Do young adolescents draw on common ground like adults?

Clara Grégoire[1] , Héloïse Retournard[2] , Stéphane Rauzy[1] , Maud Champagne-Lavau[1]
[1]Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France; {first.last}@univ-amu.fr
[2]Université Claude Bernard Lyon 1, Lyon, France

Abstract

Audience design is the ability to adapt one's language to one's interlocutor. In adolescents, this ability, essential during a conversation, is already present but not yet to the level of adult's. Furthermore, existing studies have not looked at the co-elaboration of shared knowledge between adolescents during an interaction. This paper investigates whether, with a referential communication task, adolescents aged between 11 and 13 years use their audience design in the same way as adults. To this end, we assessed their use of reference markers during the task. The results showed that adolescents used in a similar trend as adults reference markers, but it is not as optimally as adults do.

1 Introduction

In everyday conversation, speakers routinely adapt their language based on what they believe their interlocutor knows. If someone mentions their long-haired cat as a fluffy cat and then refers to "*the fluffy thing*" the listener will likely identify the referent. This process of taking into account or adapting to the specific needs and knowledge of the listener is known as *audience design* (Clark and Murphy, 1982). It also refers to *perspective-taking* which consists of attributing knowledge to the interlocutor during conversation and ensuring the appropriate use of the other's perspective (Brown-Schmidt and Heller, 2018; Clark and Murphy, 1982). Audience design implies that interlocutors involved in a conversation distinguish between common ground (or *shared knowledge*) co-constructed by them and privileged ground (or *privileged knowledge*) to ensure effective communication. Shared knowledge is established between interlocutors through mutual acceptance that certain information is shared and can be used within their conversation. Specifically, information becomes part of the shared knowledge once it is proposed by speaker A and accepted by speaker B. At that point, both interlocutors acknowledge that the information is shared (Clark and Wilkes-Gibbs, 1986). This mutual co-construction is essential and requires the active participation of both parties, along with their willingness to incorporate information into their common ground. In this way, a *mutual belief* is formed: both interlocutors are aware that the other has understood what has just been said before the conversation proceeds (Clark and Wilkes-Gibbs, 1986). While privileged knowledge refers to information known only to one person and not mutually established with the interlocutors, shared knowledge can arise from physical or linguistic co-presence, prior shared experiences, general world knowledge, or joint construction during interaction (Clark and Wilkes-Gibbs, 1986; Brown-Schmidt and Heller, 2018).

A central question is to understand how this mechanism emerges and develops from childhood to adulthood. Adolescence spans from puberty onset around age 10 to early adulthood near age 20, when individuals assume adult social responsibilities (Galvàn, 2021). This transitional phase is characterized by social and relational transformations that reshape interactions with both family and peers (Galvàn, 2021). Notably, adolescence is a critical period for the development of perspective-taking, as brain regions involved in social cognition and interaction continue to mature during this stage (Kilford et al., 2016; Galvàn, 2021). The present study focused on whether and how young adolescents appropriately use common ground and privileged ground compared to adults, especially during a *referential communication task* conducted in French.

1.1 Audience design during adolescence

Fukumura (2016) and Arvidsson et al. (2022) findings suggest that, during adolescence, audience design is still improving. First, they showed that, unlike adults, adolescents did not adjust contrastive

Proceedings of the 29th Workshop on the Semantics and Pragmatics of Dialogue, September 3–5, 2025, Bielefeld, Germany.

	Arvidsson et al. (2022)	**Fukumura (2016)**
Interlocutor	Fictitious	Real : with a confederate
Experimental groups	Adolescents aged 11-12 years Adolescents aged 15-16 years	Children aged 6-10 years Adolescents aged 11-16 years Adults
Experimental conditions	Knowledge of the fictitious character: presumed known/ presumed unknown	Different-size distractor: present / absent Target visibility: both participants / director only
Knowledge's type	Encyclopedic knowledge	From Visual co-presence

Table 1: Summary of the methodological differences between Arvidsson et al. (2022), and Fukumura (2016).

and relevant descriptions according to shared or privileged knowledge. In both studies (method detailed in Table 1), an interlocutor makes a partner describe or guess an image among 4 using either detailed descriptions or contrastive adjectives. On the one hand, the authors found that adolescents used fewer contrastive adjectives than adults in the *shared condition* (Fukumura, 2016) and more precisely, during adolescence, younger adolescents gave more detailed descriptions than older adolescents (Arvidsson et al., 2022). On the other hand, in the *privileged condition*, they found that adolescents used more contrastive adjectives than adults, and that younger adolescents gave less detailed descriptions than older adolescents. Second, both studies demonstrated that when performing the task, whether the partner was fictitious or real, young adolescents were less efficient than older adolescents or adults at taking their partner's knowledge into account, regardless of its type (*i.e.*, encyclopedic or from visual co-presence; Table 1). Adolescents thus appear to rely less on shared knowledge and more on privileged knowledge, suggesting that they are more egocentric than adults. Moreover, using a questionnaire on the presumed knowledge of fictitious characters about the pictures, given after the task, Arvidsson et al. (2022) found no differences in the attribution of knowledge levels between younger and older adolescents. While younger adolescents attributed the same knowledge to the characters as older adolescents, they did not use this information to perform the task. This suggests that they are not yet fully engaged in audience design. In line with this result, other research investigating theory of mind abilities has found that adolescents are more likely to adopt an egocentric perspective compared to adults. For instance, authors using a director task adapted from Keysar et al. (2000) have examined how adolescents adjust their perspective-taking when interacting with a fictitious interlocutor (Dumontheil et al., 2010). In these studies, participants viewed a grid of objects and were informed that certain items were occluded from the fictitious partner's view. Correct object selection required ignoring objects occluded to the other perspective. Adolescents aged 11 to 17 years (Tamnes et al., 2018; De Lillo and Ferguson, 2023) and those aged 14 to 18 years (Dumontheil et al., 2010; Symeonidou et al., 2016) made more errors in selecting the target than adults. Moreover, when multiple objects were similar to the target, adolescents aged 12 to 15 years made more errors than those aged 17 years (Humphrey and Dumontheil, 2016). Overall, research on adolescents' use of shared knowledge suggests that they engage in audience design to a lesser extent than adults. However, there is no clear consensus on the exact age at which adolescents begin to do so, as findings vary depending on the experimental paradigm used. Some studies indicate that audience design begins to develop in between 12 and 15 years (Arvidsson et al., 2022), while others suggest it continues to mature until approximately 15 to 17 years of age (Humphrey and Dumontheil, 2016). It is also important to note that all of the aforementioned studies assessed adolescents' performance in tasks without real interaction, and where shared knowledge was not co-constructed. In contrast, referential communication tasks may offer insight into how common ground is collaboratively negotiated in real-time.

1.2 Measures of audience design using the referential communication task

The referential communication task is traditionally used in the literature to assess audience design in adult population (Krauss and Weinheimer, 1964; Clark and Wilkes-Gibbs, 1986). Two participants

are presented with the same set of abstract figures (tangrams). One participant, the *director*, sees the tangrams arranged in a specific order, while the other, the addressee, views the same tangrams in a random order. The director's objective is to guide the *addressee* in rearranging the tangrams to match their own arrangement. The task is repeated across 4 to 6 trials, depending on the study. Throughout the trials, participants gradually develop shared knowledge about the tangrams, which were initially unknown (Clark and Wilkes-Gibbs, 1986).

This type of task thus requires participants to co-construct a common ground, which is constituted of descriptions proposed and used to refer to the tangrams through the task (Brennan and Clark, 1996). Development of this common ground can be assessed by examining the number of words used and the types of referential markers employed (Bovet et al., 2024). More specifically, the use of definite and indefinite references, markers of conceptualization and reconceptualization (based on the types of words and labels produced), as well as the reuse of previously established descriptions reflects how interlocutors take into account knowledge that is part of the common ground (Bovet et al., 2024). Indefinite references introduce new information (*e.g.*, "*a nice dog*" when mentioning the dog for the first time), whereas definite references refer to information presumed to be known by both interlocutors (*e.g.*, "*the fluffy cat*"). A definite reference may take the form of a proper name known to both participants or a label without an article (*e.g.*, "*next figure cat*"). A third strategy to introduce information involves employing provisional references or hedges, such as "*a kind of*" or "*a sort of*" (Brennan and Clark, 1996; Branigan et al., 2016). Hedges convey uncertainty and invite clarification or negotiation, allowing interlocutors to confirm, refine, or adjust the descriptions provided (Brennan and Clark, 1996). As interaction progresses across repeated trials with the same tangrams, interlocutors gradually replace indefinite references and hedges with definite ones from the second trial onward, reflecting the construction and the use of shared knowledge (Clark and Wilkes-Gibbs, 1986; Champagne-Lavau et al., 2009). The provisional references tend to fade as shared knowledge is jointly established, especially when the same objects are referred to repeatedly (Brennan and Clark, 1996). The reduction in the number of words used to describe tangrams across trials also demonstrates the development of shared knowledge (Isaacs and Clark, 1987; Clark and Wilkes-Gibbs, 1986). When common ground is established and used, interlocutors rely more on previously shared descriptions and shorten them. This process allows interlocutors to refer to tangrams more efficiently without needing to provide new elements or to produce new descriptions (Isaacs and Clark, 1987). This decrease in the number of words reflects increased communicative efficiency and can be used as a measure of task success.

In parallel, the reuse of previously produced words and descriptions signals that both interlocutors are drawing from shared knowledge. This not only enhances mutual understanding, but also contributes to faster and more efficient task completion (Isaacs and Clark, 1987). For instance, the reuse of descriptions has been examined in referential communication task involving either pairs of native speakers or mixed pairs including a native speaker and a second-language learner (Bortfeld and Brennan, 1997). In this study, participants completed the task six times, alternating between the roles of director and addressee. The use of shared knowledge was assessed by analyzing whether descriptions from the final two trials for each tangram were reused. The reuse of previously mentioned and jointly established elements was interpreted as evidence of successful mutual belief formation, whereas substantial changes in terminology were taken as failures to rely on common ground. The results showed that adults, whether paired with native or non-native speakers, consistently built and reused previously established descriptions in similar ways. Nadig et al. (2015) analyzed description reuse while they compared how directors described tangrams when addressing either the same addressee as in previous trials or a new one. The addressee changed on the fourth trial. The critical fourth trial enabled researchers to evaluate participants' consideration of the absence of shared knowledge with a new interlocutor. Findings revealed that adults produced longer descriptions when the addressee changed, suggesting an awareness of the lack of shared knowledge.

Conversely, when the addressee remained the same, speakers tended to reuse descriptions. Referential communication tasks have also been used to examine audience design in younger children by introducing a new addressee (*e.g.*, bystander, over hearer, or naïve participant) midway through the interaction. Depending on the role of the new addressee, the common ground shared with the di-

rector varied: the naïve participant was unfamiliar with both the task and the figures, the bystander had actively participated in the task and the over hearer had only listened to the previous interaction without taking part in it. Branigan et al. (2016) found that children aged 8–10 years often relied on presumed common ground, even when it was unwarranted, in the case of an over hearer as new addressee. Specifically, children used more definite references with the second addressee in a role of an over hearer, than what was observed in a study in adults (Wilkes-Gibbs and Clark, 1992) despite the fact that the over hearer had not seen the tangrams and thus lacked visual common ground. However, when the new addressee was either naïve or had been physically present, children appropriately adjusted their referring expressions based on the addressee's knowledge, as adults did (Wilkes-Gibbs and Clark, 1992).

To our knowledge, no study has assessed audience design in young adolescents interacting with a real interlocutor of the same age, in comparison with adults. Therefore, in the current study, we aimed to investigate whether early adolescents (aged 11–13 years) take into account and use shared knowledge established with their interlocutor during a referential communication task, in the same way as adults. Specifically, we assessed: (1) the types of referring expressions used to introduce tangram descriptions across trials (definite, indefinite, and hedges); (2) the extent to which participants reused tangram descriptions throughout the trials. The use of referential markers has been studied in younger children (Branigan et al., 2016) and adults (Clark and Wilkes-Gibbs, 1986), while description reuse has been well documented in adults (Bortfeld and Brennan, 1997; Nadig et al., 2015), but not in adolescents. We hypothesized that if adolescents were less likely to rely on shared knowledge with their interlocutor, they would use more words, indefinite markers and hedges and less definite ones than adults throughout the task. We also predicted that adolescents would show less consistency in reusing previously produced descriptions compared to adults.

2 Method

2.1 Participants

Forty adolescents aged 11 to 13 years ("Adolescence group" ($n = 40$), $M = 12.5$, $SD = 0.10$ years) and forty young adults aged 18 to 24 years ("Adult group" ($n = 40$), $M = 20.07$, $SD = 0.37$ years) took part in the study, performing the task in pairs. Each Group contained the same number of females and males ($N = 20$). All participants were fluent in French.

The adolescents were recruited from a French middle school in Briançon (France). Young adults were recruited from Aix-Marseille university and from the last year of a high school in Briançon. Participants were tested either at school or in the laboratory according to their recruitment location. The participants in each pair were friends. All participants and parents of the adolescents gave written informed consent before recruitment. Adults received financial compensation (€20), while adolescents were given a gift (*e.g.*, board game), in accordance with French ethical guidelines. The study was approved by the Ethics Committee of Aix-Marseille University (n° 2024-01-11-02).

2.2 Materials

We used the paradigm of referential communication (Clark and Wilkes-Gibbs, 1986; Champagne-Lavau et al., 2009). In this task, a participant (the "*director*") sequentially describes 10 *a priori* unknown figures (tangrams) to enable their partner (the "*addressee*") to place them in the same order. The director was presented with the tangrams arranged in a 2×5 laminated grid, while the addressee received the same grid and 10 laminated cards ($5,5 \times 9$ cm) depicting the tangrams. The task was performed 6 times, in a different pseudo-randomized order of the tangrams and constituting the six trials of the task. Participants sat face-to-face, separated by an opaque screen during the task. After each trial, the screen was removed, and both participants verified with the experimenter if their final order was the same. Each participant was provided with their own microphone to enable distinct audio recordings. The task lasted from 7 to 25 minutes.

2.3 Data coding

The six trials of the task were automatically transcribed using a script based on VOSK-API (Shmyrev and other contributors, 2020) and then manually corrected in PRAAT (Boersma and Weenink, 2024). Audio record of each participant was transcribed orthographically. Transcriptions were divided by trial and by tangram.

Two coders filtered the directors' transcripts to retain only the elements related to tangram descrip-

tions. The first filter targeted utterances preceding the actual description of the tangram (*i.e.*, references to the tangram's position and its introduction to the addressee), such as "*the next picture is*", as our focus was on the references produced specifically about the tangrams. For the same reason, we excluded questions directed to the experimenter, as they were not part of the dialogue between the two participants in the task. Finally, we filtered out digressions between participants that were unrelated to the tangrams or the task. While such digressions may contribute to the grounding process between participants, our analyses focused specifically on strategies for describing the tangrams.

A tangram description was thus defined as speech beginning when the director started speaking and ending before the addressee's first feedback. We identified the "*initial*" description, which was the director's first description before any feedback from the addressee. We also isolated the "*final*" description, which was the last description made before moving on to the next tangram. The final description could be produced by the director when the addressee accepted it (by giving feedback) or when the addressee proposed a description accepted by the director.

The corpus analyzed in this study consisted exclusively of filtered data. For each trial, we recorded the number of words, and the number of descriptions produced per tangram. The full corpus lasts 8 hours and 53 minutes and includes 3,445 descriptions produced by the director ($n = 40$), per participant across trials, $M = 14.35$, $SD = 5.71$).

2.4 Annotations

Following the approach proposed by Wilkes-Gibbs and Clark (1992) and Branigan et al. (2016), we coded directors' initial references as *definite* (definite articles, absence of article, proper nouns), and *indefinite* (indefinite articles). We also coded *hedge* forms such as "*un peu*" ("*a little*"), "*une espèce de*" ("*a kind of*"), and "*une sorte de*" ("*a sort of*"). Unlike Branigan et al. (2016), we expanded our hedges coding to encompass comparisons with "*comme*" ("*like*") and the use of the conditional "*on dirait que*" ("*it looks like*"). These hesitations and comparisons serve the same purpose of signaling the "*provisionality*" of the description until both interlocutors agree on one that satisfies them (Branigan et al., 2016).

For each tangram, we categorized the reuse of descriptions between the final description of one trial and the initial description of the subsequent in three categories derived from the proposals of Bortfeld and Brennan (1997) and Nadig et al. (2015). For each trial, we obtained 10 initial and 10 final descriptions, resulting in 2,000 instances of description reuse. We coded description reuses into three categories:

- The *identical* (IDENT) category included descriptions that conveyed equivalent information about the tangram by: (1) being word-for-word identical, (2) differing only in prepositions or determiners, or (3) using synonyms. This category reflects the reuse of knowledge already co-established, present in the participants' common ground.
- The *partial* (PARTIAL) category included descriptions that shared some elements but differed overall: (1) by including fewer elements, (2) by adding new elements, or (3) by combining both retained and newly introduced elements. This category reflects knowledge still being established and not yet fully integrated into the participants' common ground.
- The *different* (DIFFER) category included descriptions that introduced entirely new information without any shared words. This category reflects a lack of reference to the common ground.

The inter-coder reliability was strong for the annotation of reference markers (Cohen's kappa = 0.76) and the annotation of descriptions reuse (Cohen's kappa = 0.64).

In the present study, we measured the total number of words and the use of reference markers in the director's speech per trial, using three ratios: definite references to total descriptions (DEF ratio), indefinite references to total descriptions (IND ratio), and hedge references to total descriptions (HED ratio). We also assessed description reuse through three ratios: identical descriptions to total descriptions (IDENT ratio), partial descriptions to total descriptions (PARTIAL ratio), and different descriptions to total descriptions (DIFFER ratio).

2.5 Statistical analysis

We hypothesized that if the established common ground is not yet used by adolescents, they would use more words and employ more indefinite markers and hedges and less definite markers to describe tangrams than adults across all trials. To

evaluate task efficiency and the extent to which common ground was mobilized, we compared the total number of words between adolescents and adults across trials using a 2×6 repeated-measures ANOVA with Group (Adolescents, Adults) $\times$ Trials (Trials 1 to 6) as within-subject factors. We also compared the IND, DEF, and HED ratios between adolescents and adults across trials using a 2×6 repeated-measures ANOVA, with Group (Adolescents, Adults) $\times$ Trials (Trials 1 to 6) as within-subject factors.

To examine whether adolescents relied on a co-constructed common ground, we assessed whether they reused previously established descriptions or continued generating new ones across trials. We hypothesized that if adolescents did not use audience design in the same way as adults, their reuse of identical descriptions would be lower. To test this, we compared description reuse between adolescents and adults using a 2×5 repeated-measures ANOVA on the IDENT ratio, with Group (Adolescents, Adults) and Trials (Trials 2 to 6) as within-subject factors.

All data from participants assigned to the "director" condition were included in the analysis ($N = 40$). The analyses were conducted using the SPSS statistical software (IBM SPSS version 29).

3 Results

3.1 Total number of words

As shown in Figure 1, the analysis of the total number of words showed a main effect of Trial ($F(5, 190) = 122.335, p < 0.001, \eta_p^2 = 0.763$) with the number of words decreasing significantly across trials ($p < 0.001$). The difference was not significant between Trial 5 and Trial 6 ($p > 0.05$). A main effect of Group was found ($F(1, 38) = 4.137, p = 0.049, \eta_p^2 = 0.98$) with the number of words being significantly higher in adolescents than in adults across trials. The interaction Group $\times$ Trial was not significant ($F(5, 190) = 0.793, p > 0.05, \eta_p^2 = 0.020$).

3.2 Reference Markers

An illustrative example is provided in Appendix Table 2. As illustrated in Figure 2, the analysis of the DEF Ratio showed a main effect of Trial ($F(5, 190) = 97.264, p < 0.001, \eta_p^2 = 0.728$), with the DEF ratio being significantly lower in Trial 1 and Trial 2 compared to all subsequent

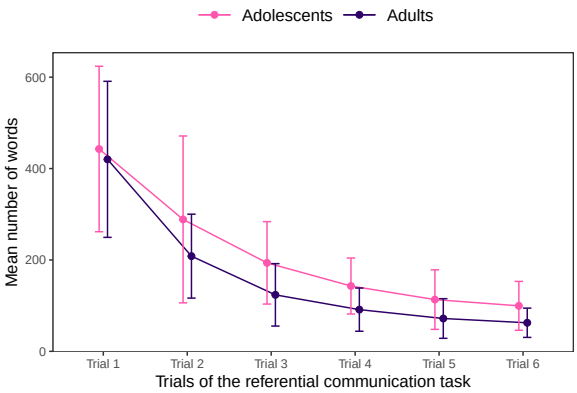

Figure 1: Mean of the total number of words by Trial and Group.

trials ($p < 0.001$) and in Trial 3 than Trial 6 ($p = 0.005$). A main effect of Group was found ($F(1, 36) = 7.007, p = 0.012, \eta_p^2 = 0.955$) with the DEF ratio being significantly lower in adolescents than in adults across trials. There was a marginal significant interaction Group $\times$ Trial ($F(5, 190) = 2.180, p = 0.058, \eta_p^2 = 0.054$) showing that in the Trial 1, there was no significant difference between adolescents and adults ($p > 0.05$), while in the subsequent trials the DEF ratio was lower in adolescents than in adults ($p < 0.04$).

The analysis of the HED Ratio showed a main effect of Trial ($F(5, 190) = 60.592, p < 0.001, \eta_p^2 = 0.615$), with the HED ratio being significantly higher in Trial 1 compared to all subsequent trials ($p < 0.001$), in Trial 2 compared to Trials 3 ($p = 0.020$), 4 ($p = 0.004$), 5 and 6 ($p < 0.001$). A main effect of Group was found ($F(1, 36) = 4.669, p = 0.037, \eta_p^2 = 0.109$), with the HED ratio being significantly higher in adolescents than in adults across trials. The interaction Group $\times$ Trial was not significant ($F(5, 190) = 1.350, p > 0.05, \eta_p^2 = 0.034$; see Figure 2).

The analysis of the IND Ratio showed a main effect of Trial ($F(5, 190) = 23.348, p < 0.001, \eta_p^2 = 0.381$), with the IND ratio being significantly higher in Trial 1 compared to all subsequent trials ($p < 0.001$), in Trial 2 compared to Trial 5 ($p = 0.025$), and Trial 6 ($p = 0.003$). A main effect of Group was found ($F(1, 38) = 6.383, p = 0.016, \eta_p^2 = 0.144$) with the IND ratio being significantly higher in adolescents than in adults across trials. The interaction Group $\times$ Trial was not significant ($F(5, 190) = 0.915, p > 0.05, \eta_p^2 = 0.024$; see Figure 2).

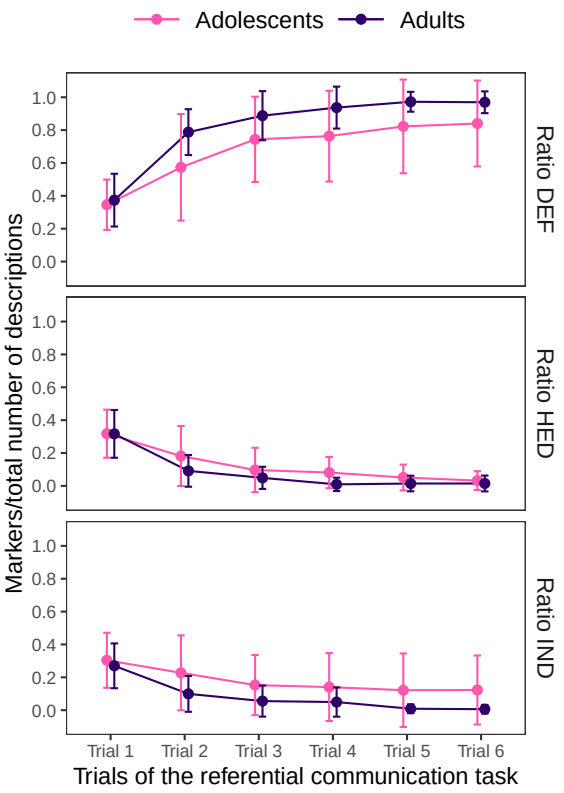

Figure 2: Ratio of each marker of reference by Markers Category, Trial and Group.

3.3 Description reuse of the descriptions across successive trials

3.3.1 IDENT ratio across trials

As shown in Figure 3, the analysis of the IDENT Ratio showed a main effect of Trial ($F(4, 152) = 92.337, p < 0.001, \eta_p^2 = 0.708$), with the IDENT ratio increasing significantly across all trials ($p < 0.001$). A main effect of group was found ($F(1, 38) = 3.991, p = 0.05, \eta_p^2 = 0.095$), with the IDENT ratio being significantly lower in adolescents than in adults across trials across trials. There was a significant interaction Group × Trial ($F(4, 152) = 5.189, p < 0.001, \eta_p^2 = 0.120$) showing that in adults, the IDENT ratio increased significantly across all trials ($p < 0.02$). In adolescents, the IDENT ratio increased significantly across all trials ($p < 0.01$) except between the Trial 3 and the Trial 4 ($p > 0.05$).

3.3.2 Reuse descriptions at Trial 4

As adolescents showed similar use of the identical category between the third and fourth trials, we focused the analysis on comparing the different categories of description reuse in the fourth trial. A two-way ANOVA with Group (Adolescents, Adults) × Category (IDENT, PARTIAL, DIFFER) was conducted on the ratio of each category to the total number of descriptions. An illustrative example is provided in Appendix Table 3.

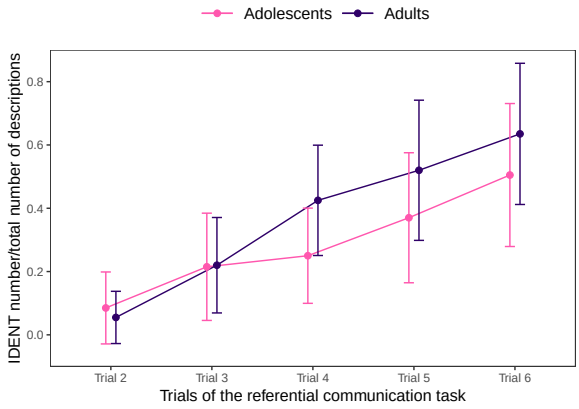

Figure 3: IDENT ratio/total number of descriptions by Trial and Group.

As illustrated in Figure 4, the analysis showed a main effect of Category ($F(5, 114) = 71.054, p < 0.001, \eta_p^2 = 0.588$), with the IDENT ratio being significantly lower than the PARTIAL ratio and significantly higher than the DIFFER ratio ($p < 0.001$), and the DIFFER ratio being significantly lower than the PARTIAL ratio ($p < 0.001$). The interaction Group × Category was significant ($F(5, 114) = 10.446, p < 0.001, \eta_p^2 = 0.155$) showing that in adults the DIFFER ratio was significantly lower than the IDENT and the PARTIAL ratios ($p < 0.001$), and the IDENT ratio did not differ with the PARTIAL ratio ($p > 0.05$). In adolescents, the IDENT ratio was significantly lower than the PARTIAL ratio ($p < 0.001$) and did not differ with the DIFFER ratio ($p > 0.05$). The DIFFER ratio was significantly lower than the PARTIAL ratio ($p < 0.001$). There was no main effect of Group ($F(5, 114) = 0.000, p > 0.05, \eta_p^2 = 0.000$; Figure 4).

4 Discussion

The aim of the present study was to assess whether young adolescents were able to use common ground co-established with their interlocutor during a referential communication task, in the same way as adults. We therefore analyzed the number of words and use of reference markers across trials, followed by the reuse of identical descriptions between trials from Trial 2 to Trial 6, and more precisely in trial 4, the use of identical, partial or different descriptions compared to those used in

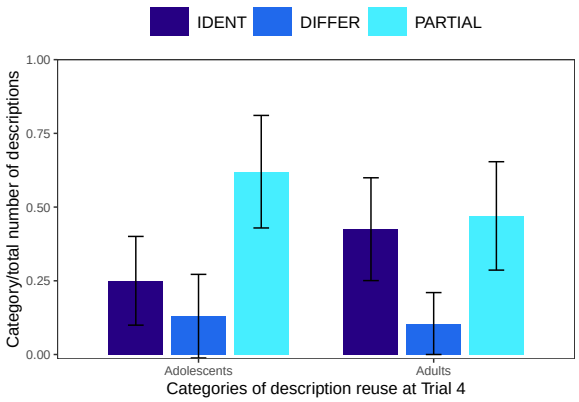

Figure 4: Ratio of each category of reuse of the descriptions between Trials 3 and 4, by Category and Group.

the trial 3.

The main results showed that, in adults, the use of reference markers corroborates findings from the literature: use of indefinites decreased (Bangerter et al., 2020) as did the hedges (Brennan and Clark, 1996), while the definites increased (Hupet et al., 1993). Adults progressively elaborated and relied on the common ground co-established with their interlocutor (Brennan and Clark, 1996). Similarly, we observed that adolescents were engaged in audience design to the extent that the indefinite references and hedges decreased in favor of an increase in definite markers, confirming the tendency observed in children aged 8 to 10 years (Branigan et al., 2016). However, across trials, adolescents used fewer hedges and indefinite references and more definite references compared to adults. Specifically, when comparing adolescents and adults, differences in the use of definite references emerged from the second trial onward, with adolescents producing fewer such markers and never reaching the level observed in adults in subsequent trials. Furthermore, although both groups produced fewer words across trials, adolescents consistently used more words than adults. These findings replicate previous results observed in adults (Clark and Wilkes-Gibbs, 1986; Isaacs and Clark, 1987) and indicate that adolescents are less efficient. These results suggest that, although young adolescents rely on the common ground established with their partner during the referential communication task, they do so less efficiently than adults. When examining the reuse of descriptions, we observed that adults increasingly reused identical descriptions as the trials progressed. This result suggests that, rather than merely reducing word production, adults tend to rely on previously established descriptions. Our observation aligns with findings from previous research in adults (Bortfeld and Brennan, 1997; Nadig et al., 2015). Similarly, among adolescents, we observed a comparable increase in the use of identical descriptions across trials, mirroring the pattern found in adults. However, from the fourth trial onward, adults consistently reused identical descriptions more frequently than adolescents, who never reached adult-level performance. The increase in identical description use suggests that both adults and adolescents had sufficiently refined their descriptions, making further simplification or elaboration unnecessary. The growing use of identical descriptions reflects task efficiency but also a more efficient use of common ground, as it is a more precise indicator of the shared knowledge established between participants. In the same way, at the midpoint of the task (Trial 4), adults used identical and partially elaborated descriptions equally, whereas adolescents continued to rely equally on new and partially elaborated descriptions.

Overall, our results do not support the idea that young adolescents are more egocentric. While previous studies have reported more pronounced egocentrism in younger adolescents compared to older ones (Arvidsson et al., 2022; Humphrey and Dumontheil, 2016), our findings indicate that adolescents are no more egocentric than adults. This discrepancy may be explained, first, by the type of shared knowledge involved. Shared knowledge was established through descriptions produced during the task and was solidified through a real-time co-construction. In contrast, previous studies manipulated encyclopedic knowledge (Arvidsson et al., 2022) or required adopting a visual perspective different from that of the participant (Humphrey and Dumontheil, 2016). Furthermore, differences observed with the director task may also be explained by variations in visuospatial or attentional cognitive abilities (Santiesteban et al., 2015). Secondly, while earlier studies found that younger adolescents were more egocentric than older ones, the attribution of knowledge was made about fictitious characters who did not respond, so their knowledge was merely assumed. This lack of interaction prevented the shared development of grounding and considered only the participant's beliefs about the character's knowledge. In contrast, our study involves real interaction and audience design engagement, allowing both interlocutors to contribute

to the co-construction and mutual adjustment of common ground. These factors could influence adolescents' performance. Our findings suggest that the cognitive processes involved are different and more demanding when the interlocutor is not physically present.

In our experimental setup, the common ground between participants consisted of both pre-existing shared knowledge (including prior knowledge about one's interlocutor and encyclopedic knowledge) and knowledge co-established during the task. It also encompasses non-verbal cues, such as gestures, gaze, and facial expressions, produced by the interlocutors (Brennan et al., 2010). However, in everyday conversation, common ground extends beyond what is explicitly said. Unlike in real-life interactions, our participants could not see each other and therefore lacked access to many of the grounding cues provided by non-verbal information. The inclusion of such cues, which contribute to mutual understanding between interlocutors in everyday conversations, could help reduce the performance gap with adults by improving the grounding process.

To conclude, our findings on referential marker use and the reuse of previous descriptions suggest that adolescents actively contribute to establishing common ground, similarly to adults. However, their strategies seem suboptimal and depend on knowledge still in the process of being co-constructed. We interpret these results as evidence that collaborative elaboration of common ground is present in young adolescents but is less efficient than in adults.

Acknowledgments

This research was funded by the French National Research Agency (ANR) under the grant ANR-24-CE28-6883-01 (COOL Project). It was carried out within the Institute of Convergence ILCB (ANR-16-CONV-0002) and has also benefited from support from the French government (France 2030), managed by the French National Agency for Research (ANR) and the Excellence Initiative of Aix-Marseille University (A*MIDEX).

References

Caroline Arvidsson, David Pagmar, and Julia Uddén. 2022. When did you stop speaking to yourself? age-related differences in adolescents' world knowledge-based audience design. *Royal Society Open Science*, 9(11):220305.

Adrian Bangerter, Eric Mayor, and Dominique Knutsen. 2020. Lexical entrainment without conceptual pacts? revisiting the matching task. *Journal of Memory and Language*, 114:104129.

Paul Boersma and David Weenink. 2024. Praat. `https://github.com/praat/praat`.

Heather Bortfeld and Susan E. Brennan. 1997. Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes*, 23(2):119–147.

Vincent Bovet, Dominique Knutsen, and Marion Fossard. 2024. Direct and indirect linguistic measures of common ground in dialogue studies involving a matching task: A systematic review. *Psychonomic Bulletin & Review*, 31(1):122–136.

Holly P. Branigan, Jenny Bell, and Janet F. McLean. 2016. Do you know what i know? the impact of participant role in children's referential communication. *Frontiers in Psychology*, Volume 7 - 2016.

Susan E. Brennan and Herbert H. Clark. 1996. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6):1482–1493.

Susan E. Brennan, Alexia Galati, and Anna K. Kuhlen. 2010. Chapter 8 - two minds, one dialog: Coordinating speaking and understanding. In Brian H. Ross, editor, *The Psychology of Learning and Motivation: Advances in Research and Theory*, volume 53 of *Psychology of Learning and Motivation*, pages 301–344. Academic Press.

Sarah Brown-Schmidt and Daphna Heller. 2018. 549perspective-taking during conversation. In *The Oxford Handbook of Psycholinguistics*. Oxford University Press.

Maud Champagne-Lavau, Marion Fossard, Guillaume Martel, Cimon Chapdelaine, Guy Blouin, Jean-Pierre Rodriguez, and Emmanuel Stip. 2009. Do patients with schizophrenia attribute mental states in a referential communication task ? *Cognitive Neuropsychiatry*, 14(3):217–239.

Herbert H. Clark and Gregory L. Murphy. 1982. Audience design in meaning and reference. In Jean-François Le Ny and Walter Kintsch, editors, *Advances in Psychology*, volume 9 of *Language and Comprehension*, pages 287–299. North-Holland.

Herbert H. Clark and Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition*, 22(1):1–39.

Martina De Lillo and Heather J. Ferguson. 2023. Perspective-taking and social inferences in adolescents, young adults, and older adults. *Journal of Experimental Psychology: General*, 152(5):1420–1438.

Iroise Dumontheil, Ian A Apperly, and Sarah-Jayne Blakemore. 2010. Online usage of theory of mind continues to develop in late adolescence. *Developmental science*, 13(2):331–338.

Kumiko Fukumura. 2016. Development of audience design in children with and without asd. *Developmental Psychology*, 52(1):71–87.

Adriana Galvàn. 2021. Adolescent brain development and contextual influences: A decade in review. *Journal of Research on Adolescence*, 31(4):843–869.

Gillian Humphrey and Iroise Dumontheil. 2016. Development of risk-taking, perspective-taking, and inhibitory control during adolescence. *Developmental Neuropsychology*, 41(1-2):59–76.

Michel Hupet, Yves Chantraine, and François Nef. 1993. References in conversation between young and old normal adults. *Psychology and Aging*, 8(3):339–346.

Ellen Isaacs and Herbert Clark. 1987. References in conversation between experts and novices. *Journal of Experimental Psychology General*, 116:26–37.

Boaz Keysar, Dale J. Barr, Jennifer A. Balin, and Jason S. Brauner. 2000. Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11(1):32–38.

Emma J. Kilford, Emily Garrett, and Sarah-Jayne Blakemore. 2016. The development of social cognition in adolescence: An integrated perspective. *Neuroscience & Biobehavioral Reviews*, 70:106–120. Series: The Adolescent Brain.

Robert M. Krauss and Sidney Weinheimer. 1964. Changes in reference phrases as a function of frequency of usage in social interaction: a preliminary study. *Psychonomic Science*, 1(1):113–114.

Aparna Nadig, Shivani Seth, and Michelle Sasson. 2015. Global similarities and multifaceted differences in the production of partner-specific referential pacts by adults with autism spectrum disorders. *Frontiers in Psychology*, Volume 6 - 2015.

Idalmis Santiesteban, Punit Shah, Sarah White, Geoffrey Bird, and Cecilia Heyes. 2015. Mentalizing or submentalizing in a communication task? evidence from autism and a camera control. *Psychonomic Bulletin & Review*, 22(3):844–849.

Nickolay V. Shmyrev and other contributors. 2020. Vosk speech recognition toolkit: Offline speech recognition api for android, ios, raspberry pi and servers with python, java, c# and node. `https://github.com/alphacep/vosk-api`.

Irene Symeonidou, Iroise Dumontheil, Wing-Yee Chow, and Richard Breheny. 2016. Development of online use of theory of mind during adolescence: An eye-tracking study. *Journal of Experimental Child Psychology*, 149:81–97. Series: Theory of Mind in Middle Childhood and Adolescence.

Christian K. Tamnes, Knut Overbye, Lia Ferschmann, Anders M. Fjell, Kristine B. Walhovd, and Sarah-Jayne Blakemore. 2018. Social perspective taking is associated with self-reported prosocial behavior and regional cortical thickness across adolescence. *Developmental Psychology*, 54(9):1745–1757.

Deanna Wilkes-Gibbs and Herbert H. Clark. 1992. Coordinating beliefs in conversation. *Journal of Memory and Language*, 31(2):183–194.

A Appendix: Examples of annotated excerpts from the dataset.

Trial	Initial description	Number of words	Annotation
1	**On dirait** quelqu'un qui va mettre un coup de pied ***It looks like*** *someone's about to kick*	10	HED
2	**Celui** qui met le coup de pied ***The one who*** *kicks*	7	DEF
3	**Le** coup de pied ***The*** *kick*	4	DEF
4	Coup de pied *Kick*	3	ABS
5	**Le** coup de pied ***The*** *kick*	4	DEF
6	Coup de pied *Kick*	3	ABS

Table 2: Example illustrating the annotation of referential markers in a participant's initial description of the same tangram.

Comparison	Initial description	Final description from previous trial	Annotation
Trial 1 to 2	Celui qui met le coup de pied *The one who kicks*	Avec le carré qui tombe là *With the square that falls here*	DIFFER
Trial 2 to 3	Le coup de pied *The kick*	Celui qui met le coup de pied *The one who kicks*	PARTIAL
Trial 3 to 4	Coup de pied *Kick*	Le coup de pied *The kick*	IDENT

Table 3: Example illustrating the annotation of intra-participant description reuse for a tangram.