discourse Processes*, *40*(1), 1–35. https://doi.org/10.1207/s15326950dp4001_1
Horton, W. S., & Gerrig, R. J. (2005b). The impact of memory demands on audience design during language production. *Cognition*, *96*(2), 127–142. https://doi.org/10.1016/j.cognition.2004.07.001
Horton, W. S., & Gerrig, R. J. (2016). Revisiting the memory-based processing approach to common ground. *Topics in Cognitive Science*, *8*(4), 780–795. https://doi.org/10.1111/tops.12216
Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*(1), 91–117. https://doi.org/10.1016/0010-0277(96)81418-1
Isaacs, E. A., & Clark, H. H. (1987). Reference in conversations between experts and novices. *Journal of Experimental Psychology. General*, *116*(1), 26–37. https://doi.org/10.1037/0096-3445.116.1.26
Jaeger, T. F. (2008). Categorical data analysis : Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007
Keselman, H. J., Algina, J., Kowalchuk, R. K., & Wolfinger, R. D. (1999). The analysis of repeated measurements : A comparison of mixed-model satterthwaite F tests and a nonpooled adjusted degrees of freedom multivariate test. *Communications in Statistics - Theory and Methods*, *28*(12), 2967–2999. https://doi.org/10.1080/03610929908832460
Keysar, B., Barr, D. J., & Horton, W. S. (1998). The egocentric basis of language use : Insights from a processing approach. *Current Directions in Psychological Science*, *7*(2), 46–49. https://doi.org/10.1111/1467-8721.ep13175613
Kiernan, K., Tao, J., & Gibbs, P. (2012). *Tips and strategies for mixed modelling with SAS/STAT procedures*. 2012 SAS Global Forum, Orlando, FL.
Knutsen, D., Bangerter, A., & Mayor, E. (2019). Procedural coordination in the matching task. *Collabra*, *5*(1), 3. https://doi.org/10.1525/collabra.188
Knutsen, D., & Le Bigot, L. (2014). Capturing egocentric biases in reference reuse during collaborative dialogue. *Psychonomic Bulletin & Review*, *21*(6), 1590–1599. https://doi.org/10.3758/s13423-014-0620-7
Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology*, *4*(3), 343–346. https://doi.org/10.1037/h0023705

Lau, I. Y., Chiu, C., & Hong, Y. (2001). I know what you know: Assumptions about others' knowledge and their effects on message construction. *Social Cognition*, *19*(6), 587–600. https://doi.org/10.1521/soco.19.6.587.20888
MacLeod, C. M. (2011). I said, you said: The production effect gets personal. *Psychonomic Bulletin & Review*, *18*(6), 1197–1202. https://doi.org/10.3758/s13423-011-0168-8
Matthews, D., Lieven, E., & Tomasello, M. (2010). What's in a manner of speaking? Children's sensivity to partner-specific referential precedents. *Developmental Psychology*, *46*(4), 749–760. https://doi.org/10.1037/a0019657
McInnes, F., & Attwater, D. (2004). Turn-taking and grounding in spoken telephone number transfers. *Speech Communication*, *43*(3), 205–223. https://doi.org/10.1016/j.specom.2004.04.001
Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, *49*(2), 201–213. https://doi.org/10.1016/S0749-596X(03)00028-7
Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*(4), 329–336. https://doi.org/10.1111/j.0956-7976.2002.00460.x
Ntsame-Mba, F., & Caron, J. (1999). Construction et gestion du terrain commun chez des enfants de sept à quatorze ans et des adultes. *Archives de Psychologie*, *67*, 37–55.
O'Neill, D. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development*, *67*(2), 659–677. https://doi.org/10.1111/j.1467-8624.1996.tb01758.x
Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–190. https://doi.org/10.1017/S0140525X04000056
Rosner, Z. A., Elman, J. A., & Shimamura, A. P. (2013). The generation effect: Activating broad neural circuits during memory encoding. *Cortex*, *49*(7), 1901–1909. https://doi.org/10.1016/j.cortex.2012.09.009
Rossnagel, C. (2000). Cognitive load and perspective-taking: Applying the automatic-controlled distinction to verbal communication. *European Journal of Social Psychology*, *30*(3), 429–445. https://doi.org/10.1002/(SICI)1099-0992(200005/06)30:3<429::AID-EJSP3><429::AID-EJSP3>3.0.CO;2-V
Sachs, J., & Devin, J. (1976). Young children's use of age-appropriate speech styles in social interaction and role-playing. *Journal of Child Language*, *3*(1), 81–98. https://doi.org/10.1017/S030500090000132X
Sacks, H., & Schegloff, E. (1979). Two preferences in the organisation of reference to persons in conversation and their interaction. In G. Psathas (Ed.), *Everyday language : Studies in ethnomethodology* (pp. 15–21). Irvington.
Schober, M. F. (2009). Spatial dialogue between partners with mismatched abilities. In K. R. Coventry, T. Tenbrink, & J. Bateman (Eds.), *Spatial language and dialogue* (pp. 23–39). Oxford University Press.
Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*(2), 211–232. https://doi.org/10.1016/0010-0285(89)90008-X
Shatz, M., & Gelman, R. (1973). The development of communication skills : Modifications in the speech of children as a function of listener. *Monographs of the Society for Research in Child Development*, *38*(5), 1–38. https://doi.org/10.2307/1165783
Slamecka, N., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology. Human Learning and Memory*, *4*(6), 592–604. https://doi.org/10.1037/0278-7393.4.6.592
Stalnaker, R. (1978). Assertion. In P. Cole (Ed.), *Syntax and semantics* (Vol. 9, pp. 315–332). Academic Press.
Warden, D. A. (1976). The influence of context on children's use of identifying expressions and references. *British Journal of Psychology*, *22*(1), 101–112. https://doi.org/10.1111/j.2044-8295.1976.tb01501.x
Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language*, *31*(2), 183–194. https://doi.org/10.1016/0749-596X(92)90010-U