Your brain says what it sees: motor mechanisms of audiovisual speech perception

Jeremy I. Skipper[1], Howard C. Nusbaum[1], Virginie van Wassenhove[2], Frederic Dick[1,3], & Steven L. Small[1]
[1]The University of Chicago, [2]The University of Maryland, College Park, [3]Birkbeck College, University of London

Introduction

▪We propose that observation of a speaker's mouth movements affects speech perception through covert production (Skipper et al., in review).

▪Thus seeing and hearing a speaker produce speech should activate the same brain regions underlying speech production.

▪These regions should be sensitive to phonetic visual information.

▪We examined the distribution of brain activity resulting from incongruent audio and visual phonetic information (the McGurk-MacDonald effect).

Methods

▪21 right handed native English speakers; event related fMRI at 3 Tesla (GE Medical Systems, Milwaukee, WI) with spiral acquisition (TR/TE 1.5/24 ms, 71° flip angle, 24 cm FOV, effective resolution 3.75 x 3.75 x 3.8mm).

▪Stimuli:

▪Congruent audiovisual (AV) /pa/, /ka/, and /ta/ and an incongruent AV stimulus (acoustic /p/ with visual /k/, designated /pk/).

▪Audio-alone (A) and video-alone (V) /pa/, /ka/, and /ta/ stimuli.

▪Imaging paradigm:

Functional scan #	Condition	Stimuli Presented	Task
Randomly ordered (1-3)	Audiovisual (AV)	/pk/, /pa/, /ka/, /ta/	Watch/listen to talker
	Audio-alone (A)	/pa/, /ka/, /ta/	Listen to talker
	Video-alone (V)	/pa/, /ka/, /ta/	Watch talker
4	Audiovisual (AV)	/pk/, /pa/, /ka/, /ta/	Watch/listen to talker
Frequency judgment: "Which stimulus did you hear most frequently in the prior run: /pa/, /ka/, or /ta/?"			
5	Active audiovisual (aAV)	/pk/, /pa/, /ka/, /ta/	Watch/listen to talker; classify syllables as /pa/, /ka/, or /ta/ with 3 alternative forced choice (3AFC)
6	Speaking (S)	/pa/, /ka/, /ta/	Read /pa/, /ka/, or /ta/ and repeat out loud

▪Behavioral measures: After scanning (1) participants indicated which consonant was heard most frequently; and (2) performed a forced choice task (see above Table).

▪Data analyses:

▪Deconvolution/regression analysis (Ward, 2001) within individuals.

▪Cortical surfaces inflated; surface-based ANOVA and conjunction/overlap analysis across individuals.

▪ROI analysis.

▪Unless otherwise stated, all imaging data is significant at $p < .05$ corrected using a false discovery rate algorithm.

Results

▪**Table 1**: Two behavioral groups differing in perception of /pk/ stimulus: (1) /ta/ perceivers (McGurk effect) and (2) /ka/ and /ta/ perceivers.

▪**Figure 1**: Congruent AV speech perception activates production areas in the pars opercularis, dorsal and ventral premotor, primary motor, subcentral gyrus and sulcus, supramarginal gyrus, and other posterior temporal sites. /pk/ activates the same regions.

▪**Figure 2**: In all frontal and temporal regions, /pk/ activation is most correlated with /ta/ activation. /pk/ was not significantly more correlated with any of the other syllables for the group who perceived both /ka/ and /ta/.

▪**Figure 3**: The time course of activation: For both groups, frontal regions correlate with perceptual reports throughout. For the /ta/ group, temporal regions show /pk/ initially correlated with audiovisual /pa/, but later with audiovisual /ta/.

▪**Figure 4**: Visual cortex activation: For the /ta/ group, visual regions (e.g., middle occipital gyrus, fusiform gyrus, calcarine sulcus) show /pk/ initially correlated with audiovisual /ka/, but later with audiovisual /ta/.

Conclusions

▪The results support a model in which speech perception is affected by information about mouth movements through the active recruitment of a motor network.

▪The activity pattern suggests that the brain has created an internal copy of the sound it *thought* it heard.

▪The results support a model of active hypothesis-testing in which motor activity simulates speech production to generate internal speech representations used for recognition (Stevens & Halle, 1967).

▪Results suggest visual cortex activity is affected by perceptual classification and auditory information.

Acknowledgments

Mucho thanks to Bernadette Brogan, Katie Barber, E Chen, Shahrina Chowdhury, Petr Hlustik, Goulven (The Frog) Josse, Robert Lyons, Daniel Margoliash, Xander Meadow, Mom, Cat Norris, David Poeppel, Ziad Saad (for much help with SUMA), Ana Solodkin, Ryan Walsh, and Nicholas (The Weem) Wymbs. Work supported by NIH RO1-DC03378

(pdf: http://www.home.uchicago.edu/~skipper/pages/posters.html)

Tables and Figures

Table 1. Behavioral results indicate two groups of participants: those who perceived /ta/ and those who perceived /ka/ or /ta/ when presented /pk/ in scanner.

	3AFC Response During Functional Scan # 5 (%)						Frequency Judgment (%)		
Group	/pa/=/pa/	/ka/=/ka/	/ta/=/ta/	/pk/=/pa/	/pk/=/ka/	/pk/=/ta/	/pa/	/ka/	/ta/
/ta/ (N=13)	97	82	96	2	15	83	23	15	62
/ka - ta/ (N=8)	100	95	94	3	61.5	35.5	12.5	62.5	25

Figure 1. A) The perception of audiovisual speech activated regions involved in speech production. Patterns for audiovisual, audio-alone, and video-alone overlapping speaking. **B) Activation pattern produced by the McGurk effect (/ta/ perception of /pk/).**

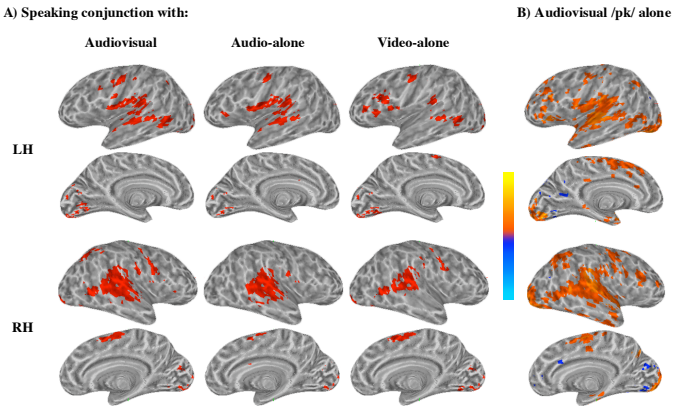


Figure 2. The distribution of activity for /pk/ in regions overlapping speech production is most correlated with the distribution of activity associated with the audiovisual syllable corresponding to the participants perception (i.e., /ta/) for the /ta/ group. Rank between /pk/ and /pa/, /ka/, and /ta/ in regions overlapping production in the two experimental groups in all **A) motor regions** and **B) temporal regions.** (*p <.008; **p<.0007)

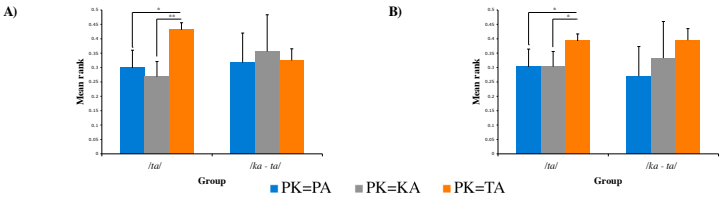
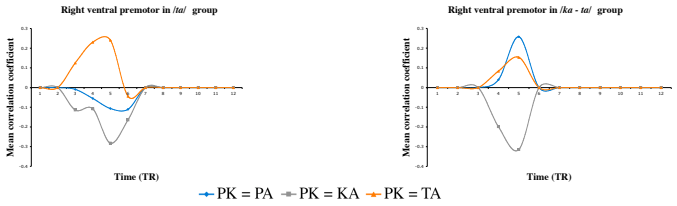


Figure 3. Temporal patterning of activation for the /ta/ group shows consistent correlation between /pk/ and /ta/ in individual regions overlapping speech production but not in the /ka - ta/ group, consistent with each group's perception. Each of the following exemplary regions show significant differences between curves as determined by nonparametric ANOVA.

A) Motor cortex:



B) "Speech perception" cortex: note how activation is initially most correlated with /pa/ (the auditory component of /pk/) but becomes increasingly correlated with /ta/ for the /ta/ group.

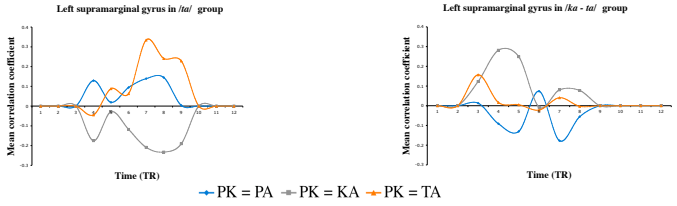


Figure 4. Temporal patterning of activation for the /ta/ group shows early correlation between /pk/ and /ka/ (the visual portion of the stimulus) but later becomes most correlated with /ta/ in primary visual cortex suggesting a feedback mechanism. Note the relative time course between Figure 3A, Figure 3B, and the below plot.