Testing Second Language Oral Proficiency in Direct and Semidirect Settings: A Social-Cognitive Neuroscience Perspective

Article *in* Language Learning · March 2011

DOI: 10.1111/j.1467-9922.2011.00635.x

LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

Testing Second Language Oral Proficiency in Direct and Semidirect Settings: A Social-Cognitive Neuroscience Perspective

Hyeonjeong Jeong
IDAC, Tohoku University

Hiroshi Hashizume
IDAC, Tohoku University

Motoaki Sugiura
IDAC, Tohoku University

Yuko Sassa
IDAC, Tohoku University

Satoru Yokoyama
IDAC, Tohoku University

Shuken Shiozaki
The Society for Testing English Proficiency

Ryuta Kawashima
IDAC, Tohoku University

This study was supported by the Japan Science and Technology Agency (JST)/Research Institute of Science and Technology for Society (RISTEX), JST/Core Research for Evolutional Science and Technology (CREST), and the Society for Testing English Proficiency (STEP). Hyeonjeong Jeong and Motoaki Sugiura were supported by a grant-in-aid for Japan Society for the Promotion of Science (JSPS) fellows and a grant-in-aid for young scientists (A) (KAKENHI 18680026). We would like to thank the team of researchers that contributed to the research reported in this article: Jamie Dunlea, Kensaku Ishimaki, Keita Nakamura, Yumi Uno, and Kazuaki Yanase. We also appreciate the editor and anonymous reviewers of *Language Learning* for their helpful comments on an earlier version of this article. Of course, all remaining errors are our own.

Correspondence concerning this article should be addressed to Hyeonjeong Jeong, Department of Functional Imaging, Institute of Development, Aging and Cancer (IDAC), Tohoku University, Seiryo-machi 4-1, Aoba-ku, Sendai 980-8575, Japan. Internet: jeong@idac.tohoku.ac.jp

DOI: 10.1111/j.1467-9922.2011.00635.x

This study used functional magnetic resonance imaging (fMRI) to identify differences in the neural processes underlying direct and semidirect interviews. We examined brain activation patterns while 20 native speakers of Japanese participated in direct and semidirect interviews in both Japanese (first language [L1]) and English (second language [L2]). Significantly greater activation was observed in the regions involved in social communication (the medial prefrontal cortex and the bilateral posterior superior temporal sulci) during the direct interview conducted in the L2 than during the semidirect interview conducted in the L2. In contrast, both the direct and semidirect interviews conducted in the L1 produced similar increases in activation in the same brain areas as those observed during the L2 direct interview. These findings suggest that the direct interview may have elicited L2 communicative ability to a greater degree than the semidirect interview. Furthermore, during the L2 direct interview, activity in the right superior temporal region, which is involved in the processing of paralinguistic features (e.g., prosody and intonation), was positively correlated with increased L2 oral proficiency. Based on our findings, we conclude that the L2 direct interview may elicit more balanced and varied aspects of communicative ability than the L2 semidirect interview.

Keywords fMRI, social communication, speaking assessment, face-to-face interaction, computer-based interaction OPI, SOPI

Introduction

The literature on the assessment of second language (L2) oral proficiency includes two widely used speaking test formats: a live (direct) format and a simulated, videotape-mediated (semidirect) format (Alderson & Banerjee, 2002; O'Loughlin, 2001; Shohamy, 1994). The direct interview format for L2 oral proficiency tests is generally conducted in real time and involves some form of face-to-face interaction with an interviewer. In contrast, the semidirect interview format removes the human interviewer to ensure that each respondent receives the same test. Although each format has advantages and disadvantages, the direct interview is thought to be more authentic than the semidirect interview (Clark, 1979) because it reflects the social conditions of human communication (i.e., face-to-face interaction). Real-time social interaction is a key element of the direct interview. In an attempt to identify the neural processes underlying direct and semidirect interviews conducted in the L2, we used functional magnetic resonance imaging (fMRI). Before discussing our study, we provide a brief review of the literature on direct and semidirect interviews in L2 assessment, followed by an overview of recent neuroimaging research on social cognition pertaining to the possible advantages of direct over semidirect interviews.

Direct and Semidirect Interviews: L2 Assessment Literature

Much of the previous research on direct and semidirect speaking test formats has centered on the comparison of several versions of the Simulated Oral Proficiency Interview (SOPI), a semidirect oral proficiency test developed by the Center for Applied Linguistics (CAL), and the Oral Proficiency Interview (OPI), a direct face-to-face oral proficiency test. Stansfield (1991) and Stansfield and Kenyon (1992) reviewed a number of concurrent validity studies conducted by CAL comparing various language versions of the SOPI and OPI. The results consistently showed high correlations between test scores for the two tests, which led Stansfield to conclude that the SOPI "has shown itself to be a valid and reliable surrogate of the OPI" (p. 200).

However, Shohamy (1994) and O'Loughlin (2001) questioned this reliance on correlations between test scores for the validation of semidirect speaking tests. To move beyond the comparison of test scores only, Shohamy conducted a detailed analysis of the content of the Hebrew version of the SOPI and OPI, in addition to examining the oral output of test-takers. Although the test scores for the Hebrew SOPI and OPI were also strongly correlated, the author found significant differences in terms of the functions and topics of the elicitation tasks. The test-takers' output also showed significant differences across a number of features, including rhetorical functions and structure, prosodic features, discourse functions, discourse markers, discourse strategies, speech moves, and the lexical density of the output.

O'Loughlin (2001) pointed out that these discrepancies may be due not only to the difference in interview formats (i.e., direct vs. semidirect) but also to a range of methodological factors, such as differences in the characteristics of the tasks, the degree of interactivity, and the preparation and response time given to interviewees. To overcome this problem and to focus more clearly on the differences inherent to the direct versus semidirect interview formats, O'Loughlin attempted to control the above-mentioned methodological factors. In his study he examined two versions of the Australian Assessment of Communicative English Skills (ACCESS) oral subtest: a direct interview version with an interviewer present to elicit output from the interviewee and a semidirect version relying on audio-taped stimuli for elicitation. Attempts were made to make the two versions as similar as possible in terms of task type, topic, and the degree of interaction, although a role-play task did require substantially more interaction in the live version (O'Loughlin, 2001). Because of these efforts to make the two versions similar, interviewee output did not differ greatly in terms of many of the features in which Shohamy (1994) noticed differences on the SOPI and OPI. However, differences in speech moves were noted in the interactive role-play task for the direct interview, and output on the direct version also showed slightly more variation in terms of prosodic features. Significant differences between the two versions were observed across tasks for lexical density (O'Loughlin, 2001).

O'Loughlin (2001) cautioned that the degree of interaction, rather than the format itself, may be the most important determining factor for lexical density, and, indeed, the largest differences in this parameter were observed in the role-play task, which included significant amounts of interaction in the live version. However, the author also noted that despite efforts to control the interviewers' output on the other tasks, the mere presence of the interviewer "called into play the candidate's interactive competence in a sense that is absent from the tape version" (O'Loughlin, 2001, p. 166). Interviewers and interviewees seemed to use both verbal and nonverbal feedback during the process of interacting in the live interview (O'Loughlin, 2001).

We argue that there may in fact be fundamental differences between direct and semidirect interviews in terms of the social dimension of interaction. In direct interviews that are conducted in real-time bidirectional settings with a live interviewer, the interviewee knows that there is a live interviewer who is listening to his/her answer. This is not the case for the semidirect interview. This social dimension of interaction may affect cognitive processes underlying L2 oral interviews. In the present study, we explored the neural processes underlying direct and semidirect interviews carried out in the L2 from a social-cognitive neuroscience perspective.

Social Communication: Neuroimaging Research

Language processing includes the processing of linguistic and paralinguistic information (e.g., prosody and tone of voice). Neuroimaging research has primarily focused on how linguistic information, such as syntax (Friederici & Kotz, 2003), semantics (Bookheimer, 2002; Gold et al., 2006), phonology (Hickok & Poeppel, 2004), and pragmatics (Braun, Guillemin, Hosey, & Varga, 2001), are processed at word, single-sentence, and discourse levels in the brain. The left inferior frontal gyrus (IFG), generally known as Broca's area, has been implicated in the strategic, controlled, or executive aspects of processing linguistic information, as reviewed in Indefrey (2006) and Skipper and Small (2005).[1] The superior temporal gyrus (STG), which contains the primary auditory cortex in the middle region and Wernicke's area in the posterior region, is concerned with the earliest stages of speech perception (Belin, Zatorre, & Ahad, 2002; Binder et al., 2000) and narrative comprehension (Braun, Guillemin, Hosey, & Varga, 2001; Xu, Kemeny, Park, Frattali, & Braun, 2005). The left inferior parietal regions, including the supramarginal gyrus and angular gyrus, play a role in phonological analyses and the storage and manipulation of speech stimuli (Paulesu, Frith, & Frackowiak, 1993; Ravizza, Delgado, Chein, Becker, & Fiez, 2004). Numerous neuroimaging studies have shown that the above fronto-temporal network, along with the Sylvian fissure (particularly in the left hemisphere), plays an important role in language processing (see Gernsbacher & Kaschak, 2003; Vigneau et al., 2006, for reviews). In contrast, the right frontal temporal networks may exert a major influence in the processing of paralinguistic features, such as prosody and tone of voice (see Lindell, 2006, for a review).

Humans utilize both linguistic and paralinguistic information to understand the pragmatic meanings of utterances, read the intentions of the speaker, and infer what the speaker is thinking and feeling during communication. This ability to infer the thoughts and feelings of others is referred to as *theory of mind* or *mentalizing* (Baron-Cohen, 1991; Premack & Woodruff, 1978). The ability to mentalize has been emphasized as a crucial factor in social communication (Bloom, 2000; Sperber & Wilson, 1995). According to neuroimaging research on social cognition, the medial prefrontal cortex, posterior superior temporal sulcus, and temporal pole underlie the mentalizing processes used in human communication (see Frith & Frith, 2007; Saxe, 2006, for reviews). For example, the medial prefrontal cortex is associated with thinking about the mental states of self and others and understanding the intentions operating in socially interactive contexts (Amodio & Frith, 2006; Kampe, Frith, & Frith, 2003; Walter et al., 2004). The temporal-parietal junction and the posterior superior temporal sulcus are related to perspective taking, reading intentions, and predicting intentions based on actions (Hamilton & Grafton, 2006; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004). A recent neuroimaging study conducted by Sassa et al. (2007) also demonstrated the involvement of these mentalizing networks in verbal communication. The authors found that the brain areas used for mentalizing showed greater activation when the participants communicated with others than when they simply described what was happening to others.

We assumed that the brain areas responsible for mentalizing are recruited to a greater degree during the direct interview than the semidirect interview. The direct interview involves some form of face-to-face interaction with an interviewer, prompting interviewees to attribute their mental states to others. In contrast, the semidirect interview is rather monologic in context; thus. mentalizing processes may be suppressed. Our review of L2 assessment and social-cognitive neuroscience research led us to the hypothesis that, compared to the semidirect interview, the direct interview may elicit active recruitment of the cortical networks underlying social communication.

Furthermore, previous neuroimaging research has repeatedly shown that increased efforts for L2 processing likely induce stronger activation in several brain regions including the left IFG than those for first-language (L1) processing (see Abutalebi, 2008; Indefrey, 2006, for a review; Jeong, Sugiura, Sassa, Haji et al., 2007). To replicate these consistent findings obtained from word-level and sentence-level processing, we attempted to examine how the type of language (L1 vs. L2) would affect brain activation during communication activities such as direct and semidirect interviews. To compare the effects of the L1 and L2 interviews, we conducted direct and semidirect interviews with the same participants in both languages and measured cortical activation during these interviews. In this way, we attempted to explore how interview type (direct vs. semidirect) and language type (L1 vs. L2) interacted. This study represents an initial attempt to explore some of the issues emerging in L2 assessment in a social-cognitive neuroscience context.

Methods

Participants

The sample consisted of 30 participants (20 men and 10 women), all of whom were right-handed native speakers of Japanese and had learned English as an L2. At the time of the experiment, all participants were undergraduate or graduate students. The mean (±*SD*) age was 22.86 ± 3.31 years (range: 19−31 years). They had started learning English at a mean age of 11.80 ± 0.35 years. All participants had studied English as a foreign language primarily in a classroom setting and had not lived outside Japan before the age of 10. Eleven of the participants had visited English-speaking countries for various personal reasons (e.g., learning English, sightseeing), with a mean length of stay of 6.05 ± 7.85 months (range: 1−24 months).[2]

All participants were required to take a placement test produced by the Society for Testing English Proficiency (STEP). This placement test provided an approximation of the test-takers' English ability in terms of the seven-level main suite of the EIKEN tests produced by STEP. To be included in the experiment, participants were required to demonstrate a minimum Grade-2-level performance. Grade 2 is an intermediate level of ability used by the Ministry of Education, Culture, Sports, Science, and Technology as a recommended benchmark for high school graduates in Japan. Written informed consent was obtained from each participant. This study was approved by the institutional review board of the Tohoku University Graduate School of Medicine, Sendai, Japan.

Two interviewers, a man and a woman, interviewed the 30 participants. The interviewers were native speakers of Japanese. Both interviewers worked as content specialists in the design and development of the EIKEN tests, and both were certified speaking test examiners for the EIKEN interview tests.

Experimental Procedure

Thirty participants participated in direct and semidirect interviews during fMRI scanning. We attempted to make all aspects of the two interview formats as parallel as possible (e.g., interview content, interviewer, time to answer, visual and audio information, and degree of interaction) to explore the differential effects of interview type on cortical activation. The demands placed on participants during the fMRI were identical between the direct and semidirect conditions. The only difference between the two conditions was the presence of a live interviewer in the direct condition versus a prerecorded interviewer in the semidirect condition. In the direct condition, participants were informed that live interviewers were listening to their answers, and the interviewers were told not to respond to the participants.[3]

During both the direct and semidirect interviews, an interviewer appeared on a video screen for 10 seconds to pose a question (see below for a sample) that the participant was to answer within 40 seconds. The direct interview was conducted in real time through a specially designed video system and dual-channel noise-canceling fiber-optical microphone (see Figure 1), although interviewers

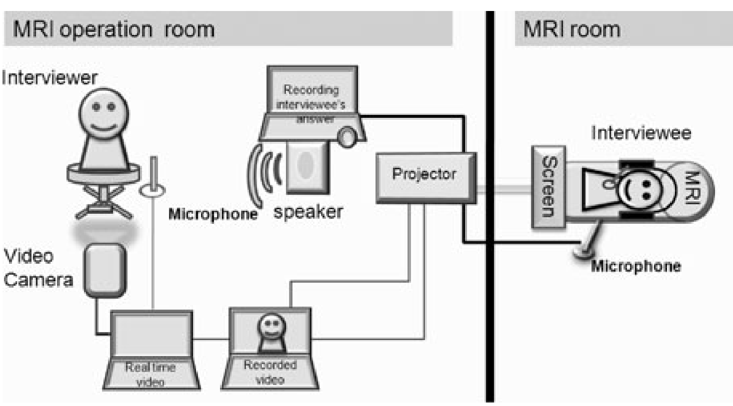


Figure 1 The experimental setup.

and interviewees were not interacting linguistically (i.e., were not asking or answering further questions). In contrast, during the semidirect interview, 50-second prerecorded videos were presented. The interviewer in this video clip asked a question for 10 seconds and then remained motionless but visible for the next 40 seconds. Thus, in both the direct and semidirect interviews, one trial lasted for a 50-second block, including the listening period (approximately 9–10 seconds) and answering period (approximately 40 seconds).

The experiment was conducted in both L2 (English) and L1 (Japanese). L2 and L1 interviews were conducted separately during two different sessions. In each language session, participants completed eight blocks consisting of four semidirect and four direct interview blocks. A rest period of 16 seconds, during which participants viewed a white cross mark on a black screen, was permitted between blocks. Each session lasted 546 seconds, including 18 seconds of rest at the beginning. The following four conditions were tested: L2 direct interview (DE), L2 semidirect interview (SE), L1 direct interview (DJ), and L1 semidirect interview (SJ). The order of interview type (direct or semidirect interview) and language (L1 or L2) was counterbalanced across participants.

In each session, participants practiced the direct/semidirect interview inside the MRI scanner to become accustomed to the experimental procedure and to learn how to minimize head movements during speech. Participants met the interviewer appearing in the direct condition outside the fMRI room and then performed the practice task inside the fMRI room. This procedure ensured that participants were aware that their interviews were being conducted in real time. In the semidirect condition, participants performed the practice task with the prerecorded movie. The gender of the examiners was varied between the conditions for each participant. The two interviewers were counterbalanced across participants and interview types. All answers were recorded with a dual-channel noise-canceling microphone (Optoacoustics Ltd., Israel).

After the fMRI experiment, each participant took a version of the full EIKEN Grade 2 face-to-face speaking test outside the MRI room. This speaking test included all the tasks normally included in the live EIKEN Grade 2 speaking test, including the read-aloud and narration tasks. Test scores from the full EIKEN speaking test were used to conduct correlation analyses between L2 proficiency scores and brain activation during L2 oral proficiency interviews.

Materials

A total of 16 questions were prepared and split across the four conditions (i.e., direct vs. semidirect interviews in L1 and L2). All 16 questions were based on questions used in past EIKEN speaking tests, which meant that a large amount of data was available in terms of difficulty. All questions were reviewed internally by a project team at STEP to ensure that they were equivalent in terms of difficulty, level of familiarity of the topics for typical Grade 2 test-takers, degree of sophistication of the topics, and appropriateness for eliciting oral responses from Grade-2-level test-takers. Typical examples of questions included "Some people say that young people's manners today are getting worse; what do you think about that?" or "Some people say that we are not doing enough to recycle things; what do you think about that?" The content of interview questions was counterbalanced between (a) interview type (direct or semidirect) and (b) language type (L2 or L1) across participants.

Data Acquisition

Scanning was performed using a 1.5-tesla Siemens Symphony system (Erlangen, Germany). Functional images were acquired using gradient echo planar image sequences with the following parameters: echo time = 50 ms, flip angle = 90°, slice thickness = 3 mm, slice gap = 0.99 mm, field of view = 192 mm, 64 × 64 matrix, and 3 × 3 × 3-mm voxels.[4] Thirty-three 3-mm-thick axial slices spanning the entire brain were obtained every 3 seconds. Excluding three dummy volumes for stabilization of the T1 saturation effect, 182 volumes were acquired for each participant in each session.[5] During functional imaging, the participant's head was secured using a foam rubber pad to minimize artifacts due to movement. In addition, T1-weighted anatomical images (thickness = 1 mm; field of view = 256 mm; 192 × 224 matrix; repetition time = 1,900 ms; echo time = 3.93 ms) were obtained from each participant to serve as a reference for anatomical correlates. For the sake of anatomical localization of hemodynamic activation effects, fMRI maps were superimposed on a mean image of the normalized T1-weighted anatomical images of all participants.

fMRI Analysis

Functional images were analyzed using the Statistical Parametric Mapping (SPM) software package, version 5 (see http://www.fil.ion.ucl.ac.uk/spm/), operating on a MATLAB platform. The fMRI analyses included (a) preprocessing of the data and (b) statistical analyses within participants and between participants.

Preprocessing[6]

First, functional images of each participant were corrected for differences in slice acquisition timing by resampling all slices in time to match the first slice (slice timing adjustment) and were realigned to the first volume to correct for head movement between volumes (realignment). Based on movement parameter estimates from the realignment procedure, we excluded participants who exceeded 3 mm translation from further analyses. Second, realigned functional images were coregistered to the T1 anatomical image from that participant. Third, the T1 images were spatially normalized to the standard Montreal Neurological Institute (MNI) brain template (available in the statistical parametric mapping software SPM5, see Friston, Ashburner, Kiebel, Nichols, & Penny, 2007) with the obtained transformation matrix applied to the coregistered functional image to allow for comparison across participants. These normalized functional images were spatially smoothed with a 9-mm Gaussian kernel (full width at half-maximum) to account for small variations in the localization of activation across participants.

Statistical Analysis

Ten members of the original pool of 30 participants were excluded. Six were excluded because the data revealed that head motion exceeded 3 mm (see Preprocessing subsection). The other four participants were excluded because they could not provide a sufficient answer for at least 1 of the 16 interview questions. Thus, the data obtained from 20 participants were analyzed to identify voxels in the image (i.e., brain regions) that were significantly activated during each interview.

Conventional first-level (within-subject) and second-level (between-subjects) analyses were performed using SPM5. In the first-level analysis, the degree of activation was estimated based on a voxel-by-voxel (i.e., each and every voxel in the image) multiple regression analysis of the time courses (Friston, Worsley, Frackowiak, Mazziotta, & Evans, 1995). A general linear model was constructed for each participant to analyze the hemodynamic response captured by functional images.

The predicted time course of the MRI signal was generated with consideration for the delay in the hemodynamic response (i.e., hemodynamic response function). Four separate regressors were used to model the hemodynamic response during each interview condition (DE, SE, DJ, and SJ). Because the involvement of cognitive processes may change dynamically during the interview period, we estimated the average degree of involvement over the entire interview period in each condition.[7] Contrast images between conditions were also generated for each participant.

We then performed statistical inferences on the contrasts of parameter estimates with a second-level between-subjects (random effects) model, using one-sample *t* tests. First, we compared brain activation during the L1 interview and L2 interview, irrespective of the type of interview (direct vs. semidirect). Second, we compared the brain activation patterns observed during the direct and semidirect interview conditions in the L1 (DJ relative to SJ), and then the same comparison was conducted for the L2 (DE relative to SE). The statistical threshold in the voxel-by-voxel analysis, assuming a search area of the whole brain, was set at $p < .001$ for height (i.e., t value of each voxel) and corrected to $p < .05$ for multiple comparisons using cluster size (i.e., number of voxels in the activated area).[8]

If there were interactions between interview type and language type, the activation profile at each activation peak was examined via two-way repeated-measures analysis of variance (ANOVA). Alpha was set at $p < .05$.

Finally, to identify the brain areas concerned with L2 oral communicative skills, we entered the contrast image (DE, SE, or DE-SE) resulting from the first-level analysis into the second-level simple regression analysis, using the L2 oral proficiency score (as determined in a face-to-face speaking test) as a predictor variable. The slope of the regression line relating L2 oral proficiency scores to brain activation was calculated at each voxel, and the t-map testing the difference of these parameter estimates from zero was thresholded at $p < .005$ for height and corrected to $p < .05$ for multiple comparisons using cluster size.

Results

Behavioral Results

The duration of speech for each interview condition was as follows: 31.87 ± 5.92 s for DE, 31.66 ± 5.47 s for SE, 32.16 ± 6.51 s for DJ, and 30.73 ± 6.38 s for SJ. Two-way repeated measures ANOVA was used to detect significant differences among these four conditions. However, no significant quantitative difference was observed between direct and semidirect interviews in terms of speech duration, $F(1, 19) = 0.95, p = .34$.[9]

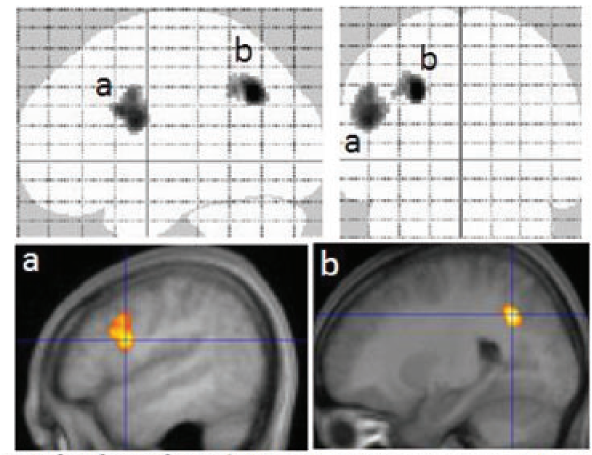
Imaging Results

We observed significant activation in the left IFG and left inferior parietal gyrus under the L2 interview condition compared to the L1 interview condition (Table 1 and Figure 2).[10] Furthermore, no brain area showed greater activation for the L1 interview condition than for the L2 interview condition. Significantly increased activation during the L2 direct interview condition (DE) compared to the L2 semidirect interview condition (SE) was observed in the medial prefrontal cortex, superior frontal gyrus, right posterior superior temporal sulcus, right IFG, left superior temporal sulcus, and cuneus (Table 2). In

Table 1 Brain areas showing greater activation during L2 versus L1 interviews

Brain region	Coordinates (*x, y, z*)	*t*	Cluster size
Left inferior frontal gyrus	−50, 6, 22	5.61	370
	−48, 14, 26	5.15	
Left inferior parietal gyrus	−24, −58, 38	7.47	232
	−30, −48, 38	4.44	

Note. Coordinates (MNI space), *t* values of peak activation, and cluster size (number of voxels) are shown for each activated area.



a. Left inferior frontal gyrus **b.** Left inferior parietal gyrus

Figure 2 Brain images showing significantly greater activation during the L2 versus L1 interview. The upper image shows a glass brain representation of activation in the sagittal and coronal planes. In the lower images, activation in (a) the left inferior frontal gyrus and (b) the left inferior parietal gyrus are superimposed on sagittal sections of the mean image for the normalized T1-weighted anatomical images of all participants. Activation was thresholded at $p < .001$ in the L2-L1 contrast and corrected to $p < .05$ for multiple comparisons using cluster size.

contrast, no significant difference in activation was detected in any brain area when we compared the L1 direct (DJ) and L1 semidirect (SJ) conditions. The activation profiles for each of the four conditions (Figure 3) revealed significant interactions between type of interview and type of language in the following areas: the medial prefrontal cortex, $F(1, 19) = 20.79, p < .001$; the superior frontal gyrus, $F(1, 19) = 13.52, p < .01$; the left posterior superior temporal sulcus, $F(1, 19) = 20.74, p < .001$; the right posterior superior temporal sulcus,

Table 2 Brain areas showing greater activation during the direct versus semidirect interview in the L2 condition

Brain region	Coordinates (*x, y, z*)	*t*	Cluster size
Medial frontal cortex	14, 56, 14	7.01	2,090
	−12, 32, 16	6.26	
	10, 24, 60	6.98	736
Superior frontal gyrus	−6, 6, 56	5.58	609
	24, 0, 56	5.50	
Right posterior superior temporal sulcus	58, −38, 8	7.46	1,285
	58, −48, 8	6.63	
Left posterior superior temporal sulcus	−58, −42, 8	6.42	1,472
	−46, −68, 14	5.56	
Right inferior frontal gyrus	32, 18, 28	5.57	878
Cuneus	−6, −74, 26	6.33	1,261

Note. Coordinates (MNI space), *t* values of peak activation, and cluster size (number of voxels) are shown for each activated area.

$F(1, 19) = 11.77, p < .001$; and the right IFG, $F(1, 19) = 12.91, p < .002$. Post hoc paired-sample *t* tests revealed significantly greater activation in the medial prefrontal cortex and bilateral superior temporal sulci during the DE, DJ, and SJ conditions than during the SE condition ($p < .05$). No significant differences in the activation patterns observed in the aforementioned areas were found (a) between the DJ and SJ conditions or (b) between the DE and DJ conditions.

The simple regression analyses revealed a significant positive correlation between the right superior temporal regions in the DE condition and scores on the L2 face-to-face speaking tests (Table 3 and Figure 4). However, no such significant correlation emerged in the SE and DE–SE contrasts.

Discussion

Differential Brain Activation During Direct and Semidirect Interviews in the L2

This study showed that the L2 direct interview produced significantly greater activation than the L2 semidirect interview in the following brain regions: the medial prefrontal cortex, the bilateral posterior superior temporal sulci, and the right frontal gyrus. The greater degree of activation in these areas during the direct interview compared to the semidirect L2 interview may be related to social communication in two ways.[11] First, participants knew that a live interviewer was listening during the direct interview; indeed, whether the

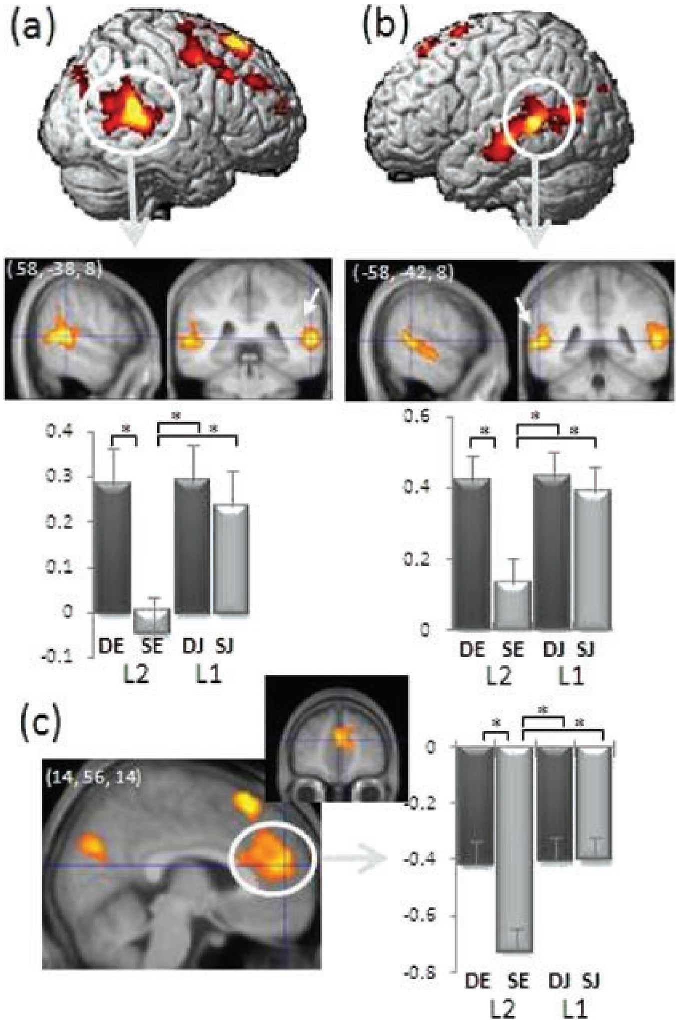


Figure 3 Brain images showing a significantly greater degree of activation in specific brain areas during the L2 interview under direct versus semidirect interview conditions. Activation profiles of peak activation in (a) the right posterior superior temporal sulcus (*x, y, z*: 58, −38, 8), (b) the left posterior superior temporal sulcus (*x, y, z*: −58, −42, 8), and (c) the medial prefrontal cortex (*x, y, z*: 14, 56, 14) are shown. The activation profile of each area represents the parameter estimates in each condition (DE, SE, DJ, and SJ). Error bars indicate standard error of the mean. $^{*}p < .05$.

Table 3 Brain area showing correlation between brain activation in the L2 direct interview condition and L2 oral proficiency scores

Brain region	Coordinates (*x, y, z*)	*t*	Cluster size
Right superior temporal cortex	68, −8, 12	4.86	526
	66, −14, 22	4.21	
	58, −26, 8	4.09	

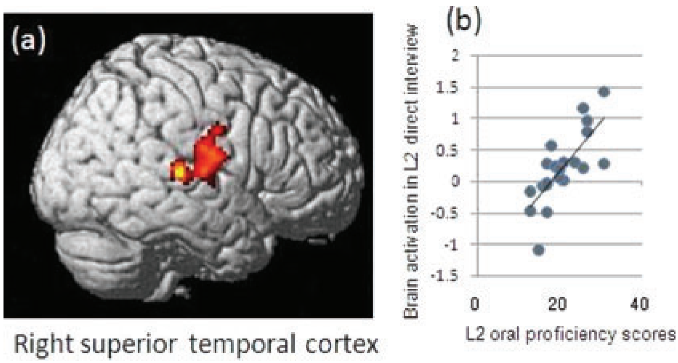


Figure 4 Results of correlation analyses for brain activation in the DE condition and oral proficiency scores determined by the EIKEN face-to-face interview. Image (a) shows significant activation in the right superior temporal cortex. Image (b) shows a positive correlation between the mean parameter estimates of peak voxels in this area for the DE condition and L2 oral proficiency scores.

interview was conducted in a real-time bidirectional setting represented the only difference between the direct and semidirect interviews. This awareness may have elicited interpersonal awareness in the direct interview. Consistent with this assumption, it has been suggested that the medial prefrontal cortex is involved in understanding communicative intention in socially interactive contexts (Kampe et al., 2003; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004; Sassa et al., 2007; Walter et al., 2004). For example, Rilling et al. (2004) made their participants believe that they were playing real-time interactive games with a human partner. Greater activation was found in these brain areas, which are related to mentalizing, when participants believed that they played games with a human partner than when the partner was a computer (see also McCabe, Houser, Ryan, Smith, & Trouard, 2001).

Second, the participants in the direct interview condition may have been sensitive to social signals, such as facial expression and eye gaze, and may have engaged in the mentalizing processes underlying social communication because they knew that the interviewer was actually present. Greater activation in the posterior superior temporal sulcus has been implicated in the processing of stimuli that elicit mentalizing with regard to both language and social signals (see Redcay, 2008, for a review). Activity in this brain region increased when the participants were engaged in comprehension and production of discourse or narratives (Braun et al., 2001; Sassa et al., 2007; Wilson, Molnar-Szakacs, & Iacoboni, 2008). Furthermore, activation of the posterior superior temporal sulcus has been demonstrated during the processing of facial expressions (Kilts, Egan, Gideon, Ely, & Hoffman, 2003), the processing of eye movements or eye gaze (Calder et al., 2002), the observation of the actions of others, and the reading of intentions from the actions of others (Saxe et al., 2004). In a study by Spiers and Maguire (2006), participants were requested to think about their own thoughts and those of others during a virtual reality game, and the right posterior superior temporal sulcus played a role in such mentalizing.

It should be noted that the L2 direct interview activated the right frontal cortex to a significantly greater degree than did the L2 semidirect interview. Although the role of this area remains open to speculation, it may be involved in the processing of pitch or accent patterns (Gandour et al., 2004; Geiser, Zaehle, Jancke, & Meyer, 2008). Activation of the right frontal area during the direct interview may reflect the controlled processing of voice tones (e.g., pitch or stress) because of the increased awareness accompanying the presence of a live interviewer.

Consistent with this reasoning, the L2 direct interview, which represents face-to-face interaction more authentically than the semidirect interview, may have involved mentalizing for participants to process paralinguistic features. Our findings indicate that the direct interview may have elicited L2 communicative language ability to a greater degree than did the semidirect interview.

Automatic Processing of the L1 and Controlled Processing of the L2

Our analyses revealed a differential effect according to interview type in the L2 interview but not in the L1 interview. This differential effect may be explained by differences in automaticity between L1 and L2 processing. In this study, we observed significantly greater activation in the brain areas associated with social communication during the L2 direct interview than during the L2 semidirect interview. Because of their limited proficiency levels in the L2 (i.e., controlled processing), the interviewees' cognitive capacity may have been recruited mainly for linguistic processing (e.g., grammatical, lexical, and phonological processing), leaving fewer cognitive resources for mentalizing.

However, the mere presence of a live interviewer in the direct interview condition may have prompted activity related to social communication in the L2 interview (O'Loughlin, 2001). During both the L1 direct and semidirect interviews, brain activity in the medial prefrontal cortex and bilateral posterior superior temporal sulci, which are thought to be responsible for social communication, increased (Figure 3). L1 linguistic processing is relatively automatic, thus freeing up the interviewee's cognitive resources for mentalizing. Therefore, brain activation patterns may not have been significantly affected by the type of interview during L1 speech.

This explanation is compatible with those offered in previous L2 studies (see Kormos, 2006, for a review). For example, Holtgraves (2007, 2008) demonstrated that L1 speakers automatically recognized speech acts (e.g., complaining, requesting, and agreeing) when reading conversational utterances, whereas L2 speakers did not demonstrate such recognition. Thus, this fMRI study indicates that the type of interview (i.e., direct vs. semidirect) may significantly affect the processes necessary for social communication depending on the degree of automaticity (i.e., L1 linguistic processing is relatively automatic, whereas L2 linguistic processing is relatively controlled).

Oral Proficiency in the L2 and the Right Hemisphere

During the L2 direct interview, participants with greater L2 oral proficiency tended to show greater engagement of the right superior temporal region.[12] This finding implies that participants with higher L2 oral proficiency levels may have employed prosodic features, such as intonation, pitch, stress, and rhythm, to communicate more effectively. In addition, they may have also utilized monitoring systems for speech generation to a greater extent than participants with lower L2 oral proficiency levels. Many previous neuroimaging studies support this interpretation. For example, the right hemisphere, including the superior temporal region, plays a pivotal role in processing the prosodic and paralinguistic aspects of speech (see Lindell, 2006, for a review; Meyer, Steinhauer, Alter, Friederici, & von Cramon, 2004). In particular, the right superior temporal region is associated with paralinguistic vocal information (Belin et al., 2002), generation of prosodic aspects of language (Dogil et al., 2002), and verbal self-monitoring during speech production (Fu et al., 2006). In previous studies, patients with right hemisphere damage were incapable of using prosodic features of speech (Pell, 1999) and were unable to appropriately organize discourse information (Weylman, Brownell, Roman, & Gardner, 1992). Additional neuroimaging research is needed to clarify the relationships between cortical mechanisms and the development of L2 communicative skills.

Concluding Remarks

This study provides neuroimaging evidence that the L2 direct interview involves a greater number of brain regions concerned with social communication than does the L2 semidirect interview. Furthermore, activity in the right superior temporal region, which is involved in processing paralinguistic features, is positively correlated with increased L2 oral proficiency during the L2 direct interview. These findings indicate that a real, face-to-face, L2 direct interview differs from a semidirect interview, at least at the cortical level. Given that social communication networks are activated during the L1 direct and semidirect conditions and the L2 direct condition, the L2 direct interview may be more appropriate than the L2 semidirect interview for eliciting L2 communicative ability.

To generalize these findings, further studies are required to examine those who have achieved advanced proficiency in L2 and those who have been exposed to L2 in real life on a daily basis. The participants in this study showed intermediate proficiency in English, had learned it as a foreign language, and tended to have very limited opportunities to speak it outside of the classroom. In addition, the L2 assessment literature includes various types of direct and semidirect interviews, as reviewed in McNamara and Roever (2006), but we used only one type of oral proficiency interview in this experiment. Therefore, additional studies are needed to examine the effect of various interview types on cortical activation. Furthermore, we treated brain activation while listening to and answering questions as an interview activity. Future studies are required to determine how the various parts of an interview (e.g., listening to questions vs. answering them) affect brain activation. Logistic problems inherent to conducting interviews inside the fMRI apparatus include the limited number of trials possible for each interview condition (due to fatigue), severe restrictions on head movement, and difficulty controlling the content of participants' responses (due to "natural" language production). Due to these limitations, our behavioral analyses were restricted to only the effect of the condition on speech duration. Future neuroimaging research in collaboration with L2 researchers is needed to investigate whether the qualitative aspects of the examinees' speech output in the direct and semidirect conditions (e.g., fluency, complexity, lexical density) reflect differences in neurological involvement. We believe that this study serves as a valuable first step in the investigation of the brain mechanisms underlying performance in L2 oral proficiency interviews.

Revised version accepted 19 January 2010

Notes

1 The IFG has multiple functions. The anterior and posterior parts of the IFG are related to syntactic processing (Embick, Marantz, Miyashita, O'Neil, & Sakai, 2000; Opitz & Friederici, 2007) and phonological processing (Poldrack et al., 1999), whereas the ventral part is associated with semantic processing (Wagner, Koutstaal, Maril, Schacter, & Buckner, 2000). The IFG also has important roles in acquiring the grammatical rules of a new language (Musso et al., 2003) and processing L2 linguistic information for those with late L2 onset or lower L2 proficiency (Indefrey, 2006).

2 Various factors, such as the age at first exposure to L2 (Perani et al., 2003), L2 proficiency level (Perani et al., 1998), and crosslinguistic influence (Jeong, Sugiura, Sassa, Haji et al., 2007; Jeong, Sugiura, Sassa, Yokoyama et al., 2007), are thought to influence cortical activation during L2 processing. In this study, the length of stay in English-speaking countries and L2 oral proficiency were taken into consideration. The length of stay in English-speaking countries was not correlated with brain activation pattern in the direct or semidirect interview in L2. See the Results and Discussion sections for the effect of L2 proficiency level on brain activation patterns.

3 Although we controlled the linguistic behaviors of the interviewers, the interviewers may have responded to participants in a nonlinguistic manner (e.g., facial expressions and nodding), which cannot be avoided in a live interview. However, this likely did not affect our findings, because we did not find any significant difference in brain activation between the L1 direct and semidirect conditions. Nonetheless, we acknowledge L2 behavioral research that has demonstrated that the nonlinguistic responses affect speech output (see Wolf, 2008). Therefore, these nonlinguistic factors in the L2 direct interview should be taken into account in further experiments.

4 Echo planar image is a rapid imaging technique for brain functional MRI that is sensitive to the local inhomogeneity of magnetic susceptibility, including that induced by the deoxy-hemoglobin (i.e., neural-activity-dependent metabolic change). A voxel is a basic sampling unit of the MRI image (see Jezzard, Matthews, & Smith, 2003 for the details).

5 The MRI signal is affected by two major time constants of the relaxation of magnetization, T1 and T2. T1 and T2 denote the time constants for longitudinal and transversal relaxation, respectively (see Jezzard et al., 2003, for the details).

6 The purpose of preprocessing is to remove various artifacts in the data, maximize the sensitivity of later statistical analyses, and increase statistical validity (Smith, 2004). We followed a standard pre-processing procedure in SPM (see Friston et al., 2007, for details).

7 Although dynamic changes in the involved cognitive processes may have increased the number of discrepancies between the model and observed time course of

activity, the subsequent increase in variation was not statistically significant in the second-level analysis.

8 In fMRI analysis, performing a statistical test at each voxel generates an enormous false positive rate when using unadjusted thresholds to detect significant activity. To avoid Type I errors, correction for multiple comparisons is necessary. Because functional imaging data are usually not spatially independent (i.e., data in one voxel are correlated with data from neighboring voxels), the Bonferroni correction may make the *p* threshold too conservative. To solve this problem, we used Gaussian random field (GRF) theory to threshold the image data. This method takes into account the spatial smoothness of the statistical map (i.e., the number of statistically independent voxels is lower; Friston, Worsley, Frackowiak, Mazziotta, & Evans, 1994). Thus, the corrected threshold under a GRF is much lower than when the Bonferroni correction is used. Finally, using GRF theory it is also possible to use cluster-based thresholding for inference. In cluster-level inference, clusters of voxels are created based on initial thresholding and then each cluster is assigned a *p*-value, which may or may not pass the final significance test (Friston et al., 2007; Smith, 2004).

9 For the purpose of this experiment, we could not avoid overt speech during scanning. Previous studies have reported that overt speech may produce susceptibility artifacts (e.g., signal drop out) in the fMRI signal (Kemeny, Ye, Birn, & Braun, 2005). Even if speech-related artifacts did play substantial roles in this experiment, no difference was observed among the four conditions in terms of duration of speech. Therefore, the influence may have been equivalent among the four conditions and comparisons of activity across direct and semidirect conditions may be valid.

10 These results are consistent with previous neuroimaging studies (Golestani et al., 2006; Indefrey, 2006; Jeong, Sugiura, Sassa, Haji et al., 2007): Increased efforts for L2 learners (particularly late learners) to process L2 linguistic information induce strong activation in these regions.

11 As an anonymous reviewer pointed out, activation in the right STG may be associated with task difficulty (Hasson, Nusbaum, & Small, 2007; Xu et al., 2005). Although we attempted to make the direct and semidirect interview conditions as similar as possible, task difficulty for the L2 direct condition may have been higher than for the L2 semidirect condition. However, the activation of several brain areas, such as the medial prefrontal gyrus and bilateral STGs, during the L2 direct condition was similar to that during the L1 condition (direct and semidirect). Considering that L1 interviews were easier than L2 interviews for our participants, task difficulty may not be sufficient to explain the greater activation in the above-mentioned brain areas (including the right STG) during both the L1 and L2 direct interviews.

12 No correlation was found between brain activation during the L2 semidirect interview and L2 oral proficiency. The contrasting findings between the direct and

semidirect interviews are interpreted as the effect of the presence of the interviewer during the direct interview on prompting brain activity related to social communication.

References

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica, 128*, 466–478.

Alderson, J. C., & Banerjee, J. (2002). State of the art review: Language testing and assessment (Part 2). *Language Teaching, 35*, 79–113.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7*, 268–277.

Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural theories of mind* (pp. 233–252). Oxford: Basil Blackwell.

Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Cognitive Brain Research, 13*, 17–26.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex, 10*, 512–528.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience, 25*, 151–188.

Braun, A. R., Guillemin, A., Hosey, L., & Varga, M. (2001). The neural organization of discourse: An H2 15O-PET study of narrative production in English and American sign language. *Brain, 124*, 2028–2044.

Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I., et al. (2002). Reading the mind from eye gaze. *Neuropsychologia, 40*, 1129–1138.

Clark, J. L. D. (1979). Direct versus semi-direct tests of speaking proficiency. In E. J. Briere & F. B. Hinofotis (Eds.), *Concepts in language testing: Some recent studies* (pp. 35–49). Washington, DC: TESOL.

Dogil, G., Ackermann, H., Grodd, W., Haider, H., Kamp, H., Mayer, J., et al. (2002). The speaking brain: A tutorial introduction to fMRI experiments in the production of speech, prosody, and syntax. *Journal of Neurolinguistics, 15*, 59–60.

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., & Sakai, K. L. (2000). A syntactic specialisation for Broca's area. *Proceeding of National the Academy of Sciences USA, 23*, 6150–6154.

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: Functional imaging and lesion studies. *NeuroImage, 20*, S8–S17.

Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology, 17*, R724–R732.

Friston, K. J., Ashburner, J., Kiebel, S., Nichols, T., & Penny, W. (Eds.). (2007). *Statistical parametric mapping: The analysis of functional brain images*. London: Academic Press.
Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2*, 189–210.
Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., & Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping, 1*, 214–220.
Fu, C. H., Vythelingum, G. N., Brammer, M. J., Williams, S. C., Amaro, E., Jr., Andrew, C. M., et al. (2006). An fMRI study of verbal self-monitoring: Neural correlates of auditory verbal Feedback. *Cerebral Cortex, 16*, 969–977.
Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., et al. (2004). Hemispheric roles in the perception of speech prosody. *NeuroImage, 23*, 344–357.
Geiser, E., Zaehle, T., Jancke, L., & Meyer, M. (2008). The neural correlate of speech rhythm as evidenced by metrical speech processing. *Journal of Cognitive Neuroscience, 20*, 541–552.
Gernsbacher, M. A., & Kaschak, M. P. (2003). Neuroimaging studies of language production and comprehension. *Annual Review of Psychology, 54*, 91–114.
Gold, B. T., Balota, D. A., Jones, S. J., Powell, D. K., Smith, C. D., & Andersen, A. H. (2006). Dissociation of automatic and strategic lexical-semantics: Functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *Journal of Neuroscience, 26*, 6523–6532.
Golestani, N., Alario, F. X., Meriaux, S., Le Bihan, D., Dehaene, S., & Pallier, C. (2006). Syntax production in bilinguals. *Neuropsychologia, 44*, 1029–1040.
Hamilton, A. F., & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *Journal of Neuroscience, 26*, 1133–1137.
Hasson, U., Nusbaum, H. C., & Small, S. L. (2007). Repetition suppression for spoken sentences and the effect of task demands. *Journal of Cognitive Neuroscience, 18*, 2013–2029.
Hickok, B., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition, 92*, 67–99.
Holtgraves, T. (2007). Second language learners and speech act comprehension. *Language Learning, 57*, 595–610.
Holtgraves, T. (2008). Automatic intention recognition in conversation processing. *Journal of Memory and Language, 58*, 627–645.
Indefrey, P. (2006). A meta-analysis of hemodynamic studies on first and second language processing: Which suggested differences can we trust and what do they mean? *Language Learning, 56*, 279–304.
Jeong, H., Sugiura, M., Sassa, Y., Haji, T., Usui, N., Taira, M., et al. (2007). Effect of syntactic similarity on cortical activation during second language processing:

A comparison of English and Japanese among native Korean trilinguals. *Human Brain Mapping, 28*, 194–204.

Jeong, H., Sugiura, M., Sassa, Y., Yokoyama, S., Horie, K., Sato, S., et al. (2007). Cross-linguistic influence on brain activation during second language processing: An fMRI study. *Bilingualism: Language and Cognition, 10*, 175–187.

Jezzard, P., Mathews, P. M., & Smith, S. M. (2003). *Functional MRI: An introduction to methods*. Oxford: Oxford University Press.

Kampe, K. K., Frith, C. D., & Frith, U. (2003). “Hey John”: Signals conveying communicative intention toward the self activate brain regions associated with “mentalizing,” regardless of modality. *Journal of Neuroscience, 23*, 5258–5263.

Kemeny, S., Ye, F. Q., Birn, R., & Braun, A. R. (2005). Comparison of continuous overt speech fMRI using BOLD and arterial spin labeling. *Human Brain Mapping, 24*(*3*), 173–183.

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage, 18*, 156–168.

Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Erlbaum.

Lindell, A. K. (2006). In your right mind: Right hemisphere contributions to language processing and production. *Neuropsychology Review, 16*, 131–148.

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences USA, 98*, 11,832–11,835.

McNamara, T. F., & Roever, C. (2006). *Language testing: The social dimension*. Oxford: Blackwell.

Meyer, M., Steinhauer, K., Alter, K., Friederici, A. D., & von Cramon, D. Y. (2004). Brain activity varies with modulation of dynamic pitch variance in sentence melody. *Brain and Language, 89*, 277–289.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Buchel, C., et al. (2003). Broca’s area and the language instinct. *Nature Neuroscience, 6*, 774–781.

O’Loughlin, K. (2001). *The equivalence of direct and semi-direct speaking tests*. Cambridge: Cambridge University Press.

Opitz, B., & Friederici, A. D. (2007). The neural basis of processing sequential and hierarchical syntactic structures. *Human Brain Mapping, 28*, 585–592.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural components of the verbal component of working memory. *Nature, 362*, 342–344.

Pell, M. D. (1999). The temporal organization of affective and nonaffective speech in patients with right-hemisphere infarcts. *Cortex, 35*, 455–477.

Perani, D., Paulesu, E., Sebastian-Galles, N., Dupoux, E., Dehaene, S., Bettinardi, V., et al. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain, 121*, 1841–1852.

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., et al. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: A fMRI study during verbal fluency, *Human Brain Mapping, 19*, 170–182.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage, 10*, 15–35.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavior Brain Science, 4*, 515–526.

Ravizza, S. M., Delgado, M. R., Chein, J. M., Becker, J. T., & Fiez, J. A. (2004). Functional dissociations within the inferior parietal cortex in verbal working memory. *NeuroImage, 22*, 562–573.

Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: Implications for the emergence of autism. *Neuroscience and Biobehavior Review, 32*, 123–142.

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage, 22*, 1694–1703.

Sassa, Y., Sugiura, M., Jeong, H., Horie, K., Sato, S., & Kawashima, R. (2007). Cortical mechanism of communicative speech production. *NeuroImage, 37*, 985–992.

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion Neurobiology, 16*, 235–239.

Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia, 42*, 1435–1446.

Shohamy, E. (1994). The validity of direct versus semi-direct oral tests. *Language Testing, 11*, 99–123.

Skipper, J. I., & Small, S. L. (2005). fMRI studies of language. In K. Brown (Ed.), *The encyclopedia of language & linguistics* (2nd ed., pp. 496–511). Oxford: Elsevier Science.

Smith, S. M. (2004). Overview of fMRI analysis. *British Journal of Radiology, 77*, S167–S175.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Malden, MA: Blackwell.

Spiers, H. J., & Maguire, E. A. (2006). Spontaneous mentalizing during an interactive real world task: An fMRI study. *Neuropsychologia, 44*, 1674–1682.

Stansfield, C. W. (1991). A comparative analysis of simulated and direct oral proficiency interviews. In S. Anivan (Ed.), *Current developments in language testing* (pp. 199–209). Singapore: Regional English Language Center.

Stansfield, C. W., & Kenyon, D. M. (1992). Research on the comparability of the Oral Proficiency Interview and the Simulated Oral Proficiency Interview. *System, 20*, 347–364.

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., et al. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage, 30*, 1414–1432.

Wagner, A. D., Koutstaal, W., Maril, A., Schacter, D. L., & Buckner, R. L. (2000). Task-specific repetition priming in left inferior prefrontal cortex. *Cerebral Cortex, 10*, 1176–1184.

Walter, H., Adenzato, M., Ciaramidaro, A., Enrici, I., Pia, L., & Bara, B. G. (2004). Understanding intentions in social interaction: The role of the anterior paracingulate cortex. *Journal of Cognitive Neuroscience, 16*, 1854–1863.

Weylman, S. T., Brownell, H. H., Roman, M., & Gardner, H. (1989). Appreciation of indirect requests by left- and right-brain-damaged patients: The effects of verbal context and conventionality of wording. *Brain and Language, 39*, 580–591.

Wilson, S. M., Molnar-Szakacs, I., & Iacoboni, M. (2008). Beyond superior temporal cortex: Intersubject correlations in narrative speech comprehension. *Cerebral Cortex, 18*, 230–242.

Wolf, J. P. (2008). The effects of backchannels on fluency in L2 oral task production. *System, 36*, 279–294.

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *NeuroImage, 25*, 1002–1015.