Applied Psycholinguistics 39 (2018), 1345–1371
doi:10.1017/S0142716418000358

An investigation of the determinants of dialogue navigation in joint activities

DOMINIQUE KNUTSEN
University of Essexefere

GILLES COL and LUDOVIC LE BIGOT
Université de Poitiers

Received: August 30, 2017 Revised: June 13, 2018 Accepted: June 14, 2018

ADDRESS FOR CORRESPONDENCE
Dominique Knutsen, University of Essex, Wivenhoe Park, Colchester CO4 3SQ United Kingdom. E-mail: dknutsen@essex.ac.uk

ABSTRACT
When people engage in joint activities together, they use dialogue, and more specifically project markers such as *yeah*, *okay*, or *uh-huh*, to coordinate entrances and exits of projects and subprojects. The purpose of the current study was to examine how two features of the dialogue situation, namely, mental load and face visibility, affect project marker production. Pairs of participants performed a collaborative puzzle game together. Mental load was manipulated through time pressure; visibility was manipulated by allowing the participants to see each other's face during the task, or not. Dialogues were transcribed and coded for project marker production. Project marker production was found to *increase* under mental load; this also depended on the role of the speaker in the dyad (Director or Matcher) and on face visibility. This sheds light on the idea that dialogue partners may behave *more* collaboratively when experiencing high levels of mental load, contributing to a better understanding of mental resource allocation in dialogue-based joint activities.

Keywords**:** collaboration; dialogue; joint activities; mental load; project markers

Human dialogue is a complex activity during which two partners or more attempt to understand each other. In order to do so, they use project markers such as *yeah*, *okay*, *uh-huh*, or *right* to navigate the interaction and keep track of its progress (e.g., Bangerter & Clark, 2003; Bangerter, Clark, & Katz, 2004; Branigan, Catchpole, & Pickering, 2011; Fox Tree, 2010; Tolins & Fox Tree, 2014). Because these markers play a central role in communication success, it is essential to identify the determinants of project marker production in dialogue. The purpose of the current study is to build on and extend previous findings by examining the effects of mental load and face visibility on project marker production, thus contributing to a better understanding of human dialogue.

In everyday life, most people have the opportunity to engage in a number of joint activities with others. For instance, two people may play a piano duet together, plan a journey around the world, or attempt to find the best time and date to schedule a meeting (e.g., Bangerter & Clark, 2003; Bangerter et al., 2004;

Clark, 1996, 2005; Garrod & Pickering, 2009; Sebanz, Bekkering, & Knoblich, 2006). Most of these activities (although not all of them) involve dialogue. For instance, planning a journey usually involves deciding how to travel, where to stay, and which places to visit (see Bangerter & Clark, 2003; Clark, 1996; Mills, 2014; Tylén, Weed, Wallentin, Roepstorff, & Frith, 2010).

Any joint activity, whatever its content and nature, may be divided into a number of projects, which must be completed jointly in order for both partners to ultimately complete the activity. For instance, if two people are trying to schedule a meeting together, they would need to find a suitable day, a suitable time, and they would need to book a room to hold the meeting. What is more, any project may itself be divided into a number of subprojects. For instance, the subproject "finding a suitable day to hold the meeting" would normally involve coming up with a number of potential dates, taking into account the attendees' availabilities and identifying the most suitable day. Once again, all of these subprojects would need to be completed before the overall project would be deemed to be completed. Just like individual activities, joint activities (and the corresponding projects and subprojects) can be conceptualized as part of the same hierarchy in which projects are nested within joint activities and in which subprojects are nested within projects (Bangerter & Clark, 2003; Clark, 1996). An example is shown in Figure 1.

Because joint activities are carried out by at least two partners, these partners need to coordinate project and subproject entrances and exits. Imagine that A and B are trying to schedule a meeting, as illustrated in Figure 1. If A considers that subproject 1.1 (coming up with several different dates) has been completed and has now moved on to subproject 1.2 (considering the attendees' availabilities), but that B considers that subproject 1.1 has not been completed (e.g., because B is not happy with A's suggestions and would like to consider alternative dates), then A and B would find it difficult to navigate the hierarchy of projects and subprojects together. The partners thus need a means of entering and exiting projects

Joint activity: Scheduling a meeting

- Project #1: Find a suitable date
    - Subproject #1.1: come up with several potential dates
    - Subproject #1.2: consider attendees' availabilities
    - Subproject #1.3: identify most suitable day
- Project # 2: Find a suitable time
    - Subproject #2.1: come up with several potential times
    - Subproject #2.2: consider attendees' availabilities
    - Subproject #2.3: identify most suitable time
- Project #3: Book a room to hold the meeting
    - Subproject #3.1: determine which rooms are available on chosen day/time
    - Subproject #3.2: select most suitable room based on number of attendees and other potential constraints

Figure 1. Diagram representing the project and subproject hierarchy emerging from the "scheduling a meeting" joint activity (adapted from Bangerter & Clark, 2003). In this example, the partners' joint activity involves three main projects, which are themselves divided into a number of subprojects. Subprojects are nested within projects, which are themselves nested within the partners' joint activity.

and subprojects jointly and explicitly. According to Bangerter and Clark (2003, see also Bangerter et al., 2004), dialogue plays a central role here. Specifically, partners use project markers to navigate across and between levels of hierarchies of joint activities, projects, and subprojects.

NAVIGATING PROJECTS AND SUBPROJECTS IN JOINT ACTIVITIES

Project markers are a subcategory of discourse markers. Discourse markers, which include words such as *yeah*, *anyway*, *possibly*, or *well* (as well as many other words), have been defined as devices used by partners to mark important multidimensional transitions in discourse (Jefferson, 1984; Louwerse & Mitchell, 2003; Schegloff, 1982). Their function is to specify the relationship between a new segment of discourse and the preceding segment. Their core meaning is procedural rather than conceptual; that is, their specific interpretation is negotiated depending on the linguistic and conceptual context in which the interaction takes place (see Fischer, 2006; Fraser, 1999). Moreover, discourse markers enable the partners to ensure that the information under discussion has been grounded properly; that is, they enable the partners to ensure that they have understood each other well enough for current purposes (Fox Tree, 2010). For instance, the discourse marker *yeah* suggests that the partner producing it has understood what was previously said; the discourse marker *I mean* may be used to fix mistakes when they occur (e.g., this discourse marker may be used to introduce reformulations in situations where an utterance was not understood correctly initially).

As for project markers, they include words and expressions such as *uh-huh*, *yeah*, *right*, *okay*, *got it*, or *all right* (Bangerter & Clark, 2003; Bangerter et al., 2004; Branigan et al., 2011; Fox Tree, 2010; Tolins & Fox Tree, 2014). Some project markers (e.g., *okay*, *got it*, or *all right*) are mainly used in *vertical* transitions across levels of the hierarchy. For instance, in Figure 1, *okay* may be used to exit subproject 1.3 and/or to enter project 2 (i.e., *okay* may be used to indicate that the most suitable date has been identified, and hence that the entire "find a suitable day" project has been completed, and/or to indicate that the dialogue partners are now moving on to the "find a suitable time" project). Other project markers (e.g., *uh-huh* and *yeah*) are mainly used in *horizontal* transitions within the same level of the hierarchy. For instance, in Figure 1, *uh-huh* may be used to exit subproject 2.1 and/or to enter subproject 2.2 (i.e., one of the dialogue partners might list all potential times and the other partner may acknowledge this list by saying *uh-huh*, thus enabling the partners to move on to considering the attendees' availabilities).[1]

Project markers, as well as discourse markers, are what Clark (1996) would call "track 2" signals. Their production does not usually contribute to the content of the interaction (which mainly includes "track 1" signals). Rather, its main function is to keep both dialogue partners on track, and to ensure that the interaction goes smoothly. It is also important to highlight here that dialogic project marker production is collaborative, that is, these markers are usually produced in order to facilitate communication. Other dialogic behaviors have been described as collaborative: for instance, audience design, or partner adaptation, is a mechanism whereby dialogue partners adapt their speech to each other's understanding (e.g., Gorman, Gegg-Harrison, Marsh, & Tanenhaus, 2013; Isaacs & Clark, 1987; Vanlangendonck, Willems, Menenti, & Hagoort, 2016). An example of project navigation is provided in Example 1, which is taken from the corpus gathered in the present study. In this study, one participant, who played the role of Director, gave instructions to another participant, who played the role of Matcher, to enable the latter to arrange the pieces of a puzzle in a predefined way. In other words, the partners' joint activity consists in completing the puzzle together. In Example 1, the participants are discussing the puzzle shown in Figure 2.

Example 1

[…]

(1) Director: then the square that you've got

(2) Matcher: *uh-huh*

(3) Director: you should match it left of that the small triangle

(4) Matcher: so the there's a so there's a gap where the parallelogram is

(5) Director: *yes*

(6) Matcher: *yeah*

[…]

(9) Director: *okay* and then underneath that square erm the medium-sized triangle erm triangle

[…]

In this extract, the Director and the Matcher are attempting to complete a project in which the square piece must be positioned correctly. This project is divided into a number of subprojects. First, the Director attempts to get the

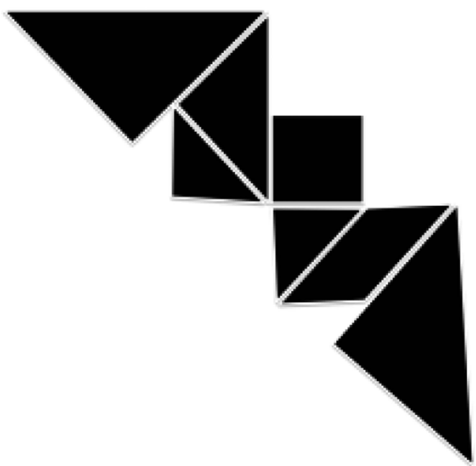

Figure 2. Puzzle discussed by the participants in Example 1.

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press

Matcher to identify the square (speech turn 1); the Matcher indicates that the subproject is completed by saying *uh-huh* (speech turn 2). Second, the Director attempts to tell the Matcher where the square is positioned (speech turn 3). However, instead of acknowledging this information and moving on to the next subproject, the Matcher initiates a new subproject in which he or she asks the Director for more information (speech turn 4); the Director completes this project by answering the Matcher's question (speech turn 5). At this point (speech turn 6), the Matcher can finally complete the subproject initiated in speech turn #3 by saying *yeah*. It is noteworthy that up until this point, only horizontal markers have been produced, suggesting that all of the Directors' instructions were conceptualized by the partners as subprojects that were all part of the same project. Finally, in speech turn 9, the Director indicates that he or she believes that the project "getting the Matcher to place the square correctly" is now complete; he or she thus produces the marker "okay" to mark a vertical transition and to move on to the next project.

DETERMINANTS OF PROJECT MARKER PRODUCTION IN DIALOGUE

As mentioned in the previous section, project markers are dialogue markers that are used by partners engaged in a joint activity to navigate hierarchies of projects and subprojects. These markers are believed to play a central role in dialogue success, as they enable partners to keep track of their progress in the subprojects and projects at hand. A systematic investigation of the determinants of project marker production was conducted by Bangerter and Clark (2003; see also Bangerter et al., 2004). In this study, the authors reanalyzed the data from a number of previous dialogue studies (Bangerter & Smolenski, 2000; Clark & Krych, 2004; Gross, Allen, & Traum, 1993) and preexisting corpora (Anderson et al., 1991; Barnard, 1974; Godfrey, Holliman, & McDaniel, 1992; Svartvik & Quirk, 1980). One of the main findings was that mutual visibility has an important influence on project marker production. For instance, in Clark and Krych's (2004) study, one participant (the Director) gave another participant (the Matcher) instructions to enable him or her to build a Lego model. In the visible condition, the Director could see the Matcher's progress during the task, but this was not the case in the hidden condition. Bangerter and Clark found that more *uh-huhs* and *okays* were produced in the hidden condition, because the participants needed to acknowledge each other's speech explicitly in this condition. Another important finding was that the type of dialogue also affects project marker production. For instance, *okay* is produced more often in the context of well-defined tasks (e.g., Clark & Krych's, 2004, Lego task), whereas *uh-huh* and *yeah* are less sensitive to the features of the interactional setting.

The purpose of the current study is to investigate project marker production further, by examining how the features of the situation in which a joint activity is carried out may affect project navigation. The first feature of interest was the mental load associated with achieving a task. Mental load characterizes the demand imposed by any task on a person's limited mental resources; when

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press

the demand exceeds the pool of resources available, the person's performance is usually affected negatively (Wickens, 2008; see also Hart & Staveland, 1988). Any dialogue involves a fairly high level of mental load, due to the demands of speaking and listening (see Boiteau, Malone, Peters, & Almor, 2014).[2] What is more, depending on the situation in which the dialogue is carried out, the mental load experienced by each partner varies (e.g., this might depend on the difficulty and the novelty of the task at hand, or on whether or not the partners feel time-pressured to complete the joint activity). A number of studies have investigated the effect of mental load or cognitive load on various features of spoken discourse. For instance, mental load may narrow the scope of advance planning in speech production (Wagner, Jescheniak, & Schreifers, 2010) and increase naming latencies (Belke, 2008). Speakers also tend to produce more sentence fragments (e.g., utterances exhibiting an incomplete syntactic structure), false starts, syntax errors (e.g., subject–verb agreement mistakes; Hartsuiker & Barkhuysen, 2007), pauses, and repetition when they experience mental load (e.g., see Berthold & Jameson, 1999; Howarth & Anderson, 2007; Villing, 2009). Mental load has also been shown to affect comprehension (e.g., mental load affects the syntactic processing of ambiguous sentences; see, e.g., Vos, Gunter, Schriefers, & Friederici, 2010). Other studies have sought to determine how mental load might interfere with higher level processes at play during dialogue such as partner adaptation. For instance, in a study conducted by Rossnagel (2000), a participant gave instructions to a confederate (an adult or a child) to build a machine model. Mental load was manipulated either by asking the participants to describe the assembly of the model from memory or by asking them to carry a mental load of seven digits as they were producing the instructions. The main finding was that partner adaptation (e.g., participants adapting their use of complex, technical terms to the confederate's likely level of knowledge) decreased as mental load increased. This finding suggests that partner adaptation does not occur automatically: rather, it mainly occurs where dialogue partners have enough resources, time, and motivation to adapt to their partners (see Cane, Ferguson, & Apperly, in press; Horton & Keysar, 1996). However, none of these studies focused on the influence of mental load on the production of project markers per se. Comparing project marker production in a situation where two partners are under mental load and in a situation where two partners are not under mental load will enable us to determine whether joint activity navigation, just like partner adaptation, depends on the amount of mental resources available to the partners.

In addition to the mental load experienced by the dialogue partners, another feature of the interaction situation that is likely to affect the partners' behavior is whether they can see each other or not; the effect of mutual visibility was thus also investigated in the present study. As mentioned previously, there is already evidence that such visibility affects dialogue navigation (Bangerter & Clark, 2003; Clark & Krych, 2004), as dialogue partners need to acknowledge information more explicitly when they cannot see each other. This is in line with the more general idea that visibility affects the way in which dialogue partners ground information (Clark & Brennan, 1991; Kraut, Fussell, & Siegel, 2003; Whittaker, 2003). However, contrary to most work on mutual visibility conducted in this field of research, we did not manipulate whether or not the participants could see each other's progress in the task at hand. Instead, we manipulated whether or not the partners could see each other's face. The purpose of this was to determine whether, just like when dialogue partners can see their partner's progress in the task at hand (Clark & Krych, 2004), face visibility also affects project marker production (e.g., partners could monitor their partner's face to determine whether or not he or she is ready to embark on the next project/subproject, thus making the production of project markers less necessary in cases where partners see each other's face). What is more, studying the influence of face visibility on project marker production has important implications for mediated communication (e.g., phone conversations and video calls). Different communication media have different features: for instance, videos calls enable dialogue partners to see each other's face whereas phone conversations do not. Thus, comparing project marker production in situations where dialogue partners can see each other's face, or not, will enable us to determine whether this specific feature is likely to affect project navigation in mediated communication.

An experiment was conducted in which pairs of participants embarked on a well-defined, goal-oriented task. Specifically, the collaborative puzzle-solving task used was a variation of the standard matching task that is often used in dialogue research (see, e.g., Achim, Achim, & Fossard, 2017; Clark & Wilkes-Gibbs, 1986; Hupet, Seron, & Chantraine, 1991; Isaacs & Clark, 1987; Krauss & Weinheimer, 1966; Yoon & Brown-Schmidt, 2014). In this variation, a Director was given the solution to a puzzle and gave instructions to a Matcher in order to enable the latter to complete the puzzle. Mental load was manipulated by putting the pairs under time pressure half of the time (see Gerjets & Scheiter, 2003; Hancock & Caird, 1993; Horton & Keysar, 1996; Howarth & Anderson, 2007; Ordonez & Benson, 1997; Paas, Tuovinen, Tabbers, & Van Gerven, 2003, for similar manipulations of mental load). In addition, half of the pairs could see each other's face while interacting whereas the other half could not. These two manipulations (mental load and face visibility) were chosen because their effects on aspects of dialogue other than project marker production are well documented in the literature (e.g., Clark & Krych, 2004; Rossnagel, 2000).

Two hypotheses were compared as to the influence of mental load on project marker production. The first hypothesis was that participants produce *fewer* project markers while under mental load than while under no mental load, as mental load decreases the amount of resources available to produce project markers. This hypothesis is consistent with the idea that, just like other collaborative behaviors such as partner adaptation, task navigation depends on the amount of mental resources available to the dialogue partners (see Horton & Keysar, 1996; Rossnagel, 2000). However, an alternative hypothesis (our second hypothesis) is that participants produce *more* project markers while under mental load than while under no mental load. Although this possibility has seldom been examined experimentally, this would be consistent with a strong collaborative view of dialogue, whereby the difficulties experienced by dialogue partners lead them to make additional individual efforts to ensure mutual comprehension (e.g., Clark, 1996; Clark & Wilkes-Gibbs, 1986).

As for face visibility, the third hypothesis tested in this study was that participants produce fewer project markers when they can see each other's face than when they cannot, as they may use their partner's facial expression (instead of project markers) to determine whether or not he or she is ready to move on to the next project or subproject.

Finally, interactions between the two factors of interest (mental load and face visibility) were also tested. Examining these enabled us to determine whether the effect of mental load depended on whether the participants could see each other's face or not, and/or whether the effect of face visibility depended on the amount of mental load experienced by the participants. The first interaction hypothesis was that the effect of mental load is reduced when the participants can see each other's face, because seeing one's partner's face means that the participants rely less on project markers, and more on visual cues, to navigate the interaction (Bangerter & Clark, 2003). The second interaction hypothesis was that the effect of face visibility is increased when the participants experience high levels of mental load, because an increase in mental load should lead dialogue partners to use as many cues as possible to make sure that the interaction goes smoothly.

METHOD

Participants

Sixty-two participants, divided into 31 pairs, initially took part in the study. There were 52 women and 10 men; their average age was 20.43 years ($SD = 3.05$). All participants were native English speakers, except 1 (the data from this participant's dyad were removed from the data set). They took part in the study for a small payment (£6) or received partial course credit. All participants signed an informed consent form before taking part in the study.

Apparatus

The interactions between the participants were recorded using a double-entry Tascam DR-40 digital voice recorder.

Materials

Eight tangram figures, each made of the same seven pieces (two small triangles, one medium triangle, two large triangles, one square, and one parallelogram) were selected randomly for use in this study (see Figure 2 for an example). Each figure was printed on a laminated A5 sheet of paper. The seven loose pieces were also printed out and laminated.

A simplified paper-and-pen version of the NASA-TLX mental workload questionnaire (see below; see also online-only Supplemental Materials) was also used as part of the manipulation check in this experiment.

Task and procedure

At the time of recruitment, the participants were informed that the experiment would last less than an hour, but they received no other information about the timing of the experiment, as this could have interfered with the time pressure manipulation. The experiment took place in a quiet experimental room. In the "with visibility" condition, the participants sat at the same table, facing each other. A low partition prevented them from seeing each other's workspace, but they could see each other's face. In the "without visibility" condition, the participants sat at two different tables and faced different directions so that they could hear but not see each other. One of the participants was informed that he or she was the Director and the other participant was informed that he or she was the Matcher in the experiment. Roles were allocated at random. The participants did not switch roles during the experiment.

Main task. The Director was given a tangram figure and gave the Matcher instructions in order to enable him or her to assemble the same figure. The Matcher was encouraged to provide feedback to the Director and to ask as many questions as necessary. In the "with visibility" condition, the participants were told that they were not allowed to look over the partition, or to draw the figures in the air. In the "without visibility" condition, the participants were told that they were not allowed to look at each other at all. The experimenter stayed in the room with them during the whole experiment to ensure that these instructions were followed. When the participants believed that they had completed a puzzle, they asked the experimenter to check that the Matcher had assembled the figure correctly. If the Matcher had got the figure right, the participants then moved on to the next figure. If the Matcher had got the figure wrong, the participants were required to try again. The eight pictures were presented to the Director in one of eight random orders.

The main task was divided into two sessions of 10 min. Half the dyads performed the first session under time pressure ("with time pressure" condition) and the second session without time pressure ("without time pressure" condition); the other half performed the first session without time pressure and the second session under time pressure. In the "with time pressure" condition, the participants were informed before they started the task that they had a maximum of 10 min to complete as many figures as possible; what is more, the experimenter would remind them of how much time they had left every 2 min, to increase time pressure. In the "without time pressure" condition, the participants were not informed that there was a time limit before they start the task; the experimenter interrupted them after 10 min.

Manipulation check. A simplified version of the NASA-TLX mental workload questionnaire (e.g., Hart, 2006; Hart & Staveland, 1988; Rubio, Díaz, Martín, & Puente, 2004; Sato et al., 1999) was administered to the participants at the end of each session in order to make sure that the mental load manipulation was effective. The questionnaire included six 20-point scales pertaining to the mental, physical, and temporal demands of the task and to the participants' perceived performance, effort, and frustration level (the full questionnaire can be found in

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press

the online-only Supplemental Materials). Each participant was required to rate his or her personal experience during the session on each scale. The participants had access to the definitions of each dimension as they completed the questionnaire. They were not allowed to communicate during this part of the experiment.

The experiment lasted approximately 45 min. The participants were fully debriefed after the end of the study.

Data coding

Project markers produced during the main task. The dialogues between the participants were transcribed (if the participants talked to the experimenter during the experiment, their queries and the experimenter's responses were not transcribed). No punctuation marks, apart from apostrophes to show possession or in contracted verb forms and negative verbs, and hyphens, were used in the transcriptions. The corpus is available online (http://researchdata.essex.ac.uk/).

Transcripts were then coded for horizontal and vertical project markers (Bangerter & Clark, 2003). The horizontal marker category included *yeah* (and similar forms such as *yes* and *yep*; for readability purposes, the term *yeah* is used to describe all three forms in the remainder of the article) and *uh-huh*. The vertical marker category included *okay*, *right*, *alright*, and *got it* (and similar forms such as *I've got it*, *got you*, and *got that*; for readability purposes, the term *got it* is used to describe all four forms in the remainder of the article). The coding scheme was based exclusively on the lexical items produced by the participants, rather than on the context in which these were produced. For instance, *yeah* was always coded as horizontal and *okay* was always coded as vertical, regardless of the context in which these were produced. The coding scheme is illustrated in Examples 2, 3, and 4.

Example 2:

(1) Matcher: so that the it it's pointing outwards yeah
(2) Director: yeah
(3) Matcher: okay

Example 3:

(1) Director: there is a the the point on the left side of it against the ground you know what I mean
(2) Matcher: no
(3) Director: erm using a triangle erm yeah face it face the triangle erm how do I say this okay alright two using two triangles face them oppositely erm as if you're gonna be creating a square from it so one facing each other okay
(4) Matcher: uh-huh

Example 4:

(1) Director: okay right let's think of this now okay so I assume that the square is probably going to be correct

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press

In Example 2, the Matcher produced one horizontal marker (*yeah*) and one vertical marker (*okay*); the Director only produced one horizontal marker (*yeah*). In Example 3, the Director produced three vertical markers (two *okays* and one *alright*) and one horizontal marker (*yeah*); the Matcher only produced one horizontal marker (*uh-huh*). Finally, in Example 4, the Director produced three vertical markers (two *okays* and one *right*).

Mental workload scores. The participants' responses on the six scales of the NASA-TLX questionnaire were averaged, thus resulting in two overall mental workload scores for each participant (one per session).

Experimental design

There were two main independent variables (IVs) in this study: visibility ("with visibility" or "without visibility") was a between-participants IV and mental load ("with time pressure" or "without time pressure") was a within-participants IV.

Two additional IVs were also included in the analyses. First, participant role ("Director" or "Matcher") was also included as an additional between-participants IV in order to determine whether or not the Directors and the Matchers produced the same proportion of project markers. What is more, to anticipate the results, it was found that the success rate differed across conditions (it was particularly low in the "with time pressure–without visibility" condition). Failure to complete the task might have affected project marker production, thus potentially interfering with the effects of visibility and mental load. In order to account for this, task success (success or failure) was also included as a within-participants IV in order to determine whether the effects of the two main IVs remained statistically significant even when this additional variable was controlled for.

The first dependent variable was the total proportion of project markers produced by each participant. This proportion was obtained by dividing the number of project markers produced by each participant by the total number of words produced by each participant, for each puzzle discussed during the experiment. The proportions of horizontal and vertical project markers were then examined separately and formed the second and third dependent variable in this study, respectively.

RESULTS

Characterization of the corpus (descriptive statistics)

Success rate. In Session 1, 13 pairs managed to complete at least one puzzle, whereas the remaining 17 pairs failed to complete a puzzle. In Session 2, 14 pairs managed to complete at least one puzzle, whereas the remaining 16 pairs failed to complete a puzzle. The average number of puzzles completed successfully in each session by each pair in all four experimental conditions is reported in Table 1.

Speech turn, word, and project marker production. The Collaborative Puzzle Game Corpus includes 83,173 words and 10,385 speech turns. A total of 6,397

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press

Table 1. *Average number of puzzles completed as a function of mental load and visibility*

	With time pressure	Without time pressure	Total
With visibility	0.64 (0.63)	0.57 (0.94)	**0.61 (0.79)**
Without visibility	0.31 (0.48)	0.56 (0.51)	**0.44 (0.50)**
Total	**0.47 (0.57)**	**0.57 (0.73)**	**0.52 (0.65)**

Note: Standard deviations are reported in brackets.

project markers were produced by the participants. There were 3,321 *yeahs*, 489 *uh-huhs*, 1978 *okays*, 343 *rights*, 178 *got its*, and 88 *alrights*. As mentioned previously, *yeah* and *uh-huh* are used as horizontal project markers, and *okay*, *right*, *got it*, and *alright* are used as vertical project markers. Thus, 59.56% of the project markers in the corpus were horizontal project markers and the remaining 40.44% were vertical project markers.

Rationale of the main analysis

The data were analyzed using linear mixed models in SAS 9.4 (GLIMMIX procedure). The main difference between mixed models and more standard analyses such as analysis of variance or regression is that mixed models include random intercepts (which account for potential variability across analysis units) and random slopes (which account for potential variability in the units' sensitivity to the IVs included in the analysis). According to Barr, Levy, Scheepers, and Tily (2013), all random effects justified by the experimental design should be included in the analysis. In this study, this would have required the inclusion of by-item (the items were the eight pictures used), by-participants, and by-dyad (where applicable) random intercepts as well as all by-item, by-participant, and by-dyad random slopes corresponding to within-units IVs. However, not all random effects significantly contribute to the model, in which case they may be removed without affecting the parameters of the model (keeping them in the analysis would cause the model to fail to converge; Kiernan, Tao, & Gibbs, 2012). Accordingly, for each analysis, we started with the maximal random effects structure. The random effects that did not significantly contribute to the model were then identified (this is performed automatically by SAS); we then ran the same model again, after having removed these random effects. The results reported hereafter always correspond to this second analysis (each report also specifies which random effects structure was used in the final analysis).

As mentioned previously, mental load was a within-participants IV: each dyad performed one session in the "with time pressure" condition and the other session in the "without time pressure" condition. Condition order was counterbalanced across dyads to prevent any order effects; however, this could have affected the results, as the participants were more aware of the temporal demands of the task in Session 2 than in Session 1. In order to discard this possibility, all the analyses reported below initially included an additional within-participants factor, session order. These analyses (which are not reported here) revealed that session had no significant influence on project marker production, nor did this factor interact with the other IVs in the design. Moreover, the patterns of results obtained were identical regardless of whether or not session was included in the analyses. This factor was consequently removed from all analyses, and the data from both sessions were analyzed together.

Finally, because the number of items (i.e., puzzles completed) differed across dyads, the numbers of observations varied across design cells. The Satterthwaite correction was therefore applied to correct the degrees of freedom used in the analyses (Keselman, Algina, Kowalchuk, & Wolfinger, 1999; Satterthwaite, 1946). Main effects were always included in the analyses, regardless of whether or not they were significant; interactions were only included when significant.

Manipulation check

The manipulation check data are reported in Appendix A. As expected, both Directors and Matchers experienced increased mental workload in the "with time pressure" condition, compared with the "without time pressure" condition, confirming that the mental load manipulation used in this experiment was effective.

Analysis #1: Effect of mental load and visibility on project marker production

The data used in this analysis are reported in Table 2. The model used to analyze the data included mental load, visibility, participant role, success, the Visibility × Mental Load interaction and the Participant × Visibility × Mental Load interaction as fixed effects. The outcome variable was the proportion of project markers produced. The random effects structure included by-participant random intercepts, by-dyad and by-item random slopes corresponding to participant role, and by-participant and by-item random slopes corresponding to success.

A significant main effect of participant role was found, $F\ (1, 15) = 76.29$, $p < .001$. As shown in Table 2, the Matchers' speech included a higher proportion of project markers than the Directors' speech. As shown in Figure 3, there was also a significant Visibility × Mental Load interaction, $F\ (1, 102) = 6.27$, $p = .014$, which was qualified by a significant Participant × Visibility × Mental Load interaction, $F\ (3,\ 109) = 3.75$, $p = .013$.

A simple main effects test was conducted to investigate these interactions further. This test revealed that the Visibility × Mental Load interaction was significant for Matchers, $F\ (3, 119) = 7.03$, $p < .001$, but that it was not significant for Directors, $F\ (3,\ 119) = 0.07$, $p = .978$. An inspection of the b coefficient revealed that in Matchers, the effect of mental load (which increased the proportion of project markers produced) was stronger in the "with visibility" condition than in the "without visibility" condition, $b = 0.12$, $p = .001$. All other effects included in the analysis failed to reach statistical significance, $p > .05$.

Table 2. *Proportion of project markers produced as a function of participant role, mental load, visibility, and success*

	With visibility			Without visibility			
Director	Without time pressure	With time pressure	**Total (with visibility)**	Without time pressure	With time pressure	**Total (without visibility)**	**Total**
Successful	0.07	0.05	**0.05**	0.06	0.05	**0.05**	**0.05**
Unsuccessful	0.04	0.04	**0.04**	0.05	0.04	**0.05**	**0.04**
Total	**0.05**	**0.04**	**0.04**	**0.05**	**0.04**	**0.05**	**0.05**

	With visibility			Without visibility			
Matcher	Without time pressure	With time pressure	**Total (with visibility)**	Without time pressure	With time pressure	**Total (without visibility)**	**Total**
Successful	0.14	0.26	**0.23**	0.22	0.18	**0.20**	**0.21**
Unsuccessful	0.17	0.30	**0.23**	0.19	0.18	**0.19**	**0.21**
Total	**0.16**	**0.28**	**0.23**	**0.20**	**0.18**	**0.19**	**0.21**

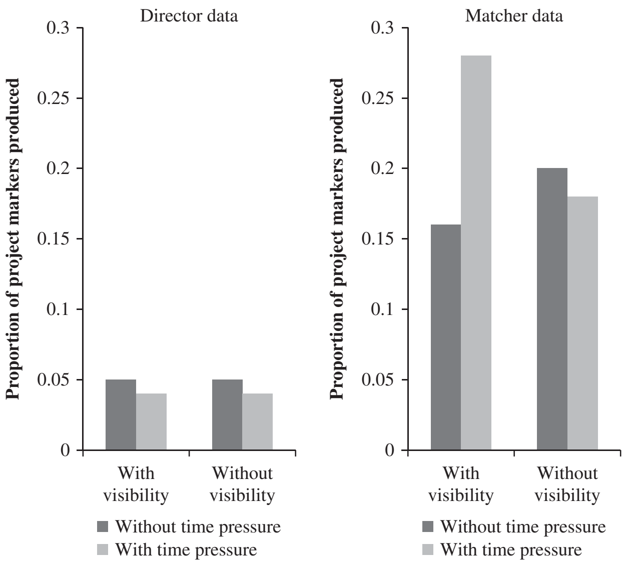

Figure 3. Proportion of project markers produced as a function of visibility and mental load by (a) Directors and (b) Matchers.

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press

Analysis #2: Effect of mental load and visibility on vertical project marker production

The data used in this analysis are reported in Table 3. The model used to analyze the data included mental load, visibility, participant role, success, and the Participant Role × Mental Load interaction as fixed effects. The outcome variable was the proportion of vertical project markers produced. The random effects structure included by-participant and by-item random intercepts, by-dyad and by-item random slopes corresponding to participant role, and by-participant and by-item random slopes corresponding to success. A significant main effect of participant role was found, F (1, 9) = 56.56, $p < .001$. As shown in Table 3, the Matchers' speech included a higher proportion of vertical project markers than the Directors' speech. A significant main effect of mental load was also found, F (1, 97) = 6.32, $p = .013$, which was qualified by a significant Participant × Mental Load interaction, F (1, 101) = 7.70, $p = .007$ (see Figure 4).

An inspection of the b coefficient revealed that the effect of mental load was stronger in Matchers than in Directors, $b = 0.04$, $p = .007$. All other effects included in the analysis failed to reach statistical significance, $p > .05$.

Analysis #3: Effect of mental load and visibility on horizontal project marker production

The data used in this analysis are reported in Table 4. The model used to analyze the data included mental load, visibility, participant role, success, and the

Table 3. *Proportion of vertical project markers produced as a function of participant role, mental load, visibility, and success*

	With visibility			Without visibility			
Director	Without time pressure	With time pressure	**Total (with visibility)**	Without time pressure	With time pressure	**Total (without visibility)**	**Total**
Successful	0.02	0.02	**0.02**	0.02	0.01	**0.02**	**0.02**
Unsuccessful	0.02	0.02	**0.02**	0.02	0.02	**0.02**	**0.02**
Total	**0.02**	**0.02**	**0.02**	**0.02**	**0.02**	**0.02**	**0.02**

	With visibility			Without visibility			
Matcher	Without time pressure	With time pressure	**Total (with visibility)**	Without time pressure	With time pressure	**Total (without visibility)**	**Total**
Successful	0.06	0.09	**0.08**	0.09	0.08	**0.09**	**0.08**
Unsuccessful	0.07	0.13	**0.10**	0.07	0.09	**0.08**	**0.09**
Total	**0.06**	**0.12**	**0.10**	**0.08**	**0.09**	**0.08**	**0.09**

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press

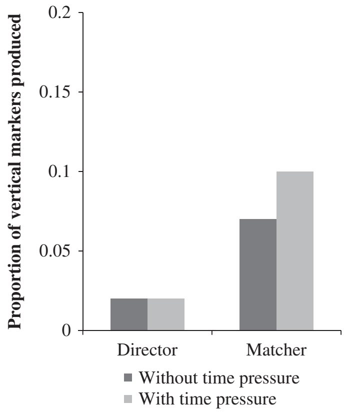

Figure 4. Proportion of vertical project markers produced as a function of participant role and mental load.

Visibility × Mental Load interaction. The outcome variable was the proportion of horizontal project markers produced. The random effects structure included by-participant and by-item random intercepts, by-dyad random slopes corresponding to participant role, and by-participant and by-item random slopes corresponding to success. A significant main effect of participant role was found, F (1, 63) = 41.28, $p < .001$. As shown in Table 4, the Matchers' speech included a higher proportion of horizontal project markers than the Directors' speech. As shown in Figure 5, there was also a significant Visibility × Mental Load interaction, F (1, 124) = 4.35, $p = .039$.

An inspection of the b coefficient revealed that the effect of mental load was weaker in the "without visibility" condition than in the "with visibility" condition, $b = -0.04$, $p = .039$. The effect of mental load was in opposite directions in both conditions: time pressure increased the production of horizontal project markers in the "with visibility" condition but decreased the production of horizontal project markers in the "without visibility" condition. All other effects included in the analysis failed to reach statistical significance, $p > .05$.

DISCUSSION

The purpose of this study was to examine the effect of mental load and face visibility on project marker production in joint activities (Bangerter & Clark, 2003; Bangerter et al., 2004). The manipulation check confirmed that the mental load manipulation was effective for both Directors and Matchers (Hart, 2006; Hart & Staveland, 1988).

The first two hypotheses tested in this study pertained to the effect of mental load on project marker production. Specifically, the first one was that dialogue partners produce *fewer* project markers when they experience increased mental load, and the second one was that dialogue partners produce *more* project markers

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press

Table 4. *Proportion of horizontal project markers produced as a function of participant role, mental load, visibility, and success*

	With visibility			Without visibility			
Director	Without time pressure	With time pressure	**Total (with visibility)**	Without time pressure	With time pressure	**Total (without visibility)**	**Total**
Successful	0.04	0.03	**0.04**	0.03	0.03	**0.03**	**0.03**
Unsuccessful	0.02	0.02	**0.02**	0.03	0.02	**0.03**	**0.03**
Total	**0.03**	**0.03**	**0.03**	**0.03**	**0.03**	**0.03**	**0.03**

	With visibility			Without visibility			
Matcher	Without time pressure	With time pressure	**Total (with visibility)**	Without time pressure	With time pressure	**Total (without visibility)**	**Total**
Successful	0.08	0.18	**0.15**	0.13	0.10	**0.11**	**0.13**
Unsuccessful	0.10	0.16	**0.13**	0.12	0.10	**0.11**	**0.12**
Total	0.10	0.17	**0.14**	0.12	0.10	**0.11**	**0.12**

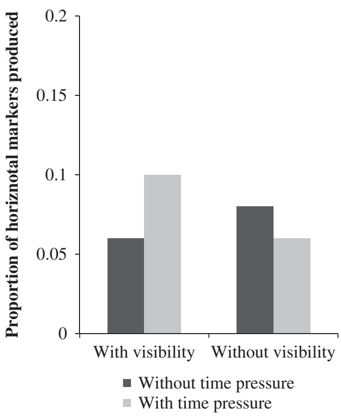

Figure 5. Proportion of horizontal project markers produced as a function of visibility and mental load.

when they experience increased mental load. The third hypothesis was that dialogue partners produce *fewer* project markers when they can see each other's face than when they cannot see each other's face. Interaction hypotheses were also tested. The corpus generated in an experiment involving spontaneous dialogues between pairs of participants enabled us to test all three hypotheses by examining the participants' production of *yeah*, *okay*, *uh-huh*, *right*, *got it*, and *alright*.

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press

Overall, the results reported here confirm and extend previous work by showing that mental load affects speech production (Belke, 2008; Berthold & Jameson, 1999; Hartsuiker & Barkhuysen, 2007; Howarth & Anderson, 2007; Villing, 2009; Wagner et al., 2010). These results are mainly in favor of our second hypothesis, as the participants' speech tended to include *more* markers in the "with time pressure" condition. However, the effect of mental load was not the same depending on whether the participants could see each other or not, and on the role each of them played in the dyad. Contrary to what was expected, the increase in project marker production was mainly observed when the participants could see each other. More specifically, when both kinds of project markers (vertical and horizontal) were analyzed together, the combined effects of mental load and visibility were mainly observed in Matchers. In the analysis that focused on horizontal project markers only, these effects were observed in participants regardless of their role in the dyad. Finally, in the analysis that focused on vertical project markers only, the effect of mental load did not depend on visibility, but it was stronger for Matchers than for Directors.

This pattern of results is in line with the idea that *in some situations* (in the current study, situations in which the participants could see each other), *some dialogue partners* (in the current study, Matchers) assume that they are more likely to fail to understand each other in situations where they experience a high mental load, leading them to navigate the interaction in a more explicit fashion. They do this in order to increase the chances of mutual comprehension even in situations where they have fewer mental resources to allocate to the interaction, highlighting the fundamentally collaborative nature of human dialogue (e.g., Clark, 1992, 1996; Clark & Wilkes-Gibbs, 1986; Isaacs & Clark, 1987; Schober, 1995). It is noteworthy that the influence of mental load on another dialogic collaborative behavior, namely, partner adaptation (e.g., Gorman et al., 2013; Isaacs & Clark, 1987; Vanlangendonck et al., 2016), is usually found to be in the opposite direction to the results reported here. Mental load is usually found to prevent dialogue partners from producing partner-adapted speech (Cane et al., in press; Horton & Keysar, 1996; Rossnagel, 2000). In this context, the current findings enrich previous research by highlighting that mental load does not necessarily undermine all aspects of collaboration during dialogue: on the contrary, mental load sometimes causes dialogue partners to behave *more* collaboratively.

This could contribute to our understanding of how dialogue partners manage to reach mutual comprehension even in situations where they have very few resources left to allocate to the interaction. In these situations, partner adaptation may be less optimal (e.g., dialogue partners might produce utterances that are less well adapted to the other person's dialogic needs, thus increasing the need for clarification and repair speech turns; Clark & Wilkes-Gibbs, 1986), but dialogue navigation would be more explicit and hence potentially more optimal, helping the participants to navigate projects and subprojects. Clark (1996) has suggested that the utterances produced during dialogue include both track 1 signals (which refer to the "official business" of the interaction) and track 2 signals (which are more concerned with interaction management), as highlighted above. Whereas

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press

the partners' attempts to produce partner-adapted speech mainly influence track 1 signals (e.g., they affect the partners' choices as to how to refer to an object or a person), project markers are track 2 signals. Thus, one possibility is that an increase in mental load causes dialogue partners to allocate fewer resources to track 1 processes in order to allocate more resources to the production of track 2 signals.

These findings also raise the question of *why* the effect of mental load would mainly be observed when the participants could see each other's face (and not when the participants could *not* see each other's face, as initially predicted). We have suggested above that producing a higher proportion of project markers while experiencing increased levels of mental load is a strategy used by dialogue partners to ensure mutual comprehension. One possibility is that when two dialogue partners can see each other's face, a number of visual cues become available to facilitate the interaction (e.g., head nods, facial expressions, gaze, etc.; see Clark & Krych, 2004; Lysander & Horton, 2012). Relying on these cues would increase the amount of mental resources available to the partners to allocate to other aspects of the interaction. For instance, they may switch to navigating the interaction in a more explicit fashion when experiencing high mental load. In other words, dialogue partners would only be able to adjust the production of project markers as a function of mental load in situations where they have enough mental resources to do so; the amount of resources available may depend on other features of the situation in which the interaction takes place, including whether or not the partners can see each other's face.

Although the results are consistent with the idea that project marker production can also be affected by face visibility, this factor only seemed to affect project marker production in an *indirect* way, through interactions with the mental load manipulation. This has two main theoretical implications. First, it confirms that simply seeing one's dialogue partner's face (rather than his or her progress in the task at hand; Bangerter & Clark, 2003; Clark & Krych, 2004) is sufficient to affect project marker production, in line with the more general idea that visibility affects information grounding (Clark & Brennan, 1991; Kraut et al., 2003; Whittaker, 2003). Second, the fact that face visibility affects project marker production only indirectly, whereas other studies have found that *workspace* visibility affects project marker production directly (Bangerter & Clark, 2003; Clark & Krych, 2004), implies that both kinds of cues are processed and used differently by dialogue partners to adjust project marker production as they interact.

In this sense, the current study provided no direct evidence in support of the third hypothesis, that is, that face visibility would result in a decrease in project marker production. This conclusion should nonetheless be nuanced in light of the fact that the participants' eye movements were not tracked in the current study, preventing us from checking that the participants actually did look at each other's face when they had the opportunity to do so. Thus, the lack of any direct effect of face visibility could be due to this cue having no strong influence on project marker production, or it could be due to this cue not being used much in the specific setting used in this study (e.g., this could be because the participants needed to look down at the puzzle most of the time). These two alternative possibilities will be examined in more detail in the future.

In addition to mental load and visibility, project marker production was also found to depend on the role played in the dyad: the Matchers' speech included a higher proportion of project markers (vertical and horizontal) than the Directors' speech. This is essentially due to the fact that the Matchers' task mainly consisted of acknowledging the instructions provided by the Directors (and potentially asking questions, in situations where the instructions provided were unclear); in this context, project markers were used by the Matchers to indicate that they were ready to move on to the next subproject and/or project. What is more, the results revealed that Matchers were more sensitive than Directors to the mental load manipulation used in this study. The first analysis (in which both kinds of project markers were taken into account) revealed that the effect of mental load on project marker production (when the participants could see each other) was mainly observed in Matchers. A similar pattern was found in the second analysis, which focused on vertical project markers only. Bangerter and Clark (2003) have pointed out that Directors and Matchers in matching tasks do not use vertical project markers for the same purposes. Directors use these markers to enter projects whereas Matchers use them to exit projects. In this context, the present results imply that increasing the mental load mainly caused project exits to be marked more explicitly. In contrast, the effect of mental load on horizontal project markers (once again when the participants could see each other) was observed in both Directors and Matchers, implying that both partners engaged in more explicit turn-by-turn grounding of subprojects when experiencing increased levels of mental load. One possible explanation for this pattern of results is that the production of vertical project markers by Matchers and the production of horizontal markers by Matchers and Directors are perceived by dialogue partners as stronger determinants of dialogue success than the production of vertical project markers by Directors, as a greater proportion of the latter is produced in situations where the partners believe that they are more likely to fail to understand each other (because of the high levels of mental load experienced).

Finally, although success in the task was taken into account in the current study, the analyses conducted revealed no reliable relationship between this variable and project marker production. This highlights the fact that although project markers contributes to dialogue success, by allowing dialogue partners to keep track of their progress in the task at hand (Bangerter & Clark, 2003), whether project marker production also influences the partners' performance on the task remains an open question.

Limitations of the study and directions for future research

The experiment presented a number of limitations that will be overcome in future research. First, face visibility could be manipulated in a different way. For instance, it is noteworthy that because the task required both participants to look at their own workspace as they interacted, they might not have looked at each other's face much (see Lysander & Horton, 2012, who reported a similar issue). Future research will involve tasks in which participants have more opportunities to look at each other's face; it will also involve manipulating both face visibility and workspace visibility in order to examine the (potentially joint) influence of both sources of information on project marker production (see Clark & Krych, 2004).

It is also noteworthy that intonation was not examined in the current study. During the coding process, we found that a number of project markers were produced using a rising intonation in this experiment. For instance, a participant might say "and now you have the triangle *yeah(?).*" In this situation, project markers were used to elicit a confirmation from one's partner. The number of confirmation-eliciting markers seemed to be quite low in this study, making it difficult to analyze this phenomenon separately, but future experiments will seek to investigate the production of such confirmation-eliciting markers further (i.e., which partner is more likely to produce them, whether both vertical *and* horizontal project markers can be used to elicit confirmation, and whether their production also increases when mental load increases and/or when the participants cannot see each other).

Another potential limitation is that, as highlighted above, all instances of *uh-huh* and *yeah* were counted as horizontal project markers and all instances of *okay*, *got it*, and *all right* were counted as vertical project markers in the current study, regardless of the context in which they occurred. Although this is in line with Bangerter and Clark's (2003) division of project markers, it is likely that at least in some situations, some project markers are used in a different way. A study in progress is currently being undertaken in which each project marker present in the corpus is coded for its function in the interaction. This work will offer more insight into how the manipulations used in the current experiment (time pressure and visibility) affect the way in which project markers are used by dialogue partners. For instance, increased time pressure could cause dialogue partners to pay more attention to the horizontal versus vertical division (e.g., to make sure that they use *okay* in vertical transitions and *uh-huh* in horizontal transitions) in order to improve communication. In contrast, visibility could cause them to pay less attention to this division, because they have access to other cues (such as their partner's gaze or facial expressions) to determine whether he or she is ready to move onto the next project or subproject.

Our findings also give rise to a number of new theoretical questions. From a linguistic point of view, we intend to further explore the specificity of the project markers studied here, and especially *yeah*, which was the most frequent project marker in the corpus (it represented almost 52% of all project markers produced). As mentioned above, *yeah* means that the partner producing it understood what was said previously, and more or less agrees with it; as such, it is processed as a horizontal project marker (Bangerter & Clark, 2003). For readability purposes, the different forms of this project marker were grouped in a single category in this study (*yeah*, *yes*, and *yep*). Yet, the morphophonological distinction between *yeah* and *yes*, as well as the range of their respective uses and frequencies (e.g., "*yes* seems to be specialized for answering questions affirmatively, and *yeah* has a number of uses, all of them corresponding to horizontal transitions"; Bangerter & Clark, 2003, p. 208), suggest that *yes* may nevertheless present vertical properties as well. For instance, a partner might say *yes* in order to mark his or her agreement with what was said previously and hence to move on to the following project. This assumption is in line with Schegloff (1982, p. 83), who makes a slight functional difference between *yes* (which signals "a full turn rather than a passing one") and *yeah* (which signals "continued attention"), even if both forms are considered as "turns continuers." A finer-grained qualitative analysis (e.g., Heiden, 2010) of the progression of the use of *yeah* and *yes* throughout the interaction will enable us to determine whether *yes* is usually processed in the same way as *yeah* or in the same way as vertical project markers.

In addition, the mental load manipulation was identical for both participants in this study; that is, both the Director and the Matcher were under high mental load or under no mental load. This made it difficult to determine whether the participants made the effort to produce project markers despite the mental load (i.e., Matcher-produced project markers and Director-produced horizontal project markers) because they experienced a high mental load themselves, or because they could see that their partner experienced a high mental load (in which case these project markers would have been produced to "help" one's partner), or potentially both. Future research will examine situations in which two dialogue partners experience different amounts of mental load. Doing so will extend the current results by examining not only *when* mental load affects project marker production but also *why* this is the case. The influence of changes in mental load throughout the interaction will also be investigated.

Conclusion

To summarize and conclude, the current study sought to offer a better understanding of joint activity navigation (Bangerter & Clark, 2003; Bangerter et al., 2004). The main finding was that mental load increases the production of project markers, depending on the role played in the dyad (for vertical project markers) or face visibility (for horizontal project markers). Slightly different patterns of results being found for vertical and horizontal project markers is in line with Bangerter and Clark's (2003) suggestion that both kinds of markers are used to do different things in the interaction. However, these findings also challenge the conclusions from studies on other collaborative behaviours, such as perspective taking (Horton & Keysar, 1996; Rossnagel, 2000), by suggesting that dialogue partners may become *more* collaborative when experiencing increased levels of mental load. These findings also shed light on how face visibility may affect dialogue in a different way than workspace visibility (e.g., Bangerter & Clark, 2003; Clark & Krych, 2004). In this sense, these findings build on and extend previous research by offering a better understanding of how the features of the communication situation may affect joint project navigation.

NOTES

1. Although Bangerter and Clark (2003) have shown that *okay*, *got it*, and *all right* are mainly used in vertical transitions whereas *uh-huh and yeah* are mainly used in

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press

horizontal transitions, different project markers may have different functions in some situations. The percentage of cases in which each marker was used either as vertical or horizontal in their analysis depended on a number of factors, including the role of the current speaker, the nature and the task, and the language used. For instance, in the US-Tangram corpus, *okay*, *all right*, and *got it* were used to acknowledge descriptions (a vertical transition) between 33.3% and 50.9% of the time, whereas it was used to answer questions (a horizontal transition) between 0.5% and 1.8% of the time only. In the current study, project markers were always coded with regard to their main function: *okay*, *got it*, and *alright* were always coded as vertical whereas *uh-huh* and *yeah* were always coded as horizontal. This point is discussed further in the Method and Discussion sections.

2. Specifically, the demands of speaking and listening increase the *dialogic* mental load, that is, the mental load associated with language processing per se during the interaction. However, interacting enables dialogue partners to coordinate on one or several joint activities (Bangerter & Clark, 2003), potentially making this activities easier to carry out, and thus reducing the *task* mental load experienced by the dialogue partners.

REFERENCES

Achim, A. M., Achim, A., & Fossard, M. (2017). Knowledge likely held by others affects speakers' choices of referential expressions at different stages of discourse. *Language, Cognition and Neuroscience*, *32*, 21–36. doi:10.1080/23273798.2016.1234059

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., …, Weinert, R. (1991). The HCRC map task corpus. *Language and Speech*, *34*, 351–366.

Bangerter, A., & Clark, H. H. (2003). Navigating joint projects with dialogue. *Cognitive Science*, *27*, 195–225. doi:10.1207/s15516709cog2702_3

Bangerter, A., Clark, H. H., & Katz, A. R. (2004). Navigating joint projects in telephone conversations. *Discourse Processes*, *37*, 1–23. doi:10.1207/s15326950dp3701_1

Bangerter, A., & Smolenski, C. (2000). *Die Wiederverwendung konzeptueller Pakte bei der Bezugnahme auf neue Informationen [Reusing conceptual pacts in referring to new information]* (No. 72). Psychological Institute, University of Basel.

Barnard, P. J. (1974). *A corpus of 757 directory enquiry conversations*. Cambridge, MA: MRC Applied Psychology Unit.

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

Belke, E. (2008). Effects of working memory load on lexical-semantic encoding in language production. *Psychonomic Bulletin & Review*, *15*, 357–363. doi:10.3758/PBR.15.2.357

Berthold, A., & Jameson, A. (1999). Interpreting symptoms of cognitive load in speech input. In J. Kay (Ed.), *User modeling: Proceedings of the Seventh International Conference, UM99*. Vienna: Springer.

Boiteau, T. W., Malone, P. S., Peters, S. A., & Almor, A. (2014). Interference between conversation and a concurrent visuomotor task. *Journal of Experimental Psychology: General*, *143*, 295–311. doi:10.1037/a0031858

Branigan, H. P., Catchpole, C. M., & Pickering, M. J. (2011). What makes dialogues easy to understand? *Language and Cognitive Processes*, *26*, 1667–1686. doi:10.1080/01690965.2010.524765

Cane, E. C., Ferguson, H. J., & Apperly, I. A. (in press). Using perspective to resolve reference: The impact of cognitive load and motivation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Clark, H. H. (1992). *Arenas of language use*. Chicago: University of Chicago Press.

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

Clark, H. H. (2005). Coordinating with each other in a material world. *Discourse Studies*, *7*, 507–525. doi:10.1177/1461445605054404

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

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

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

Fischer, K. (Ed.). (2006). *Approaches to discourse particles (Vol. 1)*. London: Elsevier.

Fox Tree, J. E. (2010). Discourse markers across speakers and settings. *Language and Linguistics Compass*, *3*, 1–13. doi:10.1111/j.1749-818×.2010.00195.x

Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, *31*, 931–952. doi:10.1016/S0378-2166(98)00101-5

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, *1*, 292–304. doi:10.1111/j.1756-8765.2009.01020.x

Gerjets, P., & Scheiter, K. (2003). Goal configurations and processing strategies as moderators between instructional design and cognitive load: Evidence from hypertext-based instruction. *Educational Psychologist*, *38*, 33–41. doi:10.1207/S15326985EP3801_5

Godfrey, J. J., Holliman, E. G., & McDaniel, J. (1992). SWITCHBOARD: Telephone speech corpus for research and development. In *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing*. San Francisco, CA: IEEE.

Gorman, K. S., Gegg-Harrison, W., Marsh, C. R., & Tanenhaus, M. K. (2013). What's learned together stays together: Speakers' choice of referring expression reflects shared experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 843–853. doi:10.1037/a0029467

Gross, D., Allen, J. F., & Traum, D. R. (1993). *The Trains 91 dialogues* (Trains Technical Note No. 92–1). Rochester, NY: University of Rochester, Computer Science Department.

Hancock, P. A., & Caird, J. K. (1993). Experimental evaluation of a model of mental workload. *Human Factors*, *35*, 413–429.

Hart, S. G. (2006). Nasa-task load index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*, 904–908. doi:10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam: North Holland Press.

Hartsuiker, R. J., & Barkhuysen, P. N. (2007). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes*, *21*, 181–204. doi:10.1080/01690960400002117

Heiden, S. (2010). The TXM platform: Building open-source textual analysis software compatible with the TEI encoding scheme. In R. Otoguro, K. Ishikawa, H. Umemoto, K. Yoshimoto, & Y. Harada (Eds.), *Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation* (pp. 389–398). Sendai, Japan: Institute for Digital Enhancement of Cognitive Development.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*, 91–117. doi:10.1016/0010-0277(96)81418-1

Howarth, B., & Anderson, A. H. (2007). Introducing objects in spoken dialogue: The influence of conversational setting and cognitive load on the articulation and use of referring expressions. *Language and Cognitive Processes*, *22*, 272–296. doi:10.1080/01690960600632796

Hupet, M., Seron, X., & Chantraine, Y. (1991). The effects of the codability and discriminability of the referents on the collaborative referring procedure. *British Journal of Psychology*, *82*, 449–462. doi:10.1111/j.2044-8295.1991.tb02412.x

Isaacs, E. A., & Clark, H. H. (1987). Reference in conversations between experts and novices. *Journal of Experimental Psychology: General*, *116*, 26–37. doi:10.1037/0096-3445.116.1.26

Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgment tokens "yeah" and "mm hm." *Papers in Linguistics*, *17*, 197–216.

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*, 2967–2999. doi:10.1080/03610929908832460

Kiernan, K., Tao, J., & Gibbs, P. (2012). *Tips and strategies for mixed modelling with SAS/STAT procedures*. Paper presented at the 2012 SAS Global Forum, Orlando, FL.

Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology*, *4*, 343–346. doi:10.1037/h0023705

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

Louwerse, M. W., & Mitchell, H. H. (2003). Toward a taxonomy of a set of discourse markers in dialog: A theoretical and computational linguistic account. *Discourse Processes*, *35*, 199–239. doi:10.1207/S15326950DP3503_1

Lysander, K., & Horton, W. S. (2012). Conversational grounding in younger and older adults: The effect of partner visibility and referent abstractness in task-oriented dialogue. *Discourse Processes*, *49*, 29–60. doi:10.1080/0163853×.2011.625547

Mills, G. J. (2014). Dialogue in joint activity: Complementarity, convergence and conventionalization. *New Ideas in Psychology*, *32*, 158–173. doi:10.1016/j.newideapsych.2013.03.006

Ordonez, L., & Benson, L. (1997). Decisions under time pressure: How time constraint affects risky decision making. *Organizational Behavior and Human Decision Processes*, *71*, 121–140.

Paas, F. G. W. C., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*, 63–71. doi:10.1207/S15326985EP3801_8

Rossnagel, C. (2000). Cognitive load and perspective-taking: Applying the automatic-controlled distinction to verbal communication. *European Journal of Social Psychology*, *30*, 429–445. doi:10.1002/(SICI)1099-0992(200005/06)30:3 < 429::AID-EJSP3 > 3.0.CO;2-V

Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology*, *53*, 61–86. doi:10.1111/j.1464-0597.2004.00161.x

Sato, N., Kamada, T., Miyake, S., Akatsu, J., Kumashiro, M., & Yume, K. (1999). Subjective mental workload in Type A women. *International Journal of Industrial Ergonomics*, *24*, 331–336. doi:10.1016/S0169-8141(98)00060-2

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics*, *2*, 110–114. doi:10.2307/3002019

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of "uh huh" and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk*. (pp. 71–93). Washington, DC: Georgetown University Press.

31w?>Schober, M. F. (1995). Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about locations? *Discourse Processes*, *20*, 219–247. doi:10.1080/01638539509544939

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*, 70–76. doi:10.1016/j.tics.2005.12.009

Svartvik, J., & Quirk, R. (1980). *A corpus of English conversation (Vol. 56)*. Lund: Liber Gleerups.

Tolins, J., & Fox Tree, J. E. (2014). Addressee backchannels steer narrative development. *Journal of Pragmatics*, *70*, 152–164. doi:10.1016/j.pragma.2014.06.006

Tylén, K., Weed, E., Wallentin, M., Roepstorff, A., & Frith, C. D. (2010). Language as a tool for interacting minds. *Mind and Language*, *25*, 3–29. doi:10.1111/j.1468-0017.2009.01379.x

Vanlangendonck, F., Willems, R. M., Menenti, L., & Hagoort, P. (2016). An early influence of common ground during speech planning. *Language, Cognition and Neuroscience*, *31*, 741–750. doi:10.1080/23273798.2016.1148747

Villing, J. (2009). Dialogue behaviour under high cognitive load. In *Proceedings of the SIGDIAL 2009 Conference: The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue* (pp. 322–325). London: Association for Computational Linguistics.

Vos, S. H., Gunter, T. C., Schriefers, H., & Friederici, A. D. (2010). Syntactic parsing and working memory: The effects of syntactic complexity, reading span, and concurrent load. *Language and Cognitive Processes*, *16*, 65–103. doi:10.1080/01690960042000085

Wagner, V., Jescheniak, J. D., & Schreifers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 423–440. doi:10.1037/a0018619

Whittaker, S. J. (2003). Theories and methods in mediated communication. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 243–286). Mahwah, NJ: Erlbaum.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*, 449–455. doi:10.1518/001872008 × 288394

Yoon, S. O., & Brown-Schmidt, S. (2014). Adjusting conceptual pacts in three-party conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 919–937. doi:10.1037/a0036161

APPENDIX A

RESULTS OF THE MANIPULATION CHECK

The data used in the manipulation check are reported in Table A.1. The model used to analyze the data included mental load, visibility, and participant role, as well as all interactions between these three factors, as fixed effects. The outcome variable was the participants' experienced mental workload, as measured by the NASA-TLX questionnaire. The random effects structure included by-dyad random intercepts, by-dyad random slopes corresponding to participant role and mental load, and by-participant random slopes corresponding to mental load.

A significant main effect of mental load was found, F (1, 28) = 16.54, $p < .001$. As shown in Table A.1, the participants experienced increased mental workload in the "with time pressure" condition than in the "without time pressure" condition. There was also a significant Participant Role × Mental Load interaction, F (1, 28) = 5.37, $p = .028$ (see Figure A.1). An inspection of the b coefficient revealed that the difference between the two mental load conditions was smaller for Directors than for Matchers, $b = -0.18$. Additional simple main effects tests revealed that the difference between the two mental load conditions was nonetheless significant for both Directors, F (1, 28) = 5.46, $p = .027$, and Matchers, F (1, 28) = 22.21, $p < .001$. All other effects in the analysis failed to reach statistical significance, all ps > .05.

These results confirm that the mental load manipulation used in this study was effective. The results also suggest that the manipulation was most effective for Matchers.

Table A.1. *Mental workload (as measured by the NASA-TLX Questionnaire) as a function of participant role, mental load, and visibility*

	With visibility			Without visibility		
	With time pressure	Without time pressure	Total	With time pressure	Without time pressure	Total
Director	13.05 (2.57)	12.81 (2.81)	**12.93 (2.65)**	13.42 (2.60)	11.46 (3.14)	**12.44 (3.00)**
Matcher	12.80 (3.09)	10.91 (2.70)	**11.86 (3.01)**	12.68 (2.16)	10.54 (2.81)	**11.61 (2.69)**
Total	**12.92 (2.80)**	**11.83 (2.87)**	**12.37 (2.87)**	**13.05 (2.38)**	**11.00 (2.97)**	**12.02 (2.86)**

Note: Standard deviations are reported in brackets.

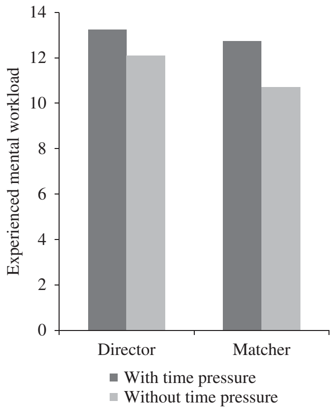

Figure A.1. Experienced mental workload as a function of participant role and mental load.

https://doi.org/10.1017/S0142716418000358 Published online by Cambridge University Press