http://eprints.gla.ac.uk/56648

Deposited on: 13 October 2011

Silent Reading of Direct versus Indirect Speech Activates Voice-selective Areas in the Auditory Cortex

Bo Yao, Pascal Belin, and Christoph Scheepers

Abstract

■ In human communication, direct speech (e.g., *Mary said: "I'm hungry"*) is perceived to be more vivid than indirect speech (e.g., *Mary said [that] she was hungry*). However, for silent reading, the representational consequences of this distinction are still unclear. Although many of us share the intuition of an "inner voice," particularly during silent reading of direct speech statements in text, there has been little direct empirical confirmation of this experience so far. Combining fMRI with eye tracking in human volunteers, we show that silent reading of direct versus indirect speech engenders differential brain activation in voice-selective areas of the auditory cortex. This suggests that readers are indeed more likely to engage in perceptual simulations (or spontaneous imagery) of the reported speaker's voice when reading direct speech as opposed to meaning-equivalent indirect speech statements as part of a more vivid representation of the former. Our results may be interpreted in line with embodied cognition and form a starting point for more sophisticated interdisciplinary research on the nature of auditory mental simulation during reading. ■

INTRODUCTION

The distinction between direct and indirect speech exists in many languages (Coulmas, 1986). Direct speech (as in *Mary said: "Gosh! The movie was terrible!"*) is assumed to entail a demonstration or depiction of the reported utterance, whereas its indirect speech counterpart (as in *Mary said that the movie was terrible*) provides a mere description of what was said (Clark & Gerrig, 1990). As a result, a direct speech demonstration is generally assumed to be more vivid and perceptually engaging than an indirect speech description. Not only has this vividness distinction been observed and discussed by linguists (Tannen, 1986, 1989), it has also been shown, for instance, that in reporting previously overheard dialogues, speakers are more likely to employ direct rather than indirect speech when instructed to be entertaining to a listener (Wade & Clark, 1993).

However, little research so far (see Bohan, Sanford, Cochrane, & Sanford, 2008, for a recent exception) has addressed the question of how the two reporting styles are represented in language comprehension, particularly during silent reading of text where no auditory stimulation or visual stimulation other than text is present. Although many of us share the intuition of hearing an "inner voice" during silent reading (specifically of direct speech statements), there has been hardly any directly measurable confirmation of this experience so far. This is surprising, given that recent embodied cognition theories (Barsalou, 1999, 2008), for example, propose that language comprehenders mentally simulate linguistically described situations based on generalized perceptual experiences they have made in the past. This suggests that, even during silent reading of text, direct speech may be more likely to activate "audible speech"-like representations than indirect speech. Crucially, aspects of the reported speaker's voice are very likely to be part of this perceptual simulation process. In other words, readers may be more likely to mentally simulate the reported speaker's voice (or aspects thereof) during silent reading of direct rather than indirect speech.

One way to test this hypothesis is by measuring "top–down" activation of the auditory cortex during silent reading of direct versus indirect speech statements. From the literature, it is known that certain areas in the auditory cortex are selectively sensitive to human voices when stimulated "bottom–up" via auditory sound clips (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000). On the other hand, studies on nonverbal (Bunzeck, Wuestenberg, Lutz, Heinze, & Jancke, 2005; Yoo, Lee, & Choi, 2001) and verbal (Jancke & Shah, 2004; Shergill et al., 2001) auditory imagery, as well as experiments on visual speech perception (also known as *lip reading*; MacSweeney et al., 2000; Calvert et al., 1997), indicate that areas within the auditory cortex are also prone to "top–down" activation without explicit stimulation from the auditory modality. Moreover, Spitsyna, Warren, Scott, Turkheimer, and Wise (2006) showed that neural pathways for silent reading and speech perception converge in the STS regions of the auditory cortex, suggesting the possibility that silent reading might recruit some auditory processing. Hence, if readers are more likely to engage in perceptual simulations of the reported speaker's

University of Glasgow

voice during silent reading of direct speech, voice-selective areas in the auditory cortex (Belin et al., 2000) should display enhanced "top–down" activation as soon as readers come across a direct speech statement (as opposed to a meaning-equivalent indirect speech statement) in written text. The following event-related fMRI experiment aimed at testing this hypothesis. Participants silently read a number of short written stories for comprehension, while their eye movements and brain activations were monitored simultaneously.

METHODS

Subjects

In total, 26 adult participants were recruited and scanned. All of them were native English speakers with normal vision and hearing, no learning or reading disabilities, and no history of neurological or psychiatric disorders. Ten participants had to be excluded from analysis due to either (a) no clear response in the voice localizer task (see Stimuli and Tasks) and/or excessive head movements during scanning (eight subjects), (b) eye-tracking data loss (one subject), or (c) less than 70% answering accuracy on comprehension questions (one subject). Data from the remaining 16 participants (age, 18–44 years; 6 men and 10 women) were valid for the final analyses. All of them were right-handed, except for one female subject.

Stimuli

Ninety short stories with different protagonists (indicated by different names) were prepared as reading materials; the stimuli can be downloaded at: www.psy.gla.ac.uk/~christop/JOCN_2011/Stimuli.pdf. Each story started with two declarative sentences to set up a scenario (e.g., *Ph.D. student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension, but Jim looked concerned.*), followed by either a direct or an indirect speech sentence (e.g., *He said: "Hmm, we really need those data in by next month for that conference."* or *He said that they really needed those data in by next month for that conference.*). Crucially, the reported clauses (underscored in the above examples) were equivalent in terms of linguistic content. Comprehension questions (e.g., *Was Ella Jim's Ph.D. student?*) were also prepared for 23 stories (ca. 25%) to assess participants' overall comprehension accuracy and to ensure that they read the stories attentively. Two lists of stimuli with counterbalanced item–condition combinations (45 direct and 45 indirect speech items per list) were constructed. Each story appeared only once per list, but in a different condition across lists. Half of the 16 valid participants saw Presentation List 1, and the other half,

Presentation List 2. The presentation order of the stories per list was randomized for each participant.

For the voice localizer session (see below), we presented blocks of vocal sounds and nonvocal sounds provided by the Voice Neurocognition Laboratory (vnl.psy.gla.ac.uk), Center for Cognitive Neuroimaging, University of Glasgow. These stimuli were the same as those employed in Belin et al. (2000) and comprised both speech (e.g., spoken vowels) and nonspeech (e.g., laughing and coughing) vocal sound clips, as well as nonvocal sound clips (e.g., telephone ringing and dog barking). The contrast in brain activity elicited by vocal versus nonvocal sounds reliably localizes voice-selective areas of the auditory cortex.

Tasks

Participants were positioned in the scanner, wearing goggles (NordicNeuroLab, Bergen, Norway) for visual presentation and eye tracking (Viewpoint Eye-Tracker, Arrington Research, Inc., Scottsdale, AZ), as well as MRI-compatible, electrostatic headphones (NordicNeuroLab, Bergen, Norway) for noise attenuation during fMRI scanning and for auditory presentation during the voice localizer session (see below). After a brief eye-tracker calibration procedure, the main reading session followed during which the participant's brain was scanned and their eye movements were recorded. Participants were instructed to read the texts silently and carefully so as to be able to answer comprehension questions, which would follow after 25% of the short stories they had read. The stimulus materials were presented using E-prime 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA). The texts were presented in a black 15-pt Courier New font on a light gray background. The whole text was centered and wrapped within 75% width of the screen, together making the reading stimuli appear as natural as possible. Each trial began with the presentation of a fixation cross in the middle of the screen for 2 sec, followed by the presentation of the text for reading. Each story was presented in a sentence-by-sentence fashion. The presentation duration per sentence display was determined as $W \times 100$ msec $+ S \times 50$ msec (wherein W refers to the number of words and S refers to the number of syllables per sentence), allowing sufficient time for reading. Mean presentation durations for the final (critical) sentence display were 7514 msec ($SD = 1412$ msec) and 7526 msec ($SD = 1354$ msec) for the direct and indirect speech condition, respectively. About 25% of the text presentations were followed by a comprehension question regarding the content of preceding story. Each such question appeared in the middle of the screen, prompting a "yes" or "no" response, which participants could provide by pressing buttons on a response box with their index or middle fingers, respectively. Figure 1A provides a schematic illustration of the experimental trial sequence. The 90 reading trials were evenly

Figure 1. Experimental procedures for (see Methods). (A) Silent reading task. (B) Voice localizer task.
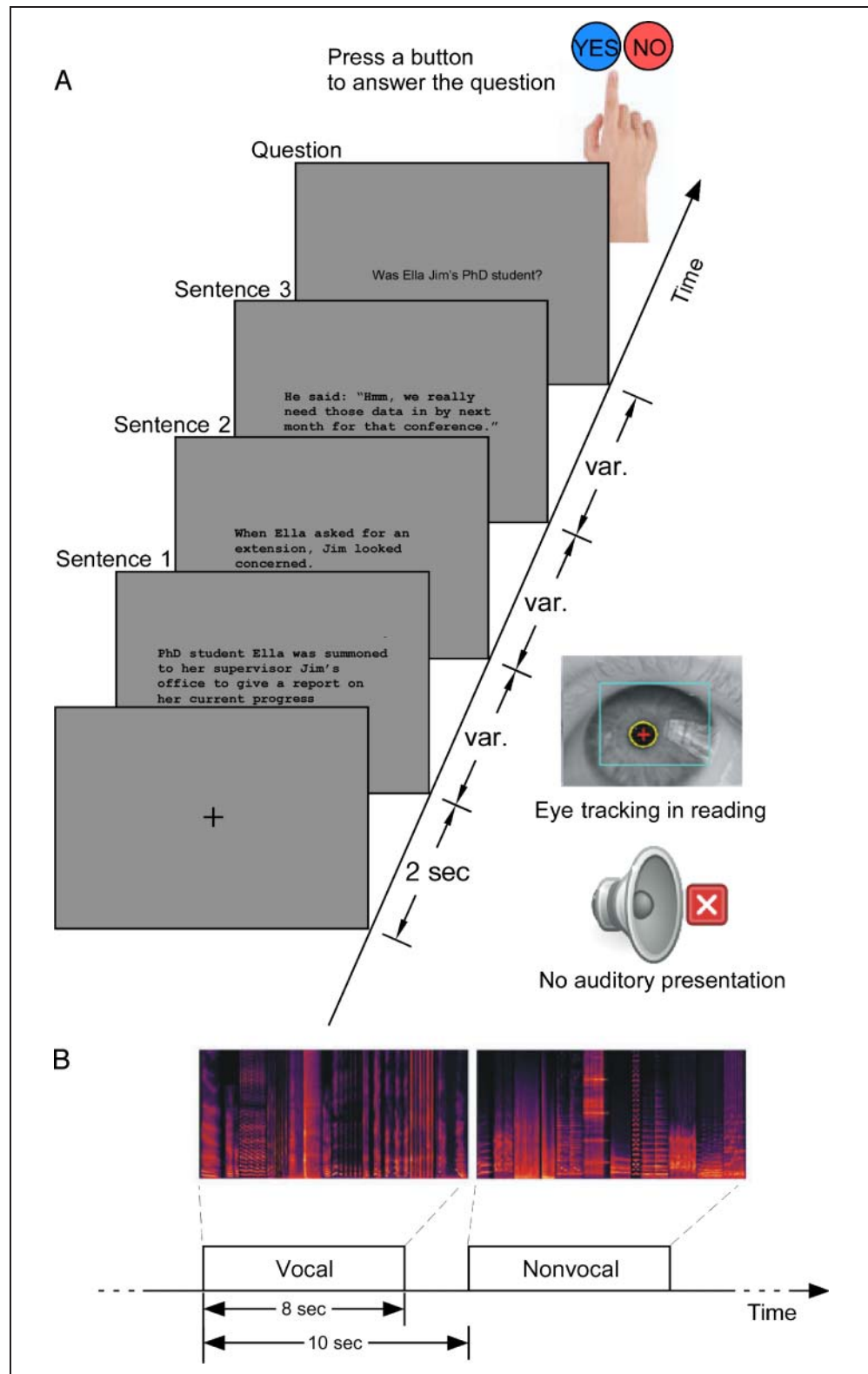
interspersed with five "baseline" trials, during which a plain fixation cross was presented in the center of the screen for 30 sec.

After the main reading session, an anatomical scan of the participant's brain was performed, which was then followed by a brief, 10-min voice localizer scanning session. During the latter, participants were instructed to close their eyes while listening to twenty 8-sec blocks of vocal and twenty 8-sec blocks of nonvocal auditory stimuli presented in an efficiency optimized, pseudorandom order along with 20 blocks without stimulation acting as a baseline (cf. Belin et al., 2000).

MRI Acquisition

Scanning was performed on a 3-T Siemens Tim Trio MRI scanner using a 12-channel head coil (Erlangen, Germany). Functional scans (for both the reading session and voice localizer session) were acquired using a T2*-weighted EPI sequence (32 slices acquired in orientation of the Sylvian fissure; repetition time (TR) = 2 s, echo time (TE) = 30 msec, matrix size = 70×70, voxel size = $3 \times 3 \times 3$ mm, field of view (FOV) = 210). T1 whole-brain anatomical scans were obtained using 3-D T1-weighted magnetization prepared rapid acquisition gradient-echo sequence (192 axial slices; matrix size = 256×256, voxel size = $1 \times 1 \times 1$ mm, FOV = 256). The average scanning time for the whole experiment was around 53 min per participant.

Data Analysis

All MRI data were analyzed using SPM8 (www.fil.ion.ucl. ac.uk/spm/, University College London). Preprocessing of functional scans included (a) head motion corrections (trilinear interpolation), whereby scans were realigned to the first volume; (b) coregistration of functional scans to their corresponding individual anatomical scans; (c) segmentation of the coregistered scans; (d) normalization of functional (3-mm isotropic voxels) and anatomical (1-mm isotropic voxels) data to Montreal Neurological Institute space; and (e) smoothing of normalized data (8-mm Gaussian kernel).

fMRI data from the anatomical and voice localizer scanning sessions were used to determine the voice-selective areas in the auditory cortex of each participant at $p < .001$ false discovery rate (FDR) corrected. The group voice localizer was obtained at $p < .001$ (uncorrected, to increase sensitivity against the background of individual differences).

For the reading scanning session, the temporal onset of a critical fMRI event was defined (via eye tracking) as the temporal onset of the first fixation in the first continuous reading of the direct or indirect speech statement in the text; its offset was defined as the temporal offset of the last fixation in the first continuous reading of the direct or indirect speech statement. Average critical fMRI event durations amounted to 3118 msec ($SD = 1175$ msec) and 3266 msec ($SD = 1209$ msec) for the direct and indirect speech conditions, respectively. The 148-msec difference in the mean durations is most likely due to of the fact that the indirect speech statements were, on average, 0.8 words longer than the direct speech statements. In fact, on a reading time per word measure, there was no appreciable difference between the two conditions (direct speech: 204 msec per word; indirect speech: 203 msec per word; $p > .5$ by paired-samples t test). Hence, when minor differences in the numbers of words are controlled for, it appears that direct and indirect speech statements were virtually identical in terms of processing difficulty.

Uncritical readings (of background sentences, comprehension questions, instructions, etc.) and events (button pressing) were specified as corresponding events in the model. The rest consisted of all fixation-cross events (including five 30-sec baseline trials and all 2-sec pretrial fixation crosses) and was regarded as baseline. The fMRI data were mapped to the human Colin atlas (sumsdb.wustl. edu/sums/) surface in Caret (brainvis.wustl.edu/wiki/index. php/Caret:About; Van Essen, Harwell, Hanlon, & Dickson, 2005; Van Essen, 2002). The mean beta estimates within ROIs were calculated by SPM toolbox easy ROI (www. sbirc.ed.ac.uk/cyril/cp_download.html) and submitted to two-tailed paired-samples t tests.

RESULTS

Answering accuracy on the posttrial comprehension questions amounted to 83% (direct speech condition) versus 82% (indirect speech condition). The 1% difference was not significant ($p > .4$ by logit binomial generalized estimating equations [GEE]; Hardin & Hilbe, 2003).

During critical events (determined via eye tracking, see Data Analysis), it was found that the direct speech condition was associated with a greater BOLD signal in voice-selective areas of the right auditory cortex than the indirect speech condition. Although both conditions were more active against the baseline, reading of direct speech elicited significantly greater activation in these areas than reading of indirect speech (see Figure 2C). For individual ROIs, the mean between-condition difference amounted to 0.240 ± 0.062 (*SE*; two-tailed paired-samples $t_{(15)} = 3.85, p = .002$); for the group ROI, the between-condition difference was 0.147 ± 0.045 ($t_{(15)} = 3.25, p = .005$). Two main clusters of enhanced activity for direct as opposed to indirect speech were located in voice-selective areas along posterior and middle parts of the right STS (Figure 2B). In addition, activation in brain areas other than the auditory cortex was found distributed in the occipital lobes, superior parietal lobules, and precuneus (Figure 2A). Although not central to our hypothesis, one might speculate that activation of those areas is part of an enriched multisensory perceptual simulation process for direct speech that also encompasses, for instance, visual aspects of the described situation. As a whole, no region showed an opposite pattern of activity, that is, direct speech was always associated with a greater BOLD signal than indirect speech. These results support the hypothesis that, during silent reading of direct speech statements, readers are more likely to engage in vivid perceptual simulations of the reported speaker's voice (or aspects thereof) than during silent reading of meaning-equivalent indirect speech statements.

DISCUSSION

Overall, our results lend objective empirical support to the intuitive experience of an "inner voice" during silent
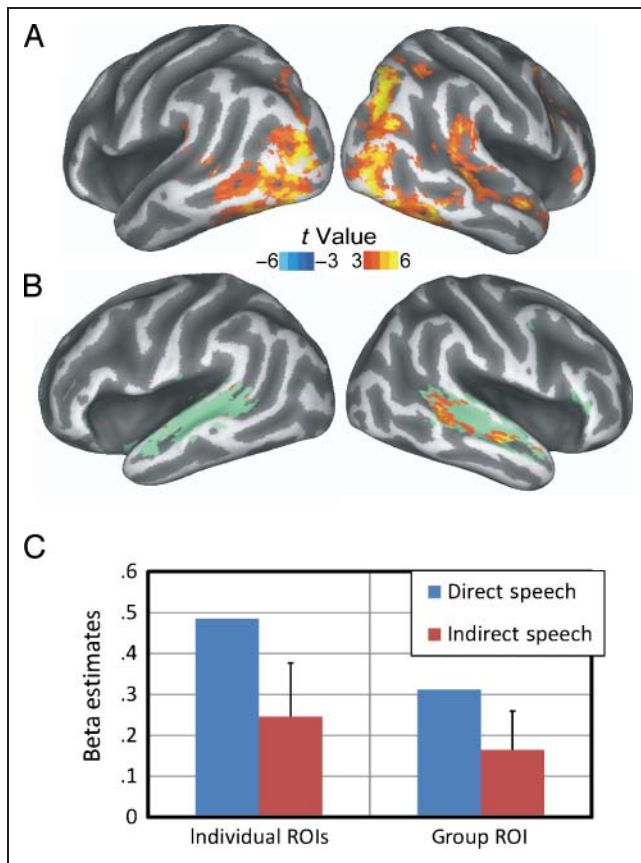
Figure 2. Illustrations of between-condition differences. (A) Regions that selectively responded to silent reading of direct as opposed to indirect speech in the whole-brain analysis. (B) Regions that selectively responded to silent reading of direct as opposed to indirect speech within the sample voice-sensitive areas (green, $n = 16$). (C) Mean signal change (against baseline) between conditions in the right voice area ROIs, determined individually (left) or using the sample average (right). The single error bar in each panel refers to the 95% confidence interval for the between-condition difference.

reading of written text, particularly during silent reading of direct speech statements. Specifically, our experiment showed that voice-selective areas in the auditory cortex become more activated during silent reading of direct speech as opposed to meaning-equivalent indirect speech statements. This finding cannot plausibly be attributed to differences in processing difficulty or comprehension performance, because there were no significant differences between the two conditions in terms of reading time per word or question-answering accuracy. Other factors, including the occasional use of direct speech exclamations or potential syntactic complexity differences in some of our direct versus indirect speech item pairs, also fail to conclusively account for the data (see www.psy.gla.ac.uk/~christop/JOCN_2011/Supplement.pdf).

Previous behavioral studies on auditory imagery (Kurby, Magliano, & Rapp, 2009; Alexander & Nygaard, 2008; Abramson, 2007) have also suggested an "inner voice" during silent reading. These studies have predominantly used an experimental setup in which particular voices are

acoustically presented to participants before the actual reading trials. In addition, participants were given explicit imagery instructions combined with cues to the identity of the to-be-imagined speaker within the actual reading materials. This type of experimental manipulation (preexposure to specific voices and/or explicit imagery cues and instructions) arguably encourages participants to imagine speaker-specific voices during silent reading. Consequently, such studies are somewhat limited in determining whether auditory imagery would also occur during "normal" language comprehension.

The current experiment focused on (a) whether silent readers activate voice-related representations even without being encouraged to do so (reading for comprehension only) and (b) whether these representations are modulated by different linguistic reporting styles (direct vs. indirect speech). A setup was used whereby no particular speaker was introduced to participants before reading the text passages and in which participants were in no way instructed to imagine voices (all they were asked to do was to read short stories and to answer questions about those stories). Each story contained a unique set of fictitious names and characters such that participants were not (or at least not obviously) led toward imagining any concrete, familiar voices. Combined with the fact that postreading comprehension questions focused on semantic content (distracting participants' attention away from the direct vs. indirect speech manipulation), it seems unlikely that our participants felt encouraged to imagine specific voices during reading. Nonetheless, they still appeared to automatically activate voice-related perceptual representations, particularly in response to reading direct speech statements.

The clearly right-lateralized brain activation pattern found in our study also suggests that voice-related representations are spontaneously activated during silent reading of direct as opposed to indirect speech. Previous studies on auditory imagery and on lip reading (Bunzeck et al., 2005; Jancke & Shah, 2004; Shergill et al., 2001; Yoo et al., 2001; MacSweeney et al., 2000; Calvert et al., 1997) mostly found bilateral (or sometimes left-dominant) top–down activation patterns within the auditory cortex. These studies employed tasks with overt metacognitive judgments (explicit imagery with or without a visual cue, shadowing, instructed rehearsing, etc.), which might recruit the left hemisphere more than would be the case in a less explicit experimental setting. The present experiment, by contrast, indicated a clearly right-lateralized locus of effect for the direct versus indirect speech comparison (Figure 2B) using a task that did not involve any metacognitive judgments.

To account for this kind of spontaneous auditory imagery during silent reading as well as its modulation through linguistic reporting style (direct vs. indirect speech), one might adopt the notion of *perceptual simulation* in language comprehension, as proposed by embodied cognition theories (Barsalou, 1999, 2008). Such theories argue that mental representations of language are grounded in perceptual experiences and actions and that perceptual

simulation (i.e., the mental reenactment of perceptual, motor, and introspective states acquired during experience with the world, body, and mind; Barsalou, 2009) is an automatic and integral part of language comprehension.

This raises the question of (a) the nature of the perceptual experiences that underlie voice simulation in response to reading direct versus indirect speech statements in text and (b) the nature of the representations that are activated during voice simulation.

Regarding the first question, we assume that accumulated experiences with direct versus indirect speech *usage* form the basis for voice simulation during silent reading. As discussed in Introduction, when speakers employ direct speech, they often mimic or dramatize aspects of the reported speaker's voice to demonstrate or depict the reported speech act; indirect speech, by contrast, is typically not used in such a vivid, demonstrative fashion, because its main function is to provide a mere description of what was said (Clark & Gerrig, 1990). Comprehension of direct speech is, therefore, more likely to be grounded in the perceptual experience of a vocal demonstration or dramatization of a reported speaker's utterance. This would explain why silent reading of direct speech is more likely to engender mental simulations of voice than silent reading of indirect speech.

The second question concerning the exact nature of the simulated voice representations is more difficult to answer at present. We conjecture that these simulated voice representations, as well as their neural correlates, overlap to a large degree with those activated during "encouraged" auditory imagery (see earlier discussion). However, compared with the speaker-specific voices that are likely to be activated during auditory imagery tasks (particularly following preexposure to concrete speech samples), the simulated voice representations reflected in the present study may be less specific, that is, they may only involve *speaker-unspecific aspects* of voice. The reason for this assumption is that our experimental setup did not encourage the imagination of speaker-specific voices. One of those speaker-unspecific aspects of voice that could underlie the present voice simulation findings is emotional prosody (suprasegmental acoustic information characterizing emotional states), which has not only been found to be associated with a right-lateralized activation pattern (cf. Wiethoff et al., 2008; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006; Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2003) but would also fit well with the notion of direct speech as vivid demonstration (the reported speaker's emotional state is often demonstrated via voice mimicking in direct speech usage). That said, the exact nature of the voice representations that are spontaneously activated during silent reading of direct versus indirect speech still remains an important question for future research.

In conclusion, our experiment showed that without being explicitly encouraged to imagine voices, readers are more likely to mentally simulate or spontaneously imagine aspects of the reported speaker's voice during silent read-

ing of direct speech as opposed to meaning-equivalent indirect speech. The results can be interpreted in line with embodied cognition (Barsalou, 1999) and the notion that direct speech is represented in a more vivid and perceptually engaging fashion than indirect speech (Clark & Gerrig, 1990). The present study is novel and important in several respects. Indeed, it is the first demonstration that top–down activation of voice-sensitive areas in the auditory cortex (Belin et al., 2000) can be modulated by linguistically (i.e., pragmatically) different reporting styles. Second, it pioneers on voice simulation during silent reading. Our results are consistent with the embodied cognition hypothesis, extending it to the auditory perceptual modality, which so far has received little attention in the relevant literature (except for studies that focused on sound-related words, see Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008; Kellenbach, Brett, & Patterson, 2001). Third, this study combined event-related fMRI with eye tracking to investigate neural correlates during on-line reading of text under relatively natural presentation conditions (e.g., contrasting with word-by-word reading); this could inspire new applications in psycholinguistic research on language comprehension, helping us to understand the kinds of mental representations activated during reading. Finally, it sheds new light on the distinction between direct and indirect speech from a cognitive neuroscience perspective, suggesting that perceptual vividness is one of the key aspects differentiating the two.

REFERENCES

Abramson, M. (2007). The written voice: Implicit memory effects of voice characteristics following silent reading and auditory presentation. *Perceptual and Motor Skills, 105,* 1171–1186.

Alexander, J. D., & Nygaard, L. C. (2008). Reading voices and hearing text: Talker-specific auditory imagery in reading. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 446–459.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577–609.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59,* 617–645.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 364,* 1281–1289.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature, 403,* 309–312.

Bohan, J., Sanford, A. J., Cochrane, S., & Sanford, A. J. S. (2008). *Direct and indirect speech modulates depth of processing*. Paper presented at the 14th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP), Cambridge, UK.

Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H. J., & Jancke, L. (2005). Scanning silence: Mental imagery of complex sounds. *Neuroimage, 26,* 1119–1127.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., et al. (1997). Activation of auditory cortex during silent lipreading. *Science, 276,* 593–596.

Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language, 66,* 764–805.

Coulmas, F. (Ed.) (1986). *Direct and indirect speech*. Berlin: Mouton de Gruyter.

Hardin, J. W., & Hilbe, J. M. (2003). *Generalized estimating equations*. London: Chapman & Hall/CRC.

Jancke, L., & Shah, N. J. (2004). "Hearing" syllables by "seeing" visual stimuli. *European Journal of Neuroscience, 19,* 2603–2608.

Kellenbach, M. L., Brett, M., & Patterson, K. (2001). Large, colorful, or noisy? Attribute- and modality-specific activations during retrieval of perceptual attribute knowledge. *Cognitive Affective and Behavioral Neuroscience, 1,* 207–221.

Kiefer, M., Sim, E. J., Herrnberger, B., Grothe, J., & Hoenig, K. (2008). The sound of concepts: Four markers for a link between auditory and conceptual brain systems. *Journal of Neuroscience, 28,* 12224–12230.

Kurby, C. A., Magliano, J. P., & Rapp, D. N. (2009). Those voices in your head: Activation of auditory images during reading. *Cognition, 112,* 457–461.

MacSweeney, M., Amaro, E., Calvert, G. A., Campbell, R., David, A. S., McGuire, P., et al. (2000). Silent speechreading in the absence of scanner noise: An event-related fMRI study. *NeuroReport, 11,* 1729–1733.

Mitchell, R. L. C., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. R. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia, 41,* 1410–1421.

Shergill, S. S., Bullmore, E. T., Brammer, M. J., Williams, S. C. R., Murray, R. M., & McGuire, P. K. (2001). A functional study of auditory verbal imagery. *Psychological Medicine, 31,* 241–253.

Spitsyna, G., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. S. (2006). Converging language streams in the human temporal lobe. *Journal of Neuroscience, 26,* 7328–7336.

Tannen, D. (1986). Introducing constructed dialogue in Greek and American conversational and literary narrative. In F. Coulmas (Ed.), *Direct and indirect speech*. Berlin: Walter de Gruyter.

Tannen, D. (1989). "Oh talking voice that is so sweet": Constructing dialogue in conversation. In D. Tannen (Ed.), *Talking voices: Repetition, dialogue, and imagery in conversational discourse* (pp. 93–133). Cambridge, MA: University Press.

Van Essen, D. C. (2002). Windows on the brain: The emerging role of atlases and databases in neuroscience. *Current Opinion in Neurobiology, 12,* 574–579.

Van Essen, D. C., Harwell, J., Hanlon, D., & Dickson, J. (Eds.) (2005). *Surface-based atlases and a database of cortical structure and function*. New Jersey: Wiley.

Wade, E., & Clark, H. H. (1993). Reproduction and demonstration in quotations. *Journal of Memory and Language, 32,* 805–819.

Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., et al. (2008). Cerebral processing of emotional prosody—Influence of acoustic parameters and arousal. *Neuroimage, 39,* 885–893.

Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. In S. Anders, G. Ende, M. Junghoffer, J. Kissler, & D. Wildgruber (Eds.), *Understanding emotions* (Vol. 156, pp. 249–268). Amsterdam: Elsevier.

Yoo, S. S., Lee, C. U., & Choi, B. G. (2001). Human brain mapping of auditory imagery: Event-related functional MRI study. *NeuroReport, 12,* 3045–3049.