Neural Correlates of Interaural Temporal Integration in Duplex Speech Perception

Merav Sabri, Einit Liebenthal, Anjali Desai, Jain Mangalathu

Medical College of Wisconsin, Milwaukee WI, USA.

INTRODUCTION

Duplex perception is a phenomenon where a single acoustic event is perceived both as a speech syllable and as a nonspeech chip (Rand, 1974; Liberman & Mattingly, 1989).

To elicit duplex perception, a pair of stop consonant-vowel (CV) syllables (ba, ga) were split into a spectral transition portion containing the distinctive cue for syllable identification - the chip, and the remaining sound structure - the base. When presented in isolation, the base is perceived as an ambiguous speech syllable not clearly identifiable as ba or ga, whereas the chip is perceived as a nonspeech glide (rising for ba, falling for ga). When presented together, the base to one ear and the chip to the other ear, duplex perception emerges consisting of a clearly recognizable syllable together with a nonspeech chip.

The stimulus onset asynchrony (SOA) between the chip and the base modulates duplex perception. As SOA increases, speech identification accuracy decreases while chip identification accuracy increases slightly or remains constant (Bentin & Mann, 1990).

The current study employed fMRI to examine the neural mechanisms mediating the perception of complex speech-like sounds using a duplex paradigm. SOA was varied parametrically to manipulate duplex perception.

METHODS

Subjects: Twenty-four healthy, normal hearing, right-handed adults.

Stimuli:

- Duplex Stimuli: Base: Identical for ba and ga; duration = 302 ms. Chip: Rising (ba) or falling (ga); duration = 46 ms.

- Normal Speech. ba and ga; duration = 295 ms.

- Isolated Chirps. Chirp (e.g., left ear).

Paradigm:

- SOA Between chip and base was 0, 20, 40, 80 ms.

- 80 trials per condition, presented in 8 runs.

- 5 stimulus conditions: 4 duplex SOA conditions, and 1 normal condition (isolated chirp or normal speech), were randomized within each run. Ear of delivery for chirp/base was counterbalanced across runs.

- 2 tasks: In Duplex Chimp, participants identified the chip portion (rising, falling) and in Duplex Speech, they identified the whole syllable (ba or ga).

fMRI Procedures:

- GE 3T scanner (GE Medical Systems, Milwaukee).

- Functional data: T2-weighted, gradient-echo EPI (Clustered acquisition, TR = 7 s, TE = 20 ms, flip angle = 77; NEX = 1), 35 axial slices, 3.0 x 3.0 x 3.50 mm³ voxels.

- Anatomical data: 3D spoiled gradient-echo sequence. Whole brain, 0.94 x 0.94 x 1.0 mm³ voxels.

- Image analysis used AFNI software package (Cox, 1996). FWHM = 6 mm.

- Cluster size threshold (5022 p, p < .05 corrected) was applied to the group t-maps thresholded at p < .05 unless specified otherwise.

- In all images, right side = Right.

RESULTS

BEHAVIORAL RESULTS

ANOVA on the accuracy scores revealed main effects of Task and SOA as well as an Interaction:

- Accuracy (across SOAs) was significantly greater in the Duplex Chirp (92%) than in the Duplex Speech task (78%).
- Duplex Chirp Task: Accuracy was unchanged as a function of SOA.
- Duplex Speech Task: Accuracy was significantly greater for SOA 0 and 20 than for SOA 40 and 80, and for SOA 40 than for SOA 80.
- ANOVA on the RT scores revealed a main effect of SOA and an Interaction.

There was a trend of longer RT in the Duplex Speech (1025 ms) than in the Duplex Chirp task (965 ms; p < .07).

- Duplex Chirp Task: RT was unchanged as a function of SOA.
- Duplex Speech Task: RT was significantly longer for SOA 40 and 80 than for SOA 0.

FMRI RESULTS I

DUPLEX SPEECH (in Orange) – BASELINE (in Blue)

DUPLEX CHIRP (in Orange) – BASELINE (in Blue)

MAIN EFFECT OF TASK: DUPLEX SPEECH (in Orange) – DUPLEX CHIRP (in Blue).

FMRI RESULTS II

DUPLEX SPEECH 0 (in Orange) - NORMAL SPEECH (in Blue).

DISCUSSION

- Duplex speech identification was less accurate with increased SOA in the range of 20 to 80 ms, replicating the findings by Bentin and Mann, 1990. These results confirm that perception of speech was indeed disrupted by the temporal separation between the chip and base.

- In the duplex chip task, the maximal overlap at SOA 0 between F1 and F2 can potentially interfere with perception of the F2 transition cue. This task can become easier with the increase in SOA (Bentin and Mann, 1990). However, we did not observe such improvement, possibly due to ceiling effects in performance.

- Bold activation in the duplex chip compared to the duplex speech task was significantly stronger in the left intraparietal lobe, a region implicated previously in representation of discrete objects (Cusack, 2005), but also in suppression of task-irrelevant information (Wojciulik & Kanwisher, 1999). The greater signal in the left anterior gyrus, however, is due to a stronger negative activation in this region in the duplex chip compared to the duplex speech task, in line with greater involvement of the default network in less demanding tasks (Raichle et al., 2001).

- Activation in the duplex speech 0 compared with the natural speech task was stronger in the left planum temporale, suggesting its involvement in interaural integration.

- Activation in the superior temporal gyrus, bilaterally, increased proportionally to SOA only in the duplex speech task, implicating this region in temporal integration of auditory spectral information.

These results provide new insight regarding the respective roles of superior temporal regions in temporal and interaural integration across spectral bands.

References: