Brain and Cognition 154 (2021) 105801

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

The neural correlates of referential communication: Taking advantage of sparse-sampling fMRI to study verbal communication with a real interaction partner

Amélie M. Achim [a,f,*], Isabelle Deschamps [b], Élisabeth Thibaudeau [c,f], Alexandra Loignon [c,f], Louis-Simon Rousseau [c,f], Marion Fossard [d], Pascale Tremblay [e,f]

[a] *Département de Psychiatrie et Neurosciences, Université Laval, Pavillon Ferdinand-Vandry, (room 4873), 1050, Avenue de la Médecine, Quebec City, QC G1V 0A6, Canada*
[b] *Human Services, Georgian College, 825 Memorial Avenue, Orillia, ON L3V 6S2, Canada*
[c] *École de Psychologie, Université Laval, Pavillon Félix-Antoine-Savard (room 1116), 2325, rue des Bibliothèques, Université Laval, Quebec City, QC G1V 0A6, Canada*
[d] *Institut des sciences logopédiques, Université de Neuchâtel, Rue Pierre-à-Mazel 7, CH-2000 Neuchâtel, Switzerland*
[e] *Département de réadaptation, Université Laval, Pavillon Ferdinand-Vandry, (room 4247), 1050, Avenue de la Médecine, Quebec City, QC G1V 0A6, Canada*
[f] *Centre de recherche CERVO / CERVO brain research Centre, 2601, de la Canardière, Québec, QC G1J 2G3, Canada*

ARTICLE INFO

Keywords:
Referential communication
Collaboration
Theory of mind
Mentalizing
Social cognition
Sparse sampling

ABSTRACT

This paper introduces an innovative functional magnetic resonance imaging (fMRI) protocol to study real verbal interactions while limiting the impact of speech-related movement artefacts. This protocol is based on a sparse sampling acquisition technique and allowed participants to complete a referential communication task with a real interaction partner. During verbal interactions, speakers adjust their verbal productions depending on their interlocutors' knowledge of the referents being mentioned. These adjustments have been linked to theory of mind (ToM), the ability to infer other's mental states. We thus sought to determine if the brain regions supporting ToM would also be activated during a referential communication task in which participants have to present movie characters that vary in their likelihood of being known by their interlocutor. This pilot study establishes that the sparse sampling strategy is a viable option to study the neural correlates of referential communication while minimizing movement artefacts. In addition, the brain regions supporting ToM were recruited during the task, though specifically for the conditions where participants could adjust their verbal productions to the interlocutor's likely knowledge of the referent. This study therefore demonstrates the feasibility and relevance of a sparse-sampling approach to study verbal interactions with fMRI, including referential communication.

1. Introduction

There is convincing evidence that during verbal interactions, speakers adjust the words that they use as a function of the knowledge held by the person they are interacting with (i.e. their interlocutor). For instance, a seminal study by Isaacs and Clark (1987) revealed that speakers more often use the names of New York monuments when they have to present these monuments to someone who is from New York, compared to when they have to present the same monuments to someone who is not from New York. That study as well as subsequent referential communication studies (Achim, Achim, & Fossard, 2017; Achim, Fossard, Couture, & Achim, 2015; Heller, Gorman, & Tanenhaus, 2012; Isaacs & Clark, 1987; Wu & Keysar, 2007) showed that speakers sometimes use names when presenting referents (e.g. places, objects or characters) that their interlocutor did not previously know, but then typically add descriptors along with the names, which allows the interlocutor to identify the target referent (Achim et al., 2015; Heller et al., 2012). It thus seems that even if speakers can be affected by egocentric biases linked to their own perspective or own knowledge (e.g. using privilege information that the interlocutor does not possess; Wu & Keysar, 2007), they nonetheless also adjust their verbal productions depending on their interlocutor's knowledge of the referents.

In a previous study (Achim et al., 2015), we sought to determine whether speakers would also adjust the way they present movie characters when their interlocutor's knowledge of these characters has not been previously established and hence has to be estimated in real-time.

---

* Corresponding author at: Centre de recherche CERVO, 2601, de la Canardière (F-4561-4), Québec, QC G1J 2G3, Canada.
E-mail address: amelie.achim@fmed.ulaval.ca (A.M. Achim).

https://doi.org/10.1016/j.bandc.2021.105801
Received 9 April 2021; Received in revised form 11 August 2021; Accepted 17 September 2021
Available online 9 October 2021

The results revealed that participants indeed adjusted their choices of referring expressions, adding descriptors along with the names more often when presenting referents (here movie characters) that the other person was less likely to know (e.g. “it is Leonidas from the movie 300, he has a beard and a red cape”) than when presenting very well-known characters that the other person was hence very likely to know (e.g. “it’s Harry Potter”). In addition, we observed that speakers adjusted the amount of information that they provided, with less information provided when the other person was very likely to know the character and more information provided when the other person was less likely to know the character (in the example above, the name “Leonidas”, the movie title “300”, and two pieces of descriptive information, “beard” and “red cape”). Interestingly, these adjustments were significantly correlated with participants’ performance on a typical, story-based theory of mind (ToM) task (the Combined stories task; Achim, Ouellet, Roy, & Jackson, 2012; Achim & Thibaudeau, 2018; Thibaudeau, Legendre, Villeneuve, Cellard, & Achim, 2018). ToM refers to the ability to infer the mental states of others, and in that study (Achim et al., 2015) the participants who showed better ToM abilities also adjusted their verbal productions to a greater extent during the social interactions. ToM thus seems implicated in assessing the interlocutor’s likely knowledge for new referents presented by the speakers during verbal interactions, at least when presenting referents for which prior knowledge by the interlocutor has never previously been established.

At the neurobiological level, ToM judgments involve a relatively consistent set of brain regions including the medial prefrontal cortex, the precuneus, the temporo-parietal junction and the superior temporal sulcus (Lavoie, Vistoli, Sutliff, Jackson, & Achim, 2016; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014; Spreng, Mar, & Kim, 2009). While it could be expected that the brain regions supporting ToM would be recruited when a speaker interacts verbally with another person whose likely knowledge is relevant to adjust communicative choices, technical challenges have limited the number of studies addressing the neural correlates of verbal interactions or referential communication. First, brain imaging techniques are sensitive to movement artefacts, including those resulting from speech production. Second, during MRI protocols, participants are lying on their back in a small noisy space, which makes verbal interactions challenging. Because of these challenges, very few fMRI studies have examined live verbal interactions between two people. Finding ways to study the neural correlates of language production during real verbal interactions is thus an important challenge to overcome in order to study the communicative aspects of language production. While we here focus on adjustments linked to the interlocutor’s likely knowledge and are interested in the eventual involvement of the brain regions also linked to ToM, finding ways to study language production in the context of an interaction with a real interlocutor is important more globally in order to favour ecological validity of language production studies, i.e. ensure that “the methods, materials, and setting […] approximate the real-world that is being examined” (Schilbach, 2015, p.159).

An interesting approach used in at least three previous fMRI studies of real verbal interactions was to examine brain activation during the phase in which speakers prepare their verbal utterances (Kuhlen, Bogler, Brennan, & Haynes, 2017; Willems et al., 2010) in order to limit the impact of artefacts linked to speech production. The study by Willems et al. (2010) revealed greater activation in the dorsomedial prefrontal cortex (dmPFC) in a condition in which the interlocutor needed to identify a target word from the verbal instructions provided by the speaker compared to when the interlocutor already knew the target word. The study by Kuhlen et al. (2017) asked participants to give instructions to an interaction partner through a live audiovisual stream or to speak outside of a conversational context (allegedly to calibrate the microphone). Using multivariate pattern analyses, the authors identified three brain regions that showed different patterns of activation between their two conditions, namely the ventromedial prefrontal cortex (vmPFC) and the left and right ventrolateral prefrontal cortex. Finally, the study by Vanlangendonck et al. (2018) used a referential communication task in which participants in the scanner described a target object either to a listener outside of the scanner who needed to identify the target object in a grid, or to themselves. Activation again emerged in the medial prefrontal cortex when contrasting the trials that involved (communicative trials) versus the trials that did not involve the interaction partner (non-communicative-trials). Overall, these pioneer studies all suggested an implication of the medial prefrontal cortex, a brain region often linked to ToM, when participants prepare what they will say to an interaction partner with whom they are performing a collaborative task (versus non-interactive conditions).

Additionally, the study by Vanlangendonck et al. (2018) also contrasted different types of referential communication trials that the speakers performed with the interaction partner. In all trials, some distracter objects were visible to both participants (common ground) while other distracter objects were visible only by the speaker (privileged ground). The authors manipulated whether speakers had to consider the objects in priviledge ground to communicate efficiently with the listener, and they observed greater activation in the dmPFC and bilateral temporoperietal junctions for the trials in which priviledge ground information was relevant to adjust communication. These brain regions partially overlaped with those activated by a ToM localizer task, and the authors suggested that the ToM brain network “plays a crucial role when speakers have to consider which information they share with their addressee” (Vanlangendonck et al., 2018).

In addition to these studies targeting the preparation phase, at least two studies looked at brain activation during a task in which pairs of participants communicated verbally (Jasmin et al., 2019; Spiegelhalder et al., 2014). In Spiegelhalder et al. (2014), the task did not allow for an interaction between the two partners, who in turn spoke about personal events (using a noise cancelling microphone system) or listened passively while the other participant spoke. The analyses targeting the period when the participants were speaking (versus baseline) revealed activation in brain areas previously linked to language production, including the primary motor, premotor, supplementary motor and cerebellar areas. No significant activation was found in brain areas linked to ToM, likely reflecting the non-interactive nature of the task. In Jasmin et al. (Jasmin et al., 2019), participants engaged in informal conversations with the experimenter during the full fMRI run and, given “the lack of appropriately spaced baseline periods during naturalistic conversation” (p. 811), the analyses focused on the pattern of functional connectivity. The study included healthy participants and people with autism, and increased interregional correlations were observed in people with autism relative to the healthy controls.

Other previous fMRI studies have allowed their participants to communicate through other, non-verbal means such as eye gaze, hand movements or facial expressions, either directly or through a video camera (e.g. Bilek et al., 2015; Cavallo et al., 2015; Guionnet et al., 2012; Koike et al., 2016; Redcay et al., 2013; Redcay et al., 2010; Stolk et al., 2014). Some of these studies involved the realization of a collaborative task with an interaction partner (Bilek et al., 2015; Redcay et al., 2013; Stolk et al., 2014). For example, in the study by Bilek et al. (2015), participants used eye gaze to inform the other person about the location on the screen of a target that the other person could not see. These studies mainly used hyperscanning (i.e. scanning two participants concurrently) and targeted interindividual synchronization of brain activation, rather than reporting the patterns of activation linked to the task demands per se. Overall, these studies paved the way for the study of real social interactions using fMRI, yet verbal interactions with a real interaction partner have been limited by the challenge of controlling for movement artefacts on the fMRI images and by the noisy environment.

The current study aimed to introduce the use of sparse sampling fMRI to study brain activation during a referential communication task. Sparse sampling fMRI acquisitions include silent delays between volume acquisitions, thereby allowing overt vocal responses without motion-induced artefacts (Birn, Bandettini, Cox, Jesmanowicz, & Shaker, 1998; Gracco, Tremblay, & Pike, 2005; Hall et al., 1999a, Hall et al., 1999b). This strategy had not yet been employed to study real-time verbal interactions.

In our referential communication task performed during fMRI acquisition, the speaker (in the scanner) presented a series of movie characters to his/her interlocutor (participating to the task in the control room nearby). While we expected to observe brain activation linked to speech production (e.g. motor areas, left inferior frontal gyrus, etc.) and to the perception of the visual stimuli (visual areas), we were particularly interested to establish whether the brain regions linked to ToM judgments (i.e. the medial prefrontal cortex, precuneus, temporo-parietal junction and/or superior temporal sulcus) would also be recruited during our referential communication task. More specifically, we aimed to determine if activation in these regions would emerge either across all referential communication trials, or would come into play more specifically for trials in which the interlocutor's knowledge of the movie characters, or lack thereof, had to be estimated in real-time.

2. Method

2.1. Participants

Fourteen healthy participants took part in this pilot study (9 men; mean age = 24.5 years, SD = 4.0, range = 20–34; mean of 17.5 years of education, SD = 3.6; range = 11–24). All were right-handed (average laterality quotient = 79.9, SD = 14.9 (Oldfield, 1971), had French as their first language, had normal or corrected-to-normal vision and no self-reported history of speech, voice, language or neurological disorders. The study was approved by the Research Ethics Board of the Centre intégré universitaire de santé et de services sociaux de la Capitale-Nationale (neuroscience and mental health division; project #339–2013) and all participants provided informed consent before being enrolled in the study.

2.2. Overview of the procedure

The experiment entailed two visits. During the first visit, participants completed a series of cognitive tasks and questionnaires (not used for the current study). In preparation for the experimental task, they were presented with a large set of image stimuli and were asked to report whether they knew the movie/TV characters presented in each image. For the characters that they knew, they were also asked to identify the character. As detailed below, this was done to create individualized stimuli sets for each participant.

The second visit consisted in the fMRI session during which participant completed a referential communication task with a confederate (E. T.), who was presented as being a naive participant. A practice version of the task was first performed with the confederate outside of the scanner, followed by the scanning session.

2.3. Stimuli used for the experimental task

The stimuli consisted in portrait pictures of characters from movies and TV shows. A series of online surveys were conducted to retain only the stimuli that presented either a character that most women in their twenties in Quebec City typically know (the 'typically known' category) or a character that women in their twenties in Quebec City would not typically know (the 'typically unknown' category). These surveys targeted women in their twenties given that the confederate with whom participants performed the referential communication task was a woman in her twenties. Further details about these surveys are presented as Supplementary material.

For the referential communication task performed during fMRI scanning, the following categories of stimuli were included:

(1) Characters that the participant knew and that are typically known according to our surveys (known-known condition, KK);
(2) Characters that the participant knew and that are not typically known according to our surveys (known-unknown condition, KU);
(3) Characters that the participant did not know and that are not typically known according to our surveys (unknown-unknown condition, UU).

The characters that the participant did not know and that are typically known according to our surveys were not retained for the task for several reasons. First, participants knew most of the typically known characters. Second, the adjustments in verbal productions identified in our previous studies (Achim et al., 2017; Achim et al., 2015) focused on the characters that the participants knew, given that if you do not know a character you can only provide descriptive information. Third, if someone does not know a character, he/she is not in a position to differentiate whether most people would know that character or not, and hence a single category of characters that the participants themselves did not know seemed sufficient, also allowing a greater number of trials for our conditions of interest (KK, KU and UU).

For each of our three conditions of interest, we aimed to retain 25 characters. In order to reach this number for most participants, our stimulus database included 35 typically known characters and 180 typically unknown characters (including male and female characters). The greater number of stimuli for the typically unknown characters reflects the fact that the participants themselves knew a smaller proportion of these characters and as a result we needed more stimuli to reach 25 characters to include in the KU category. For four participants, some trials nonetheless had to be replaced by fixation crosses since they did not know 25 characters either for the KK (N = 2 with 22 or 21 trials) or the KU condition (N = 2 with 23 or 20 trials).

2.4. Referential communication task

During the fMRI task, the participant communicated verbally with the interlocutor (a confederate presented as being another participant) using a high-quality MRI compatible optical omnidirectional microphone (MO-2000, Sennheiser). As shown in Fig. 1A, the same computer was used to project the stimuli on the screen that the participant saw in the scanner and to concurrently display three images on the screen that the interlocutor saw while seated in the control room outside the magnet room. For each trial, participants were asked to present the character verbally to the interlocutor so that she could identify the target character within her set of three images (the target character and two distracters). The three images appearing on the screen of the interlocutor were randomly selected, with the only constraint that the three images were either all of likely known or all of likely unknown characters. The interlocutor listened to the participant through earphones and selected the target character that she could identify by pressing the corresponding response key on her keyboard. Participants were aware that the interlocutor had to select the target character among three images, but they were not informed that there were different conditions to the task.

The interlocutor was a 24-year-old woman who knew all the typically known characters and was trained to act as if she knew none of the typically unknown characters. More specifically, for the typically unknown characters (identified on the interlocutor's screen by red numbers above each of the three images) she was trained to use only the descriptive information provided by the participant as cues to identify the target character, and to disregard any information that was not visually presented in the image, such as the name of the character, the name of the movie or TV show, or any other information linked to the movie or TV show. This manipulation ensured that all participants interacted with an interlocutor who displayed a typical level of knowledge about movie characters. Participants were not aware of which

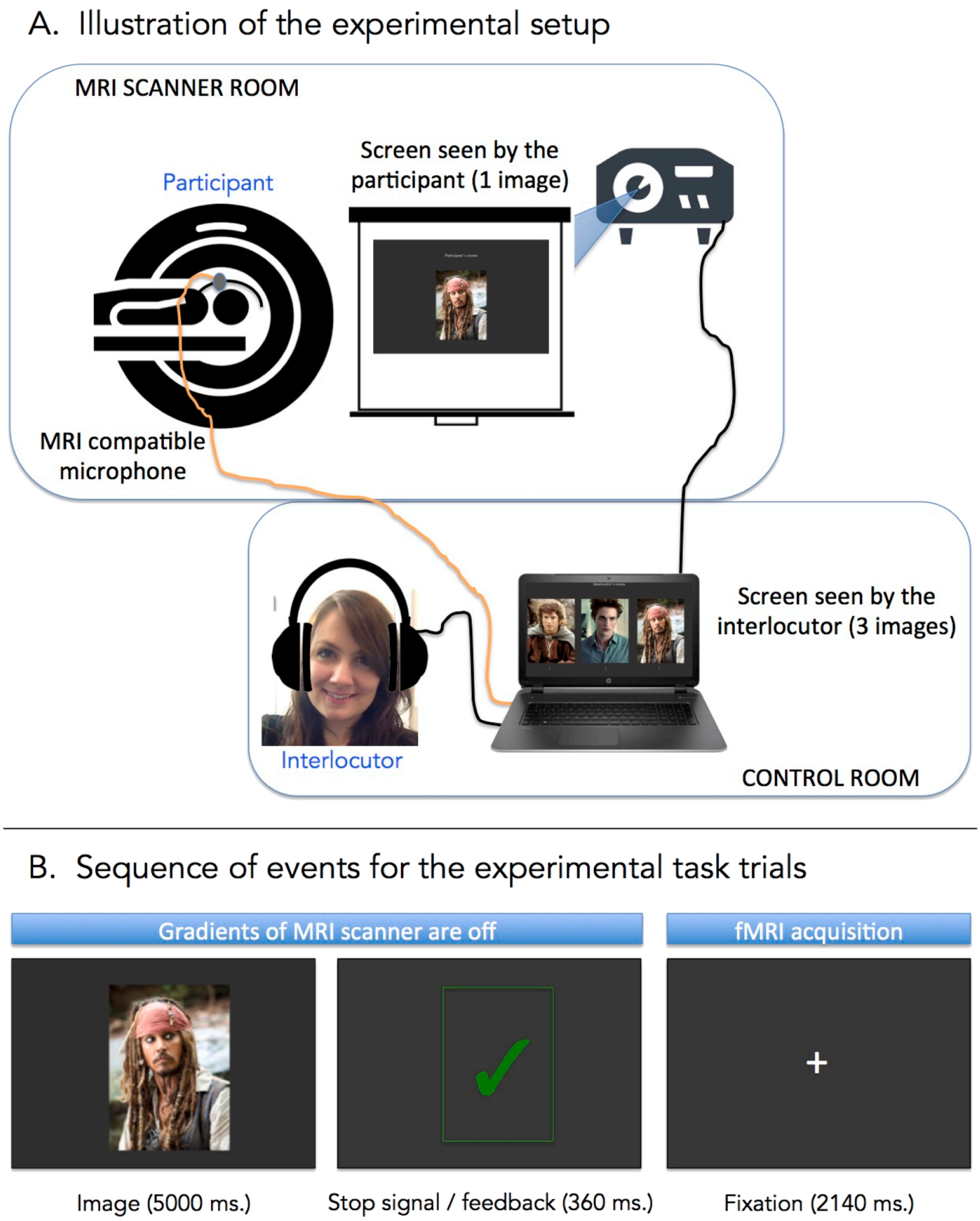

Fig. 1. Illustration of the experimental setup (A) and illustration of the sequence of events presented one after the other during each task trial.

characters the interlocutor knew. Thus, to estimate the likelihood that the interlocutor would know each new character, they had to rely on: (1) their own judgments about how well-known a character was and/or (2) their own estimation of the level of knowledge about movie/TV characters held by the interlocutor, as suggested by the feedbacks from previous trials.

As presented in Fig. 1B, each trial began with the presentation of the target stimulus on the participant's screen for 5000 ms. The scanner gradients were turned off during the presentation of the stimulus (i.e. sparse sampling fMRI acquisition), allowing the participant to present the character verbally without interference from the noise and without creating movement artifacts in the fMRI images. Then, a stop signal was presented for 360 ms, which indicated to the participant to stop talking before the acquisition of the next volume began. This stop signal also provided feedback on whether the interlocutor had identified the target character among the three images displayed on her screen. Next, a fixation cross was presented for 2140 ms while an fMRI volume was acquired. These time intervals were determined such that the fMRI volumes be acquired at the time when the hemodynamic response is expected to peak (i.e. about 5–6 s after the beginning of the trials), also allowing sufficient time for the participants to verbally present the characters. Twenty-five (25) fixation trials were also included in the design. For these trials, a fixation cross was presented for 5360 ms with the gradients turned off and remained on the screen for an additional 2140 ms during which the volume was acquired. Experimental and control trials were presented in a pseudo-random order, and the design was optimised using Optseq2 (http://surfer.nmr.mgh.harvard.edu/optseq/). The task was implemented in the software Presentation 16.5, which was also used to record the interlocutor's response choices. The same computer also recorded the verbal productions of the participants using Audacity 2.0.5.

2.5. Image acquisition

High-resolution anatomical and functional data were acquired with a 3 T Philips Achieva TX MRI scanner at the Clinic IRM Québec-Mailloux in Quebec City. The structural scans were acquired with a T1-weighted MPRAGE sequence (TR/TE = 8.2/3.7 ms, isotropic voxel size = 1 mm^3, 256 × 256 matrix, 180 slices/volume, no gap). Single-shot EPI images were acquired using a SENSE factor of 2 to reduce the number of phase encoding steps and a sparse sampling acquisition paradigm wherein each volume was followed by a period during which the gradients were turned off ("silent period"; Eden, Joseph, Brown, Brown, & Zeffiro, 1999; Edmister, Talavage, Ledden, & Weisskoff, 1999; Gracco et al., 2005; Hall et al., 1999a, Hall et al., 1999b). A total of 113 functional images were acquired in a single run (TR/TE = 7500/30 ms, 40 axial slices, voxel size = 3 mm^3, no gap; matrixSD = 80 × 80, FoV = 240 × 240 mm; 2140 ms of scan time followed by 5360 ms of silence). Four dummy scans were used at the beginning of the sequence to allow the MRI signal to stabilize.

2.6. Behavioural data processing and analyses

The verbal productions were first transcribed verbatim. Then, for each character, each element of information provided in the verbatim was separately coded by two independent research assistants using the procedure developed by Achim et al. (2015). For example, if a participant said "it is Leonidas from the movie 300, he has a red cape and a black beard", this referring expression[1] would be coded as including two pieces of *Character-related information* (i.e. the name of the character "Leonidas" and the movie title "300") and two elements of *Descriptive information* ("red cape" and "black beard"). Additional elements in the verbal productions were categorized as *Other,* which included incomplete statements, mostly observed at the end of the trials and linked to the appearance of the stop signal (e.g. "he has a…"), false information (e.g. naming the wrong character), or information about the actors that participants were asked not to mention so that they focus on presenting the characters.

ANOVAs were used to compare the three task conditions (KK, KU, UU) for:

1) The total number of elements of information per trial;
2) The number of elements of information separately for each category (Character-related information, Descriptive information and Other);
3) The percentage of trials in which the characters were presented using only character-related information, only descriptive information or both character-related and descriptive information;
4) The percentage of characters correctly identified by the interlocutor.

T-tests were then used to decompose the significant effects for each pairs of conditions.

2.7. fMRI analysis

2.7.1. Pre-processing

The images were first visually inspected and no problem was detected at this stage, leading to the inclusion of all participants in the subsequent analyses. The time series were then spatially registered, motion-corrected, de-spiked, mean normalized and smoothed with a Gaussian 6 mm FWHM filter using AFNI (Cox, 1996). For motion correction, all time points occurring during excessive motion (i.e. >1 mm/degree) were excluded (Johnstone et al., 2006) from the regression using AFNI's *censor* function, which led to the exclusion of 6.66% of the volumes.

2.7.2. Individual-level analysis

Three separate regressors were created, one for each experimental condition (KK, KU, UU). Additional regressors included the mean, linear and quadratic trend components as well as the six motion parameters (x, y, z and roll, pitch and yaw). A 1-parameter block basis response function (fixed-shape regression; AFNI model BLOCK of length 5 sec, which corresponded to picture duration) was used to fit our statistical model and BOLD signal. The anatomical and functional datasets were spatially normalized to the Talairach version of the Colin_N27 template (TT_N27) using the 12-parameter affine transform implemented in AFNI (@auto_tlrc program). The T1 image was first normalized to the template, and then the T2 images were normalized to the normalized T1 images. The group analyses were performed on the participants' beta values resulting from the first level analysis.

2.7.3. Design-specific hypotheses and group-level analyses

To test whether the brain regions supporting ToM were activated during referential communication trials in general, the first analysis consisted in the conjunction of the three experimental conditions, which highlights the brain regions showing significant activation or significant deactivation relative to baseline across all conditions.

To determine if the brain regions supporting ToM are more specifically active when the other person's knowledge or lack thereof needs to be taken into consideration, we used a whole-brain one-way repeated measures ANOVA with conditions as the within-subject factor (KK, KU, UU), followed by *t*-tests comparing each pairs of conditions.

For all group analyses, a cluster correction for multiple comparisons was implemented using AFNI's 3dClustSim. This procedure revealed that a family-wise error (FWE) rate of $p < 0.05$ is achieved with a minimum cluster size of 112 contiguous voxels each significant at $p < 0.01$.

3. Results

3.1. Behavioural results

The behavioural results and statistics are presented in Table 1. Despite the time limitation (5000 ms. to present each character), the number of information used to present each character was closely matched to that observed in our prior behavioural study in which there was no time limit (Achim et al., 2015). More specifically, participants used a mean of 1.79 pieces of information for the KK condition in this study versus 1.83 for the likely-known characters in our previous study, and 2.13 and 2.27 pieces of information for the KU and UU conditions in this study versus 2.29 for the likely-unknown characters in our previous study. There were some instances where participants interrupted what they were about to say because of the stop signal, but it was not very frequent (2.9 instances per subject on average across all three conditions, with a range of 0–8 instances for the different subjects).

Participants used significantly more elements of information in their referring expressions when presenting characters that are not typically known (KU and UU) than when presenting well-known characters (KK). When they did not know the characters (UU), participants used more elements of descriptive information and fewer elements of character-related information (i.e. almost none) than for the two conditions for which they knew the characters (KK and KU). When the participants knew the characters, the amount of character-related information was greater for characters that the interlocutor was likely to know (KK) than not likely to know (KU). The use of descriptive information showed the reverse pattern, with more descriptive information for KU than KK.

The same analyses were also repeated looking at the number of words used to present the characters, instead of the number of pieces of information, and these results are provided in the Supplement, Table S1.

---

[1] Expression used by a speaker to identify a target item or person (here a movie character).

Table 1
Behavioural results.

	Mean (SD)			One-way ANOVA		*t*-tests comparing the pairs of conditions		
	KK	KU	UU	F	*p*	KK vs. KU	KK vs. UU	KU vs. UU
A. Number of pieces of information used per character*								
Total	1.83 (0.43)	2.13 (0.44)	2.27 (0.49)	8.22	0.002	$p = .002$	$p = .008$	p = .200
By type:								
Character-related	1.28 (0.22)	1.01 (0.32)	0.01 (0.15)	192.66	<0.001	p = 0.001	$p < .001$	p < .001
Descriptive	0.43 (0.34)	0.97 (0.53)	2.06 (0.45)	116.84	<0.001	$p < .001$	$p < .001$	p < .001
Other	0.12 (0.19)	0.14 (0.21)	0.20 (0.33)	1.50	0.243	–	–	–
B. Percentage of trials including the different types of information*								
Trials including character-related AND descriptive information	31.5% (23.4)	42.6% (25.1)	0.6% (1.7)	25.9	<0.001	$p = .030$	$p < .001$	p < .001
Trials with only character-related information	65.8% (25.2)	37.8% (31.2)	0.0% (0.0)	47.6	<0.001	$p < .001$	$p < .001$	p = .001
Trials with only descriptive information	2.7% (5.5)	19.7% (18.8)	99.3% (1.7)	351.0	<0.001	$p = .002$	$p < .001$	p < .001
C. Identification of the target character by the interlocutor								
Recognition rate	98.8% (2.7)	47.8% (20.1)	84.5% (14.9)	64.5	<0.001	$p < .001$	$p = .003$	$p < .001$

*See Table S2 in the Supplement for the same analyses performed with non-parametric tests.

As presented in Table 1B, there were more trials in which character-related information was accompanied by descriptive information for the KU condition than for the KK condition, where character-related information was more often used alone. The UU condition also significantly differed from the other two conditions since descriptive information was almost exclusively used for these trials (see Table 1B).

As presented in Table 1C, the interlocutor identified a significantly greater proportion of characters in the KK condition relative to both the KU and UU conditions, and in the UU condition relative to the KU condition. The incorrect trials were mainly misses (trials for which no response was provided within the allocated time), with very few incorrect answers (<2% of trials).

3.2. fMRI results: similarities between the conditions (conjunction)

As shown in Fig. 2A, the conjunction analysis revealed one large cluster of positive activations, with 14,067 voxels activated in all three conditions. This cluster included the bilateral lingual/occipital cortex, bilateral precentral and postcentral gyrus, left middle and inferior frontal gyrus, bilateral superior temporal gyrus, thalamus, striatum and cerebellum. The conjunction also identified a cluster of negative activation in the anterior cingulate (239 voxels).

3.3. fMRI results: differences between the conditions (ANOVA and t-tests)

As shown in Fig. 2B and listed in Table 2A, the ANOVA revealed

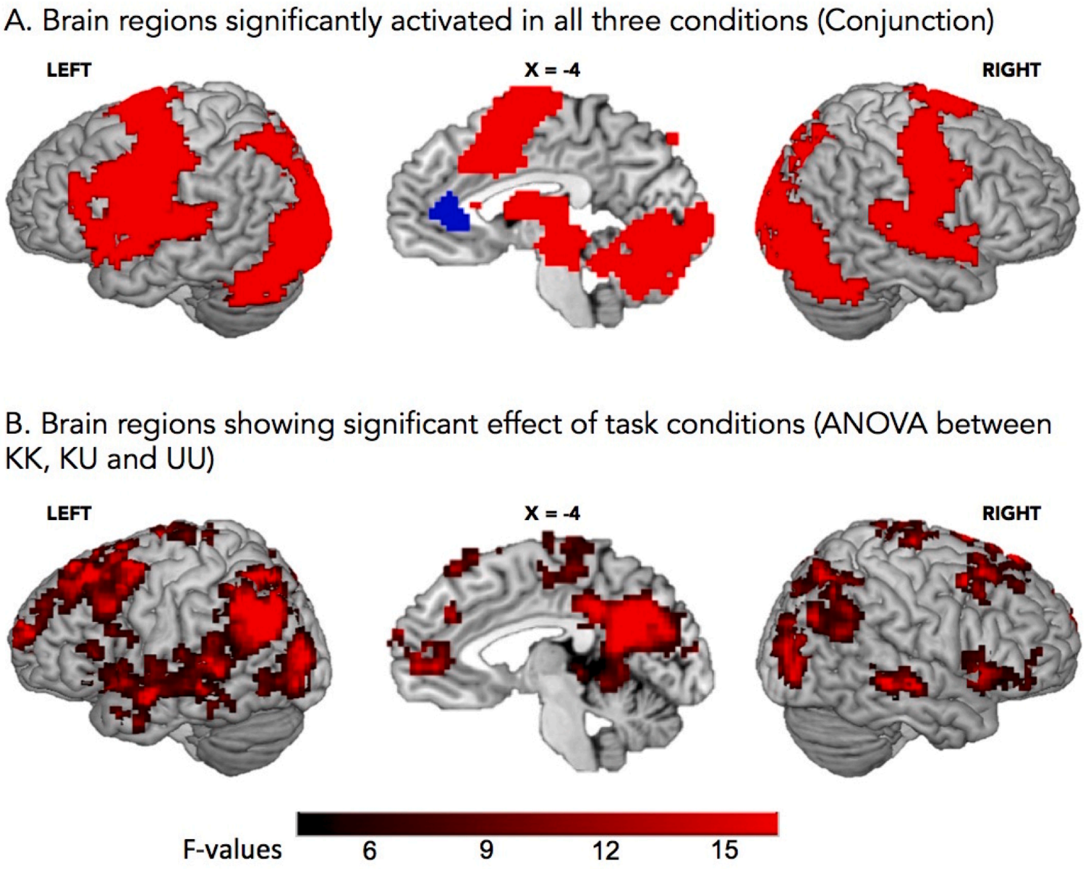

Fig. 2. Results from the conjunction analysis (A) and from the overall comparison between the three task conditions (B).

Table 2
fMRI results.

Brain regions	Hemisphere	x	y	z	F	Nb. of voxels
A. ANOVA comparing KK, KU and UU						
Precuneus, cuneus and posterior cingulate	Left	−4	−52	17	59.5	1679
Temporo-parietal junction (TPJ), superior/middle temporal gyrus, inferior frontal gyrus and insula	Left	−46	−64	35	47.0	1419
Superior and middle frontal gyrus, medial prefrontal cortex (mPFC), anterior cingulate	Left	−22	26	50	28.6	1299
Middle and inferior occipital cortex	Left	−25	−61	44	36.0	756
Middle and inferior occipital cortex	Right	23	−73	29	34.2	609
Inferior frontal gyrus and anterior insula	Right	41	23	−4	16.2	202
Temporo-parietal junction (TPJ)	Right	44	−70	29	15.4	175
Middle/superior frontal gyrus	Right	29	17	50	19.4	144
Middle/superior temporal gyrus	Right	65	−19	−4	21.2	119
B. Contrast between KK and UU						
KK > UU						
Temporo-parietal junction (TPJ), superior/middle temporal gyrus, inferior frontal gyrus and insula	Left	−46	−67	35	11.85	1279
Precuneus, cuneus, posterior cingulate, and cerebellum	Left	−4	−52	17	8.28	1088
Dorsolateral prefrontal cortex (DLPFC) and medial prefrontal cortex (mPFC)	Left	−10	26	59	7.91	716
Paracentral lobule	Left/Right	−1	−31	59	5.54	301
Temporo-parietal junction (TPJ)	Right	44	−73	32	6.01	140
Medial prefrontal cortex (mPFC) and anterior cingulate	Left/Right	−1	56	2	5.40	138
Middle/superior temporal gyrus	Right	65	−19	−4	5.41	129
Inferior frontal gyrus and anterior insula	Right	44	20	−1	4.98	113
UU > KK						
Middle occipital gyrus, fusiform gyrus and superior parietal cortex	Left	−28	−85	8	−9.50	603
Middle occipital gyrus and superior parietal cortex	Right	26	−70	29	−10.04	488
C. Contrast between KU and UU						
KU > UU						
Precuneus, cuneus and posterior cingulate	Left	−4	−52	17	8.70	778
Dorsolateral prefrontal cortex (DLPFC)	Left	−31	5	56	7.55	459
Temporo-parietal junction (TPJ)	Left	−49	−67	20	7.81	310
Medial prefrontal cortex (mPFC) and anterior cingulate	Left	−10	56	5	5.74	245
UU > KU						
Superior parietal cortex	Left	−25	−61	41	−8.28	193
Middle occipital gyrus and superior parietal cortex	Right	23	−73	29	−6.47	409
Middle occipital gyrus, fusiform gyrus and cerebellum	Left	−28	−79	11	−7.08	398
D. Contrast between KK and KU						
KK > KU						
Inferior frontal gyrus and anterior insula	Right	53	14	−1	5.49	144
Temporo-parietal junction (TPJ)	Left	−52	−49	32	5.51	140
Inferior frontal gyrus and anterior insula	Left	−46	11	−1	5.73	114

several brain regions that showed significant differences in brain activation across the conditions (KK, KU, UU). These regions included the bilateral temporo-parietal junction (TPJ), the precuneus and the medial prefrontal cortex (mPFC), as well as bilateral areas of the temporal, frontal and occipital cortex.

The results of the paired-sample *t*-tests between each pair of conditions are presented in Fig. 3 and Table 2 (B-D). The contrast between KK and UU revealed greater activation for the KK condition in the bilateral TPJ, the precuneus, the mPFC, the left dorsolateral prefrontal cortex (DLPFC), the bilateral inferior frontal gyrus extending into the insula, as well as the bilateral middle/superior temporal gyrus. Greater activation in the UU condition (versus KK) was observed in two clusters encompassing respectively the left and right middle occipital gyrus, fusiform gyrus and extending into the superior parietal cortex.

The contrast between KU and UU revealed greater activation for the KU condition in the left TPJ, the precuneus, the mPFC and the left DLPFC. Greater activation in the UU condition (versus KU) was observed in the occipital cortex and superior parietal cortex bilaterally, forming a single cluster in the right hemisphere and two separate clusters in the left hemisphere (see Table 2 and Fig. 3).

The contrast between KK and KU revealed greater activation in the KK condition in the left TPJ and bilaterally in clusters at the junction between the inferior prefrontal cortex and the insula. No region showed significantly greater activation for KU relative to KK.

4. Discussion

This pilot study introduces a novel fMRI research paradigm taking advantage of sparse-sampling fMRI to examine the neural correlates of referential communication. While it can be challenging to study online verbal interactions between a participant in the MRI scanner and a real interlocutor because of the noise and artefacts related to speaking, our paradigm allowed for real verbal communication while minimising the interference from noise and movement artefacts.

The referential communication task used for this study required that the participants present a series of movie characters to their interlocutor, and we first examined the brain activation that was common to all three task conditions of the referential communication task (KK, KU and UU relative to baseline) using a conjunction analysis. This analysis revealed a single, widespread cluster of activation consistent with the nature of the task, which involves the presentation of visual stimuli (pictures of movie/TV characters) as well as the planning and production of complex overt verbal responses (see Fig. 2A).

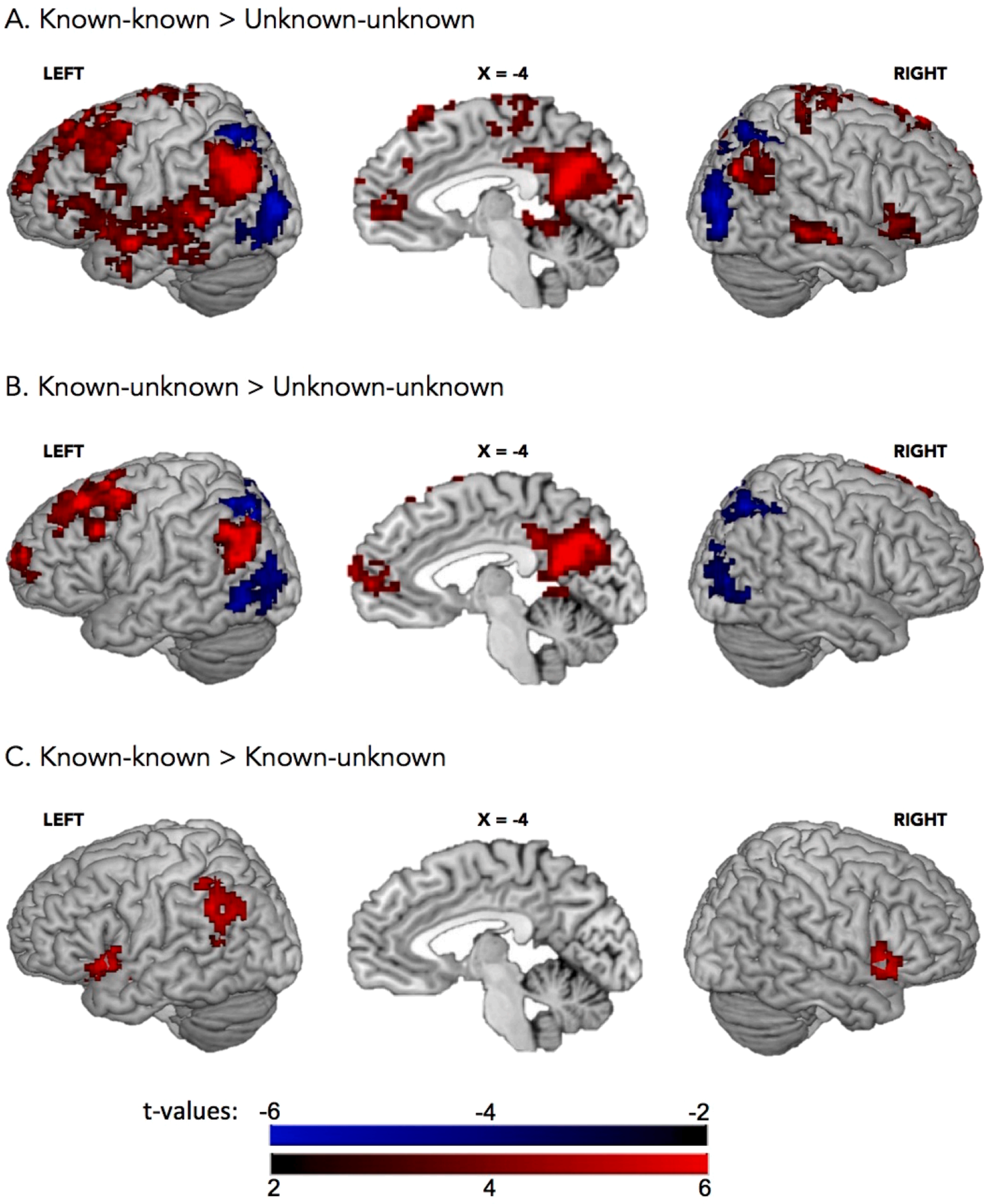

Fig. 3. Results from the comparisons between each pairs of condition.

We had also anticipated that if ToM is involved in taking the other person into account during verbal interactions, then the brain regions associated with ToM (Lavoie et al., 2016; Schurz et al., 2014; Spreng et al., 2009) should also be solicited during our interactive referential communication task, designed to trigger spontaneous verbal adjustments based on the other person's likely knowledge about the characters from different movies or TV shows (Achim et al., 2017; Achim et al., 2015). Previous behavioural studies (Achim, Guitton, Jackson, Boutin, & Monetta, 2013; Champagne-Lavau et al., 2009) and theoretical models (Brennan, Galati, & Kuhlen, 2010) have suggested that the collaborative verbal adjustments that occur during referential communication could be linked to ToM abilities. It was thus of particular relevance to examine whether the brain regions typically involved during ToM tasks (the bilateral TPJ, mPFC and precuneus) would also be solicited during our referential communication task. As will be further discussed below, activation was indeed observed in these brain regions, though not across all task conditions, but rather specifically for the conditions in which participants could make verbal adjustments to take the other person's likely knowledge into account (KK > UU and KU > UU).

4.1. Effect of experimental conditions

At the behavioural level, when the participants did not know the characters on a given trial (UU condition) they could only use descriptive information to present that character to their interlocutor. This is reflected in the observation that nearly all of the UU trials (99.3%) comprised only descriptive information. In contrast, when the participants knew the character, verbal adjustments to the other person's likely knowledge were possible and were indeed observed in participants' use of descriptive and/or character-related information, with significant differences in the use of different types of information between the KK and KU conditions (see Table 1). More specifically, character-related information (e.g. their name or their role in the movie) was more often used when the interlocutor was more likely to know the character (KK condition) whereas descriptive information was typically included when the interlocutor was less likely to know the character (KU condition). This pattern of results in the choice of referring expressions is consistent with that of prior behavioural studies (Achim et al., 2015; Gorman, Gegg-Harrison, Marsh, & Tanenhaus, 2013; Heller et al., 2012) and crucially occurs even if the participant himself knows all the characters in the KK and KU conditions, likely reflecting an adjustment to the interlocutor's likely knowledge of the characters.

At the neural level, the bilateral TPJ, mPFC and precuneus showed increased activation specifically for the KK and KU conditions, i.e. the two conditions in which the participants had the possibility to adjust how they presented the characters to their interlocutor (i.e. using either character-related information, descriptive information or both), as compared to the UU condition in which such adjustments were not possible (since participants can only describe the characters that they do not know). Together with the observation of a lack of common activation in these regions in the conjunction analysis between all three conditions, this pattern of results suggests that ToM-related brain regions may be specifically recruited when judgments about the other person's likely knowledge are useful to adjust verbal responses.

In contrast, activation in ToM-related brain regions was significantly less prominent when the participants did not know the characters (the UU condition). Initially, we had considered that ToM judgments about the interlocutor's likely knowledge could potentially occur even when participants did not know the characters. For example, the participants could have assumed that the interlocutor was unlikely to know a character that they did not themselves know. However, these judgments would have been little informed, and would not have allowed the participants to adjust their choices of referring expressions because only descriptive information could be used to present characters that are unknown. The lesser activation in ToM-related brain regions for this condition suggests that these regions are activated specifically when verbal production can be consequently adjusted, lending support to the suggestion by Vanlangendonck et al. (2018) that the ToM brain network comes into play specifically when speakers have to consider which information they share with their addressee.

Importantly, ToM is not a cognitive process per see, but rather a complex ability that involves the coordinated recruitment of several cognitive processes (e.g. Achim et al., 2020; Lavoie et al., 2016; Schaafsma, Pfaff, Spunt, & Adolphs, 2015; Schurz & Perner, 2015). Hence, while this study supports the idea of a link between referential communication and ToM, more work will be required to firmly establish which cognitive processes are involved during referential communication. It is noteworthy that the processes supported by the mPFC, TPJ and precuneus may not be specific to ToM and are likely to encompass other aspect of social or even non-social cognition (see Adolphs, 2009; Molapour et al., 2021 (In Press)). Nonetheless, the current study allowed us to identify that specific conditions trigger activation in the same brain regions typically observed for ToM tasks, namely those conditions in which referential adjustments are possible, and this is a major finding of this study.

It is also worth mentioning that direct comparison of brain activation between the KK and KU conditions revealed greater activation in the left TPJ (as well as in the bilateral anterior insula) for the KK condition, while no brain region showed significantly greater activation for the KU condition. At the behavioural level, collaboratively adjusting one's verbal productions to an interlocutor's likely knowledge involves finding a balance between providing enough information and not saying too much (Grice, 1975). During our referential communication task, the type of information that was useful to the interlocutor depended on her knowledge of each character. For the KU condition, character-related information was not sufficient to guide identification of the target character by the interlocutor and descriptive information was thus needed (e.g. "he has a white hat"). For the KK condition, on the other hand, character-related information was sufficient for identification, and adding descriptive information was unnecessary. The KU and KK conditions thus differed in the type of collaborative referential adjustments that they favoured. While the additional TPJ activation observed here for the KK condition could potentially suggest additional ToM processing for that condition, several alternate hypotheses could also explain this result. Even if the role of this brain region in ToM judgments is well established (Samson, Apperly, Chiavarino, & Humphreys, 2004; Schurz et al., 2014), the TPJ is also recruited by a range of other tasks including those targeting empathic perspective-taking (Vistoli, Achim, Lavoie, & Jackson, 2016) and attentional control (Geng & Vossel, 2013). Future studies are needed to further understand the role of the TPJ during referential communication and the range of conditions that lead to an increased recruitment of this brain region.

4.2. Additional observations beyond brain regions linked to ToM

An interesting, unexpected result for the UU condition is the greater activation observed in the occipital cortex, lateral precuneus and fusiform gyrus relative to KK and KU. Given the role of these regions in visual perception and visual attention (Ganis, Thompson, & Kosslyn, 2004), activation in these areas for the UU condition could reflect additional examination of the visual characteristics of these characters, which is unsurprising given that participants could only use descriptive information to present the characters on these trials. This idea would certainly be relevant to further examine, for example using the eye-tracking methodology that can provide detailed information about the time-course of underlying processes.

Another interesting observation is the greater activation in the left DLPFC for KK > UU and KU > UU, i.e. both conditions in which participants could adjust their use of character-related or descriptive information. In the literature, activation in this region is linked to cognitive control (Niendam et al., 2012), and the left DLPFC is not consistently activated during ToM tasks or for other kinds of social judgments. Interestingly, a previous study revealed activation of the DLPFC during ToM judgments specifically when participants had to inhibit a concurrent interpretation in order to make the correct judgment/response (Lavoie et al. 2016). Similar processes could be at play during referential communication, and the paradigm that we introduce here could certainly be modified to more directly test that hypothesis.

Finally, the direct comparison between the KK and KU conditions revealed greater activation not only in the TPJ but also in the anterior insula, which is part of the salience network. Menon and Uddin (2010) suggested that the anterior insula is involved in detecting stimuli for which there is a need to initiate additional attentional control. The greater activation for the KK condition could thus reflect the need to inhibit the production of superfluous descriptive information that is relevant only in the other two conditions (i.e. for 67% of the trials in the task). Here again, this suggestion awaits empirical validation and the current paradigm offers great opportunities for future studies that will take advantage of the methods to manipulate a range of factors involved in verbal interactions.

Overall, this pioneer pilot study allowed us to identify a set of brain regions involved during referential communication and to introduce a paradigm that will allow further investigations into this understudied topic.

4.3. Strengths, limitations and future directions

The main strength of this study is that it introduces a research paradigm that allows for real-time overt verbal interactions between a participant and a real interlocutor, using a sparse sampling acquisition sequence to limit the impact of noise and movement artefact. The main limitation is that this was a pilot, feasibility study with a relatively small sample size, and as such, the current results pave the way for future studies in larger groups of participants but should be interpreted with caution given the small sample size.

The use of a confederate who did not provide verbal feedback could also be considered as another limitation. More specifically, using a confederate is considered as a possible source of influence on the results of verbal interaction tasks, especially if the confederate does not really need to do the task (e.g. if the images to identify are the same with all the participants) or if they are made aware of the pattern of response that would support the hypothesis being tested (Kuhlen & Brennan, 2010, 2013). In the present study, however, the confederate really did perform the task of identifying the target image among the set of three images, since the targets and distractors were different for each participant. The feedback provided to the participant was however controlled for by restricting it to a visual feedback provided with a fixed timing. While allowing real-time feedback (verbal and/or non-verbal) would certainly make the task more interactive and naturalistic, it would also increase the difficulty of standardising the feedback between the different participants and, more importantly, between the different conditions since the confederate may act less naturally when having to disregard knowledge that she actually possesses, which applies only to the likely-unknown characters that the participants themselves know.

A third limitation is that the referential communication task used for this study included a single attempt at presenting each character, which means that when the target character was not initially identified by the confederate, there was no further occasion to build towards mutual understanding as would happen through turn-taking in more natural conversations. While this single-attempt strategy allowed us to focus on the neural correlates linked to the initial verbal presentation of new items (here movie/TV characters), it would also be interesting to study the neural correlates of conversational repairs in future studies, for example by repeating trials for which the target character has not been initially identified. In addition, future studies could also seek to determine if the mental model that participants have about the knowledge of their interlocutor evolves as the task progresses. While it is likely to occur even when the interlocutor displays typical knowledge of the characters (i.e. knowing many characters in the KK conditions and few in the KU condition), such effect could be even more prominent if the interlocutor rather displayed overall high or low knowledge of the characters across all conditions.

A fourth limitation that could be addressed in future studies is that we did not control for the level of knowledge that participants had about the different characters. It is thus possible that they had greater knowledge (social knowledge) of the characters in the KK than the KU conditions (e.g. we likely know more about Harry Potter than about Leonidas from the movie 300, even if we saw both movies and can recognize both characters). Future studies could thus attempt to quantify how familiar participants are with the different characters and take it into account in the analyses.

Overall, this study introduces a new method that paves the way to study different aspects of referential communication using fMRI, a recognized need in order to further our understanding of the neural mechanisms involved during real social interactions (Redcay & Schilbach, 2019).

CRediT authorship contribution statement

Amélie M. Achim: Conceptualization, Methodology, Resources, Writing – original draft, Visualization, Project administration, Funding acquisition. **Isabelle Deschamps:** Formal analysis, Writing – original draft, Writing – review & editing. **Élisabeth Thibaudeau:** Investigation, Project administration, Writing – review & editing. **Alexandra Loignon:** Methodology, Writing – review & editing. **Louis-Simon Rousseau:** Investigation, Writing – review & editing. **Marion Fossard:** Conceptualization, Methodology, Writing – review & editing, Funding acquisition. **Pascale Tremblay:** Conceptualization, Methodology, Resources, Writing – review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by a pilot grant from the Quebec Bio-imaging Network (QBIN) and additional funds from the Natural Sciences and Engineering Research Council of Canada (NSERC #341631-2013) and the Social Sciences and Humanity Research Council of Canada (SSHRC #410-2011-1628). The Fonds de Recherche du Québec – Santé supported our team through salary awards to AMA and PT, and through a studentship to ET. Technical support for protocol development and data acquisition was provided by the "Consortium d'imagerie en neuroscience et santé mentale de Québec" (CINQ) via a platform support grant (#3456) from the Brain Canada Foundation. The funding agencies had no involvement in the scientific aspects of the study or its publication.

The authors would like to thank Christine St-Pierre, Camille Plaisance, Alexandra Michaud and Catherine Lévesque who worked on different aspects of the project, including the creation of the stimuli, data collection, and the transcription and coding of the behavioural data.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bandc.2021.105801.

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, Cognitiion and Neuroscience, 32*(1), 21–36.

Achim, A. M., Fossard, M., Couture, S., & Achim, A. (2015). Adjustment of speaker's referential expressions to an addressee's likely knowledge and link with theory of mind abilities. *Frontiers in Psychology, 6*, 823. https://doi.org/10.3389/fpsyg.2015.00823

Achim, A. M., Guitton, M., Jackson, P. L., Boutin, A., & Monetta, L. (2013). On what ground do we mentalize? Characteristics of current tasks and sources of information that contribute to mentalizing judgments. *Psychological Assessment, 25*(1), 117–126. https://doi.org/10.1037/a0029137

Achim, A. M., Ouellet, R., Roy, M. A., & Jackson, P. L. (2012). Mentalizing in first-episode psychosis. *Psychiatry Research, 196*(2–3), 207–213. https://doi.org/10.1016/j.psychres.2011.10.011

Achim, A.M., & Thibaudeau, É. (2018). Test des Histoires Combinées — v2018R. Guide d'administration et cahier-réponse. Available at https://cervo.ulaval.ca/pages_perso_chercheurs/achim_a/Tests.html. Retrieved from.

Achim, A. M., Thibaudeau, É., Haesebaert, F., Parent, C., Cellard, C., & Cayouette, A. (2020). La cognition sociale: Construits, évaluation et pertinence clinique. *Revue de neuropsychologie, 12*, 46–69.

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology, 60*, 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514

Bilek, E., Ruf, M., Schäfer, A., Akdeniz, C., Calhoun, V. D., Schmahl, C., … Meyer-Lindenberg, A. (2015). Information flow between interacting human brains: Identification, validation, and relationship to social expertise. *Proceedings of the National Academy of Sciences, 112*(16), 5207–5212. https://doi.org/10.1073/pnas.1421831112

Birn, R. M., Bandettini, P. A., Cox, R. W., Jesmanowicz, A., & Shaker, R. (1998). Magnetic field changes in the human brain due to swallowing or speaking. *Magnetic Resonance in Medicine, 40*, 55–60.

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 301–345). Burlington, MA: Academic Press/Elsevier.

Cavallo, A., Lungu, O., Becchio, C., Ansuini, C., Rustichini, A., & Fadiga, L. (2015). When gaze opens the channel for communication: Integrative role of IFG and MPFC. *Neuroimage, 119*, 63–69. https://doi.org/10.1016/j.neuroimage.2015.06.025

Champagne-Lavau, M., Fossard, M., Martel, G., Chapdelaine, C., Blouin, G., Rodriguez, J.-P., & Stip, E. (2009). Do patients with schizophrenia attribute mental states in a referential communication task? *Cognitive Neuropsychiatry, 14*(3), 217–239.

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*(3), 162–173.

Eden, G. E., Joseph, J. E., Brown, H. E., Brown, C. P., & Zeffiro, T. A. (1999). Utilizing hemodynamic delay and dispersion to detect fMRI signal change without auditory interference: The behavior interleaved gradients technique. *Magnetic Resonance in Medicine, 41*, 13–20.

Edmister, W. B., Talavage, T. M., Ledden, P. J., & Weisskoff, R. M. (1999). Improved auditory cortex imaging using clustered volume acquisitions. *Human Brain Mapping, 7*(2), 89–97.

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Brain Research. Cognitive Brain Research, 20*(2), 226–241. https://doi.org/10.1016/j.cogbrainres.2004.02.012

Geng, J. J., & Vossel, S. (2013). Re-evaluating the role of TPJ in attentional control: Contextual updating? *Neuroscience & Biobehavioral Reviews, 37*(10 Pt 2), 2608–2620. https://doi.org/10.1016/j.neubiorev.2013.08.010

Gorman, K. S., Gegg-Harrison, W., Marsh, C. R., & Tanenhaus, M. K. (2013). What's learned together stays together: Speakers' choice of referring expression reflects shared experience. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 39*(3), 843–853. https://doi.org/10.1037/a0029467

Gracco, V. L., Tremblay, P., & Pike, B. (2005). Imaging speech production using fMRI. *Neuroimage, 26*(1), 294–301. https://doi.org/10.1016/j.neuroimage.2005.01.033

Grice, H. (1975). Logic and conversation. In P. C. J. M. (Ed.), Syntax and semantics, Vol 3, Speech and Acts (pp. 41–58). New York: Academic Press.

Guionnet, S., Nadel, J., Bertasi, E., Sperduti, M., Delaveau, P., & Fossati, P. (2012). Reciprocal imitation: Toward a neural basis of social interaction. *Cerebral Cortex, 22*(4), 971–978. https://doi.org/10.1093/cercor/bhr177

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliot, M. R., … Bowtell, R. W. (1999). Sparse temporal sampling in auditory fMRI. *Human Brain Mapping, 7*, 213–223.

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., … Bowtell, R. W. (1999). "Sparse" temporal sampling in auditory fMRI. *Human Brain Mapping, 7*(3), 213–223.

Heller, D., Gorman, K. S., & Tanenhaus, M. K. (2012). To name or to describe: Shared knowledge affects referential form. *Topics in Cognitive Science, 4*(2), 290–305. https://doi.org/10.1111/j.1756-8765.2012.01182.x

Isaacs, E. A., & Clark, H. H. (1987). References in conversations between experts and novices. *Journal of Experimental Psychology: General, 116*(1), 26–37.

Jasmin, K., Gotts, S. J., Xu, Y., Liu, S., Riddell, C. D., Ingeholm, J. E., … Martin, A. (2019). Overt social interaction and resting state in young adult males with autism: Core and contextual neural features. *Brain, 142*(3), 808–822. https://doi.org/10.1093/brain/awz003

Johnstone, T., Ores Walsh, K. S., Greischar, L. L., Alexander, A. L., Fox, A. S., Davidson, R. J., & Oakes, T. R. (2006). Motion correction and the use of motion covariates in multiple-subject fMRI analysis. *Human Brain Mapping, 27*, 779–788.

Koike, T., Tanabe, H. C., Okazaki, S., Nakagawa, E., Sasaki, A. T., Shimada, K., … Sadato, N. (2016). Neural substrates of shared attention as social memory: A hyperscanning functional magnetic resonance imaging study. *Neuroimage, 125*, 401–412. https://doi.org/10.1016/j.neuroimage.2015.09.076

Kuhlen, A. K., Bogler, C., Brennan, S. E., & Haynes, J. D. (2017). Brains in dialogue: Decoding neural preparation of speaking to a conversational partner. *Social Cognitive and Affective Neuroscience, 12*(6), 871–880. https://doi.org/10.1093/scan/nsx018

Kuhlen, A. K., & Brennan, S. E. (2010). Anticipating distracted addressees: How speakers' expectations and addressees' feedback influence storytelling. *Discourse Processes, 47*(7), 567–587. https://doi.org/10.1080/01638530903441339

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review, 20*(1), 54–72. https://doi.org/10.3758/s13423-012-0341-8

Lavoie, M. A., Vistoli, D., Sutliff, S., Jackson, P. L., & Achim, A. M. (2016). Social representations and contextual adjustments as two distinct components of the Theory of Mind brain network: Evidence from the REMICS task. *Cortex, 81*, 176–191.

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: A network model of insula function. *Brain Structure & Function, 214*(5–6), 655–667. https://doi.org/10.1007/s00429-010-0262-0

Molapour, T., Hagan, C. C., Silston, B., Wu, H., Ramstead, M., Friston, K., & Mobbs, D. (2021). Seven computations of the social brain. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsab024

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience, 12*(2), 241–268. https://doi.org/10.3758/s13415-011-0083-5

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Redcay, E., Dodell-Feder, D., Mavros, P. L., Kleiner, M., Pearrow, M. J., Triantafyllou, C., … Saxe, R. (2013). Atypical brain activation patterns during a face-to-face joint attention game in adults with autism spectrum disorder. *Human Brain Mapping, 34*(10), 2511–2523. https://doi.org/10.1002/hbm.22086

Redcay, E., Dodell-Feder, D., Pearrow, M. J., Mavros, P. L., Kleiner, M., Gabrieli, J. D. E., & Saxe, R. (2010). Live face-to-face interaction during fMRI: A new tool for social cognitive neuroscience. *Neuroimage, 50*, 1639–1647. https://doi.org/10.1016/j.neuroimage.2010.01.052

Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nature Reviews Neuroscience, 20*(8), 495–505. https://doi.org/10.1038/s41583-019-0179-4

Samson, D., Apperly, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Left temporoparietal junction is necessary for representing someone else's belief. *Nature Neuroscience, 7*(5), 499–500. https://doi.org/10.1038/nn1223

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences, 19*(2), 65–72.

Schilbach, L. (2015). The neural correlates of social cognition and social interaction. *Brain Mapping, 3*, 159–164.

Schurz, M., & Perner, J. (2015). An evaluation of neurocognitive models of theory of mind. *Frontiers in Psychology, 6*, 1610.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews, 42*, 9–34. https://doi.org/10.1016/j.neubiorev.2014.01.009

Spiegelhalder, K., Ohlendorfb, S., Regen, W., Feige, B., van Elst, L. T., Weiller, C., … Tüscher, O. (2014). Interindividual synchronization of brain activity during live verbal communication. *Behavioural Brain Research, 258*, 75–79. https://doi.org/10.1016/j.bbr.2013.10.015

Spreng, R. N., Mar, R. A., & Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience, 21*(3), 489–510. https://doi.org/10.1162/jocn.2008.21029

Stolk, A., Noordzij, M. L., Verhagen, L., Volman, I., Schoffelen, J. M., Oostenveld, R., … Toni, I. (2014). Cerebral coherence between communicators marks the emergence of meaning. *Proceedings of the National Academy of Sciences of the United States of America, 111*(51), 18183–18188. https://doi.org/10.1073/pnas.1414886111

Thibaudeau, É., Legendre, M., Villeneuve, K., Cellard, C., & Achim, A. (2018). Reliability of two measures of social cognition, the Combined stories task (COST) and the Situation task. *Psychiatry Research, 262*, 63–69.

Vanlangendonck, F., Willems, R. M., & Hagoort, P. (2018). Taking common ground into account: Specifying the role of the mentalizing network in communicative language production. *PLoS ONE, 13*(10), Article e0202943.

Vistoli, D., Achim, A. M., Lavoie, M. A., & Jackson, P. L. (2016). Changes in visual perspective influence brain activity patterns during cognitive perspective-taking of other people's pain. *Neuropsychologia, 85*, 327–336. https://doi.org/10.1016/j.neuropsychologia.2016.03.020

Willems, R. M., de Boer, M., de Ruiter, J. P., Noordzij, M. L., Hagoort, P., & Toni, I. (2010). A dissociation between linguistic and communicative abilities in the human brain. *Psychological Science, 21*(1), 8–14. https://doi.org/10.1177/0956797609355563

Wu, S., & Keysar, B. (2007). The effect of information overlap on communication effectiveness. *Cognitive Science, 31*(1), 169–181. https://doi.org/10.1080/03640210709336989