Journal of Neurolinguistics xxx (2011) 1–8

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

# Functional activation for imitation of seen and heard speech

Julia R. Irwin [a,b,*], Stephen J. Frost [a], W. Einar Mencl [a], Helen Chen [a], Carol A. Fowler [a,c]

[a] Haskins Laboratories, 300 George Street, Suite 900, New Haven, CT 06511, USA
[b] Southern Connecticut State University, USA
[c] University of Connecticut, USA

ARTICLE INFO

*Article history:*
Received 21 February 2011
Received in revised form 10 May 2011
Accepted 15 May 2011

*Keywords:*
fMRI
Audiovisual
Speech perception
Speech production
Imitation
McGurk effect

ABSTRACT

This study examined fMRI activation when perceivers either passively observed or observed and imitated matched or mismatched audiovisual ("McGurk") speech stimuli. Greater activation was observed in the inferior frontal gyrus (IFG) overall for imitation than for perception of audiovisual speech and for imitation of the McGurk-type mismatched stimuli than matched audiovisual stimuli. This unique activation in the IFG during imitation of incongruent audiovisual speech may reflect activation associated with direct matching of incongruent auditory and visual stimuli or conflict between category responses. This study provides novel data about the underlying neurobiology of imitation and integration of AV speech.



Visible speech influences what listeners hear (Irwin, Whalen, & Fowler, 2006). Evidence that both visual and auditory speech signals influence perception is demonstrated in the "McGurk effect" (McGurk & MacDonald, 1976). This effect occurs when a video of a speaker's face and recording of a speaker's voice are manipulated such that they mismatch; if both face and voice are presented simultaneously the listener often reports hearing a third (non-presented) utterance that integrates the visual and auditory information (e.g. visual /da/ + auditory /ma/ are often heard as /na/). Although the auditory and visual signals are available to different perceptual modalities (and in the case of the McGurk effect, are incongruent), they are perceived as a unitary event.

* Corresponding author. Haskins Laboratories, 300 George Street, Suite 900, New Haven, CT 06511, USA. Tel.: +1 203 865 6163x311; fax: +1 203 865 8963.
*E-mail address:* Julia.Irwin@haskins.yale.edu (J.R. Irwin).

0911-6044/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jneuroling.2011.05.001

Substantial neurobiological evidence indicates that the superior temporal sulcus (STS), the superior temporal gyrus (STG) and the inferior frontal gyrus (IFG) are involved in the perception of audiovisual speech (Skipper, Nusbaum, & Small, 2005). The STS has been shown to play a role in perception of both visual (Calvert, Campbell, & Brammer, 2000) and audiovisual speech (Skipper et al., 2005). Incongruent audiovisual speech (both mismatched auditory and visual tokens and temporally asynchronous tokens) elicits activation in the STS and IFG (Olson, Gatenby, & Gore, 2002). Moreover, several neuroimaging studies have explicitly examined incongruent auditory and visual "McGurk" stimuli (Jones & Callan, 2003). For example, Olson et al. (2002) report increased activation in the left inferior frontal gyrus (IFG), including Brodmann area 44 and in bilateral STG extending to STS for mismatched audiovisual words (e.g. a visual "vet" and audio "bat", led to a percept of "vat"). Jones and Callan (2003) presented both matched and mismatched vowel-consonant-vowel (VCV) stimuli to perceivers. These stimuli, which were either synchronous, or with a 400 ms offset, led to increased activation in the left occipito-temporal junction, supramarginal gyrus (SMG), and bilateral inferior parietal lobule (IPL). Hasson, Skipper, Nusbaum, and Small (2007) found that when participants heard an auditory version of a stimulus from a given category, and were then presented with an exemplar of that category elicited from a McGurk stimulus, they showed a repetition suppression response (reduced activation). This finding indicates that the percept that results from incongruent auditory and visual speech elicits similar activation in the brain as a canonical auditory token of that category. Further, visual areas in the pars opercularis of the inferior frontal gyrus showed increased response to auditory information in audiovisual contexts than when the auditory information was presented alone, indicating particular sensitivity to the auditory signal in an audiovisual context.

Language users not only integrate phonetic information across perceptual modalities, they also produce phonetic information when they talk. Goldstein and Fowler (2003, p.159–207; see also Liberman & Whalen, 2000) note that, for communication by language to succeed, talker-listeners require a "common currency" between phonetic forms produced and perceived. Areas implicated in the production of auditory speech are activated during speech perception, indicating that the motor system is involved in auditory speech perception (Wilson, Saygin, Sereno, & Iacoboni, 2004). Moreover, a few studies have used neuroimaging to explicitly examine the areas that are activated in both perception and production of speech to identify both distinct areas and areas of overlap. Buchsbaum, Hickok, and Humphries (2001) assessed perception and production of auditory speech using fMRI. The results from that study show left posterior superior temporal gyrus activation during both perception and production, providing evidence of substantial overlap for perception and production of auditory speech in this region. Skipper, Wassenhove, Nusbaum, and Small (2007) compared activation of perception and production of consonant-vowel (CV) syllables. For the production task, the participants read the CV stimuli aloud from written text. Areas of activation, including activation in the anterior superior temporal (STa) cortex, posterior superior temporal areas (STp) and the ventral premotor cortex (PMv) for the incongruent McGurk syllable during perception were more similar to areas of activation of the congruent (called the veridical) syllable both when perceived and when spoken by the participant than activation to either the auditory or visual syllables that were combined to create it. That is, the neural response to a /ta/ percept elicited from a visual /ka/ + auditory /pa/ is more like a matched visual and auditory /ta/ than either a matched /pa/ or/ka/; further the neural response to a visual/ka/+ auditory/pa/ is also more like that of a /ta/spoken by the participant than either the /pa/ or /ka/.

From infancy, humans imitate the facial movements that they observe (Meltzoff & Moore, 1997). Meltzoff and Moore (1997) argue that there is a "supramodal representation" of stimulus input that allows them to imitate the actions of others. Early language development likely takes place in this audiovisual context, with infants imitating the productions of their caregivers. Iacoboni, Woods, Brass, and Bekkering (1999) have proposed a "direct matching" hypothesis as the basis for imitation of motor tasks, where an increase in brain activity occurs when an action is both observed and produced over when it is only produced or only observed. With regard to AV speech, Skipper et al. (2007) hypothesize that visual speech information influences heard speech via a process of "analysis-by-synthesis", in which a motor plan is recruited based on what the speaker might have been attempting to produce. In the current study, we examined the underlying neurocircuitry associated with imitation of audiovisual speech, in particular for incongruent, mismatched McGurk speech. Toward this end, we examined areas of activation for both perception of and productions elicited by imitation of audiovisually produced speech, both congruent (matched) audiovisual speech and incongruent (mismatched McGurk) speech.

Imitation was measured by contrasting two different conditions. In one condition, participants were asked to passively observe what the speaker said, and, in another, they were asked to observe and repeat what the speaker said. We examined audiovisual matched and audiovisual mismatched (McGurk) speech. The current work provides novel data about the underlying neurocircuitry involved in the perception and imitation of audiovisual speech.

## 1. Method

Because we were interested in AV integration, twenty-three participants (12 females, 11 males) completed a behavioral prescreen to identify those who were likely to integrate mismatched auditory and visual (or McGurk) stimuli. In this task, incongruent audiovisual speech stimuli were presented to listeners, an auditory /ba/ dubbed over visual /va/, which can lead to a visually influenced percept of /va/, and auditory /ma/ dubbed over visual /da/, which can lead to the visually influenced percept /na/. Participants who reported a visually influenced syllable at least 80% of the time were eligible for the functional imaging portion of the study.

### *1.1. Participants*

Seventeen participants (10 females and 7 males) met criteria in the behavioral pretest phase and received fMRI scans. Of these, data from two were eliminated due to motion and data from another participant were eliminated due to scanner malfunction. Therefore, we report data from 14 participants (7 male, 7 female), ranging in age from 18 to 44 years (mean age 26.8 years). All participants were native speakers of American English and reported normal hearing and vision.

### *1.2. Stimuli and task*

Stimuli were created from a digital video of a female native speaker of American English producing the consonant-vowel (CV) syllables /ma/, /na/, /ba/, /da/ and/va/. The stimuli were digitally edited to create four stimulus conditions: visual only stimuli (created by removing the audio track from the audiovisual stimuli), auditory only stimuli (created by removing the video track from the audiovisual stimuli), congruent audiovisual stimuli, and incongruent audiovisual stimuli. Congruent audiovisual stimuli consisted of one video token of a CV with a different audio token of the same CV (i.e. one token of a video /ma/ and a different token of an audio/ma/). The incongruent audiovisual speech stimuli were of two types, auditory /ba/ dubbed over visual /va/, which can lead to a visually influenced percept of /va/, and auditory /ma/ dubbed over visual /da/, which can lead to the visually influenced percept /na/. All stimuli were 1.6 s in duration. The video was captured and presented at 30 frames per second. The auditory stimuli were sampled at 44.1 kHz and high-pass filtered at 44 Hz to reduce ambient noise. In addition to the experimental conditions, catch trials, consisting of a silent video of the speaker with a hash mark superimposed over the lips, were also presented on 6% of trials to enforce ongoing attention to the video. When presented with these trials, participants were instructed to press a button on a response pad. All participants responded appropriately to at least 30 of the 32 catch trials.

### *1.3. fMRI procedure*

Participants received eight functional imaging runs in an event-related design, employing jittered trial durations (4, 5 s) and occasional longer null trials. Across runs, participants received 64 trials of each of the stimulus types (audio only, video only, AV congruent, AV incongruent,) and 32 catch trials as described above. During perception (perceive only) runs, participants were instructed to simply pay attention to the movie; during production (perceive and produce) runs, participants were further instructed to repeat what the speaker said. Order of perception and production runs alternated across runs with the initial run type counterbalanced across participants.

Please cite this article in press as: Irwin, J. R., et al., Functional activation for imitation of seen and heard speech, *Journal of Neurolinguistics* (2011), doi:10.1016/j.jneuroling.2011.05.001

A total of 1488 echoplanar functional images were acquired from each subject on a Siemens 3T scanner located at the Yale University Magnetic Resonance Research Center. Data acquisition and analysis followed standard procedures for univariate voxelwise analysis (c.f., Frost et al., 2009; Pugh et al., 2008); mapwise false discovery rate (FDR) correction to the $p < 0.01$ level (Genovese, Lazar, & Nichols, 2002); mapping to MNI standard space (Papademetris, Jackowski, Scultz, Staib, & Duncan, 2003, p. 788–795), and Region Of Interest (ROI) definition for the IFG and STG/STS with MRICro and PickAtlas (Lancaster et al., 2000; Maldjian, Laurienti, Kraft, & Burdette, 2003). The IFG/Broca's area region was defined within PickAtlas using the Brodmann Area 44 ROI, with dilation of two voxels to account for the extremely thin gray matter region definition. The region encompassing the posterior superior temporal gyrus and sulcus (STG/STC) was defined using the superior temporal gyrus ROI (which also includes the sulcus), and eliminating anterior voxels (MNI $y <= +22$ mm).

## 2. Results

### 2.1. Perception and production of unimodal and congruent AV speech syllables

Because our primary questions involved the congruency (mismatched or incongruent McGurk syllables vs. matched or congruently dubbed syllables) effect within the AV conditions, we only briefly summarize the simple evoked responses to the more typical unimodal auditory, unimodal visual, and congruent AV stimuli here. Unimodal auditory perception of acoustic stimuli activated the bilateral superior temporal gyrus, as well as the left lingual gyrus and precuneus. Areas activated during the perception of visual stimuli included bilateral middle occipital, fusiform, and lingual gyri, as well as portions of the inferior temporal and precentral gyri. Both congruent and incongruent AV stimuli elicited activation in the areas identified in the two unimodal conditions, including bilateral STG, bilateral fusiform gyrus, bilateral middle occipital gyrus, bilateral precentral gyrus, and bilateral cuneus.

Overt production during imitation of auditory speech showed additional significant activations (relative to the auditory perceive only condition) in the speech-motor system, including the bilateral precentral gyrus and the inferior frontal gyrus extending into the medial frontal gyrus and insula. Additional production-related activations were observed in the inferior parietal cortex, cingulate, and putamen. Overt production during visual speech produced higher activation (relative to the visual perceive only condition) in this same set of speech-motor regions, with the addition of the left medial temporal gyrus and left thalamus. Overt production during congruent AV stimuli similarly produced higher activations (relative to the AV perceive only condition) in this same general set of speech-motor regions, primarily the IFG, middle and superior frontal gyrus, and insula; but also included the bilateral inferior parietal lobule.

### 2.2. Perception and production of congruent vs. incongruent audiovisual speech stimuli

The primary goal of the current work was to assess the neural signature of imitation of matched and mismatched audiovisual speech. We employed a brain-wide analysis to broadly assess loci of McGurk effect correlates (activations for incongruent minus congruent stimuli) in the perceive only and the imitation (i.e., perceive and produce) conditions. Fig. 1A shows areas of increased activity for the incongruent stimuli in the perceive only condition. These include the posterior aspects of the superior temporal gyrus and sulcus bilaterally, the left middle temporal gyrus, and the right precentral gyrus. In the perceive and produce condition (Fig. 1B), a slightly different pattern was observed. The left superior temporal gyrus and sulcus again activated more for the incongruent stimuli, albeit in a greatly reduced area. Additionally, portions of the left inferior frontal and precentral gyrus were engaged more strongly. Finally, some areas showed higher activity levels for the congruent stimuli, including the left superior frontal gyrus, left superior occipital gyrus, and bilateral parahippocampal gyrus (not visible in this surface view).

We also queried the two defined regions of interest, IFG and STG/STS, (Calvert et al., 2000; Jones & Callan, 2003; Olson et al., 2002; Skipper et al., 2005) with specific planned comparisons and interaction tests. In the left IFG region (Fig. 2A), initial planned comparisons concur with the observations from the brain-wide analysis: in the perceive only condition, there was only a marginally significant difference between congruent and incongruent trials, $F(1,13) = 3.25$, $p = 0.09$ In the perceive and produce

Please cite this article in press as: Irwin, J. R., et al., Functional activation for imitation of seen and heard speech, *Journal of Neurolinguistics* (2011), doi:10.1016/j.jneuroling.2011.05.001

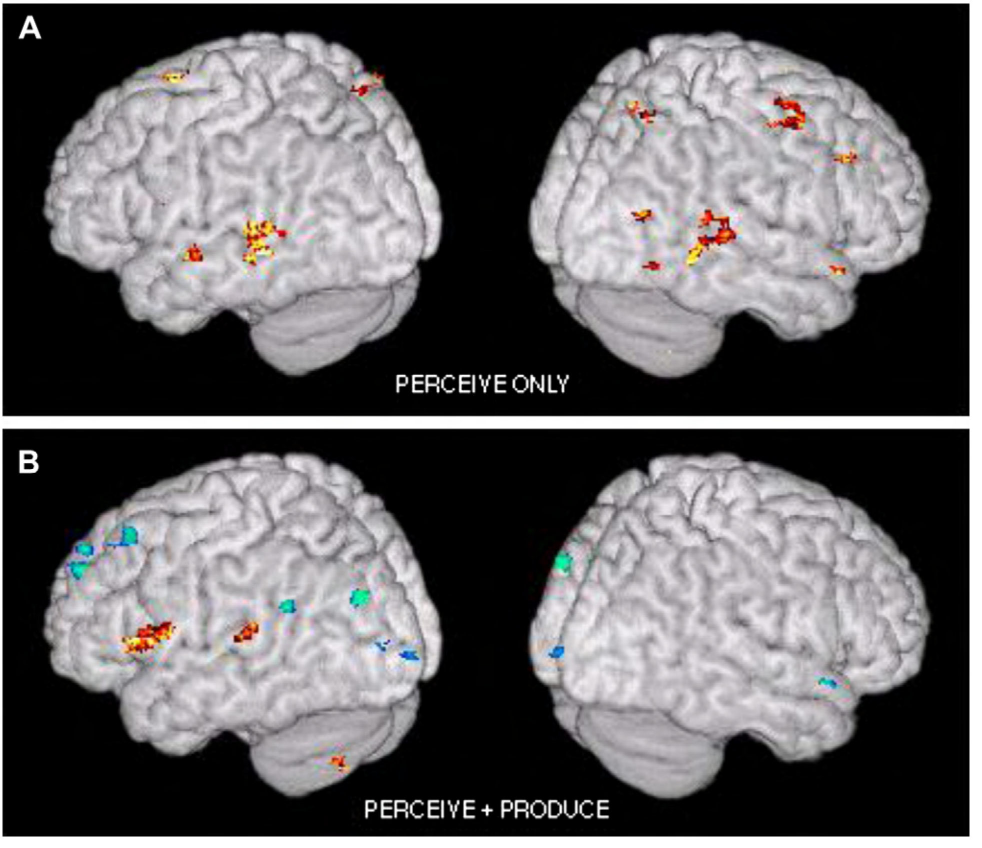


**Fig. 1.** (A, top): AV incongruent trials (McGurk trials) vs. AV congruent trials in the perceive only condition. (B, bottom): AV incongruent trials vs. AV congruent trials in the perceive and produce condition. Higher activations for incongruent trials are shown in yellow/red, and higher activations for congruent trials are shown in green/blue. Threshold is $p < 0.01$, FDR corrected, with an additional cluster threshold of 15 contiguous voxels. Activations are superimposed on a 3-D rendering of the template brain, showing views from behind the left hemisphere (left side) and right hemisphere (right side). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

condition however, incongruent stimuli elicited greater activation $F(1,13) = 9.83$, $p = 0.0079$. In the overall model, only the main effects of task and congruency were significant, with higher activation during production $F(1,13) = 21.98$, $p = 0.0004$, and higher activation for incongruent trials $F(1,13) = 10.46$, $p = 0.0065$. The direct interaction of task with congruency failed to reach significance, however ($F < 1$). In summary, the IFG exhibited higher activation for McGurk-type mismatched stimuli when production of the utterance was required, but only a marginal effect of this mismatch was observed during passive perception.

In the left STG/STS region, the opposite pattern was observed. Planned comparisons indicate higher activation during incongruent trials in the perceive only condition $F(1,13) = 6.85$, $p = 0.0213$, but no significant difference in the perceive and produce condition $F(1,13) = 2.80$, ns. Similar to the IFG however, only the main effects of task and congruency were significant in the overall model, again with higher activation during production $F(1,13) = 19.32$, $p = 0.0007$ and higher activation for incongruent stimuli $F(1,13) = 5.74$, $p = 0.0323$. The direct interaction of task and congruency again failed to reach significance ($F < 1$). For the STG/STS ROI, we observed significant increases for McGurk-type stimuli during perceive only trials, and did not observe this effect when production was additionally required. However, we did observe a smaller focal effect in the left posterior STG (see Fig. 1B); but this was not large enough to drive significance in the larger ROI.

Please cite this article in press as: Irwin, J. R., et al., Functional activation for imitation of seen and heard speech, *Journal of Neurolinguistics* (2011), doi:10.1016/j.jneuroling.2011.05.001

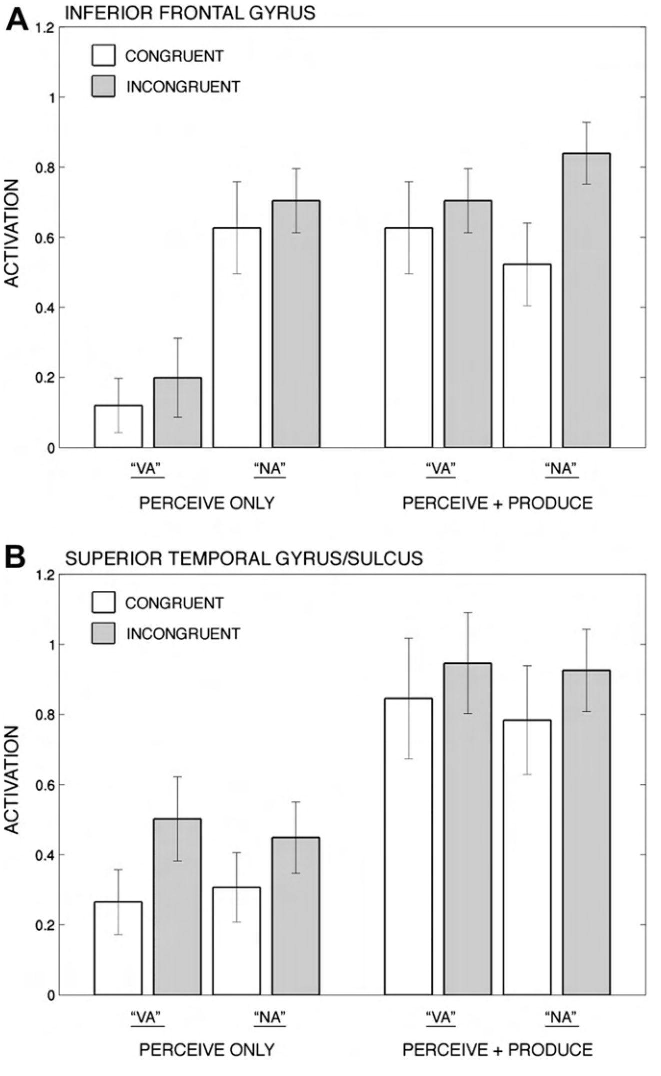


**Fig. 2.** (A, top): Mean activations, with +/− 1 standard error bars, for the eight different AV trial types in the inferior frontal gyrus ROI. (B, bottom): Mean activations and standard errors for AV trial types in the posterior superior temporal gyrus/sulcus ROI.

## 3. Discussion

A comparison of perception and production during auditory, visual, and audiovisual speech perception indicated increased activation for production in bilateral STG, bilateral precentral gyrus, and the right inferior parietal lobule (IPL). These results are consistent with previous findings indicating that brain areas involved in speech production are recruited during audiovisual speech perception (Skipper et al., 2005; Skipper et al., 2007) and that there is significant overlap in activation during perception and production of speech in the STS/STG (Buchsbaum et al., 2001). Further, within audio-visual speech perception, the STS was implicated in integration as previously reported by Calvert et al. (2000). Both the recruitment of production areas of the brain during perception and the overlap between brain areas during perception and imitation are broadly consistent with the idea of a common currency between perception and production of speech (Goldstein & Fowler, 2003, p.159–207).

Significantly different patterns of activation were shown for congruent and incongruent (McGurk) stimuli, with increased activation in the incongruent condition for production in the left precuneus, left IFG, left insula and the left SMA (supplementary motor area). In the imitation condition, incongruent stimuli elicited greater activation in the IFG. Thus, for the mismatched McGurk-type stimuli, greater IFG activation is observed when the participant was asked to imitate the CV syllable. The "direct matching" hypothesis (Iacoboni et al., 1999) predicts that an increase in brain activity will be shown when an action is both observed and produced over when it is only produced or only observed. Because we asked participants to passively observe or to imitate what the speaker said, this direct matching may elicit increased activation in particular for the McGurk stimuli because speech-motor areas must be activated for the ambiguous, mismatched auditory and visual signals (e.g. activation of both /ga/ and /ba/ for a dubbed McGurk stimulus that includes both). This finding is consistent with behavioral findings that show that McGurk stimuli lead to reduced goodness ratings and slower reaction times in comparison to matched auditory and visual syllables (Brancazio, 2004; Green & Kuhl, 1991). This could be interpreted as consistent with motor theories of speech perception (Liberman & Mattingly, 1985; Skipper et al., 2007) which posit a central role for the motor system in perception of speech (but see Lotto, Hickok, & Holt, 2009).

The current findings may also indicate that the IFG is involved in mediating competition in speech categories for audiovisual speech, as has been previously shown in auditory speech (Myers, Blumstein, Walsh, & Eliassen, 2009). For example, greater IFG activation has been observed when there is competition between competing semantic alternatives (Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). Accordingly, we observed the greatest amount of activation in IFG when participants were asked to produce what the speaker said after viewing the *mismatched* (McGurk) audiovisual stimuli. Skipper et al. (2007) hypothesize that AV speech elicits a motor plan during production of the phoneme that the speaker might have been attempting to produce, or an "analysis-by-synthesis". Because increased activation in the IFG occurred specifically in the production condition, there may be competition among elicited potential speech-motor responses (for example, auditory /na/ and visual /da/ sometimes led to /na/ and sometimes /la/ percepts). Because the inferior prefrontal cortex has been implicated in processing of between and within category differences in auditory speech stimuli (Myers et al., 2009), our stimuli may reveal this competition between possible percepts (e.g. auditory /na/ and visual /da/ could be identified as /na/ or /la/) at the level of the speech-motor actions when the listener must produce the syllables during imitation. In contrast, the passive nature of the perception task does not require the listener to choose between categories, as is required in the production task.

The current findings confirm previous research showing that the neural circuitry involved in speech production is recruited during speech perception and that there is significant overlap in activation during perception and production of speech in the STS/STG. The findings also confirm that the STS is involved in integration of audiovisual speech during perception (Buchsbaum et al., 2001; Calvert et al., 2000; Skipper et al., 2005). In addition, the results provide novel data about the underlying neurobiology of imitation and integration of AV speech. Unique activation in the IFG during imitation of incongruent audiovisual speech was revealed, which may reflect direct matching or conflict between category responses.

## Acknowledgments

This work was supported by NIH grant HD-001994 (C. Fowler, PI), DC-00403 (C. Best, PI) and DC-007339 (J. Irwin, PI) to Haskins Laboratories. Thanks go to Nicole Landi and Vince Gracco for discussions about the manuscript. Additional thanks go to Stephanie Murana and Dana Albert for assistance in preparing the manuscript.

## References

Brancazio, L. (2004). Lexical influences in audiovisual speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 30*(3), 445–463.

Buchsbaum, B. R., Hickok, G., & Humphries, C. (2001). Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science: A Multidisciplinary Journal, 25*(5), 663–678.

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology, 11*(12), 1110.

Frost, S. J., Landi, N., Mencl, W. E., Sandak, R., Fulbright, R. K., Tejada, E., et al. (2009). Phonological awareness predicts activation patterns for print and speech. *Annals of Dyslexia, 58*, 78–97.
Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage, 15*, 870–878.
Goldstein, L., & Fowler, C. (2003). *Articulatory phonology: a phonology for public language use*. Meyer A. Berlin: Mouton de Gruyter.
Green, K., & Kuhl, P. K. (1991). Integral processing of visual place and auditory voicing information during phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance, 17*(1), 278–288.
Hasson, U., Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2007). Abstract coding of audiovisual speech: beyond sensory representation. *Neuron, 56*, 1116–1126.
Iacoboni, M., Woods, R. P., Brass, M., & Bekkering, H. (1999). Cortical mechanisms of human imitation. *Science, 286*, 2526–2528.
Irwin, J., Whalen, D., & Fowler, C. (2006). A sex difference in visual influence on heard speech. *Perception & Psychophysics, 68*(4), 582–592.
Jones, J. A., & Callan, D. E. (2003). Brain activity during audiovisual speech perception: an fMRI study of the McGurk effect. *Neuroreport, 14*(8), 1129–1133.
Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., et al. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping, 10*, 121–131.
Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*, 1–36.
Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences, 4*(5), 187–196.
Lotto, A. J., Hickok, G. S., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences, 13*(3), 110–114.
Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage, 19*, 1233–1239.
McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746–748.
Meltzoff, A., & Moore, K. (1997). Explaining facial imitation: a theoretical model. *Early Development and Parenting, 6*(3–4), 179–192.
Myers, E. B., Blumstein, S., Walsh, E., & Eliassen, J. (2009). Inferior frontal regions underlie the perception of phonetic category invariance. *Psychological Science, 20*(7), 895–903.
Olson, I. R., Gatenby, J. C., & Gore, J. C. (2002). A comparison of bound and unbound audiovisual information processing in the human cerebral cortex. *Cognitive Brain Research, 14*(1), 129–138.
Papademetris, X., Jackowski, A. P., Scultz, R. T., Staib, L. H., & Duncan, J. S. (2003). *Computing 3D non- rigid brain registrations using extended robust point matching for composite multisubject fMRI analysis*. Peters TM. Berlin: Springer-Verlag.
Pugh, K. R., Frost, S. J., Sandak, R., Landi, N., Rueckl, J. G., Constable, R. T., et al. (2008). Effects of stimulus difficulty and repetition on printed word identification: an fMRI comparison of non-impaired and reading disabled adolescent cohorts. *Journal of Cognitive Neuroscience, 20*(7), 1146–1160.
Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: motor cortical activation during speech perception. *Neuroimage, 25*(1), 76–89.
Skipper, J. I., Wassenhove, V. V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex, 17*, 2387–2399.
Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of the left prefrontal cortex in retrieval of semantic knowledge: a re-evaluation. *Proceedings of the National Academy of Science, 94*, 14792–14797.
Wilson, S., Saygin, A. P., Sereno, M., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience, 7*(7), 701–702.

Please cite this article in press as: Irwin, J. R., et al., Functional activation for imitation of seen and heard speech, *Journal of Neurolinguistics* (2011), doi:10.1016/j.jneuroling.2011.05.001