Tactile Communication of Speech

RLE Group
Sensory Communication Group

Sponsor
National Institutes of Health/National Institute on Deafness and Other Communication Disorders
Grant 2 R01 DC00126, Grant 1 R01 DC000117

Project Staff
Louis D. Braida, Lorraine A. Delhorne, Nathaniel I. Durlach, Theodore M. Moallem, Charlotte M. Reed, E. Courtenay Wilson

Visiting Scientists and Research Affiliates
Parivash Ranjbar

Goals and Significance

The long-term goal of this research is to develop tactual aids for persons who are profoundly deaf or deaf-blind to serve as a substitute for hearing in the reception of speech and environmental sounds. This research can contribute to improved speech reception and production, language competence, and environmental-sound recognition in such individuals. This research is also relevant to the development of improved tactual and haptic displays for a broad class of applications (including virtual-environment and teleoperator systems in addition to sensory aids).

Research over the past year has focused on two projects: (1) perceptual interactions of near-threshold stimuli in the tactual and auditory sensory modalities, and (2) evaluation of a new system for vibrotactile transduction of speech for use in speechreading aids for the deaf.

Current Studies

Perceptual Studies of the Integration of Auditory and Tactile Stimulation

Accumulating evidence from neuroanatomy, neurophysiology, imaging, and perceptual studies indicates that strong interactions occur among the senses in the way in which humans interact with the environment. This challenges the commonly held notion of specialization of the senses for their own stimulus types. Much of the ongoing work in the multimodal area has been concerned with auditory-visual and to a smaller degree visual-tactile interactions. Much less effort has been devoted to auditory-tactile interactions (e.g., Soto-Faraco and Deco, 2009).

Our previous research in this area has been concerned with investigations of the perceptual interactions between auditory and tactile stimulation in subjects with normal hearing and no known tactile deficits using detection paradigms (Wilson, 2009; Wilson et al., 2009, 2010a). In all experiments, auditory tones were presented diotically over headphones in diotic broadband noise (necessary to eliminate any potential auditory cues arising from the vibrotactile apparatus) and tactile stimuli were delivered to the left middle fingertip through a high-performance single-channel electromagnetic vibrator. The experiments were designed to measure responses to stimuli presented through each modality separately and in combination. Using models of detection, the unimodal results were compared to the bimodal results to determine whether there was evidence of facilitative responses for bimodal stimulation.

Specifically, for 500-msec, 250-Hz signals we have shown significant increases in detectability that are independent of relative auditory-tactile phase when the auditory and tactile stimuli are presented simultaneously, suggesting that the envelopes, and not the fine structure, of the two signals interact in a facilitative manner (Wilson et al., 2009). Additionally, we have also shown asymmetric changes in detectability when the two signals are presented with temporal asynchrony: when the auditory signal is presented first, detectability is not significantly greater than in A-alone or T-alone conditions, but when the tactile signal is presented first, detectability is significantly greater for almost all values of SOA employed (Wilson et al., 2009). These differences are consistent with the neural mechanics of auditory-on-auditory masking (e.g., Plack and Oxenham, 1997) and tactile-on-tactile masking (e.g., Gescheider and Migel, 1995). Our previous work has also investigated the effects of stimulating frequency in each modality on auditory-tactile integration (Wilson et al., 2010a). These results indicate that the highest rates of combined modality performance are obtained when the stimulating frequencies in the two modalities are equal or closely spaced and are within the range of the Pacinian receptor system. Furthermore, the results show evidence of critical-band filtering in the integration of auditory and tactile inputs that is broader for the tactile compared to the auditory sense.

Our results were compared with three models of integration: Optimal Signal Channel Model (OSCM), Pythogorean Sum Model (PSM), and Alegebraic Sum Model (ASM). In the phase and temporal asynchrony experiments, the measurements were more successfully modeled by the PSM approach than by the OSCM or ASM approach. In studies exploring the effects of frequency, the ASM provided a close fit to data for conditions where A and T frequencies were equal or closely spaced and within the Pacinian range; the PSM tended to model conditions with larger frequency spacings; and the OSCM provided a good fit to 50-Hz stimulation in both modalities (suggesting that integrative effects may not extend to frequency regions conveyed by non-Pacinian receptors).

Over the past year, we have continued this research through two studies: (1) effects of varying the intensity of a 250-Hz auditory or vibrotactile stimulus on detection of a combined-modality signal; and (2) effects of stimulating frequency in the auditory and tactile modalities on the loudness of suprathreshold, combined-modality signals.

Effects of Intensity Variations on Auditory-Tactile Integration

Subjects: Five subjects (age range of 21 to 39 years) with normal hearing and no-known tactile deficits participated in these experiments.

Stimuli and Block Diagram: The auditory stimuli were 250-ms pure tones presented bilaterally over headphones in a background of white noise whose level was 50 dB SPL. The tactile stimuli were sinusoidal vibrations at presented through a single-channel vibrator to the fingerpad of the left middle finger. All signals had a duration of 500 msec with 20-msec rise/fall times. In A+T conditions, the auditory and tactile signals were presented with a starting phase of 0 degrees and equal onset and offset times. A basic block diagram for the presentation of the auditory and tactile signals is provided in Figure 1.

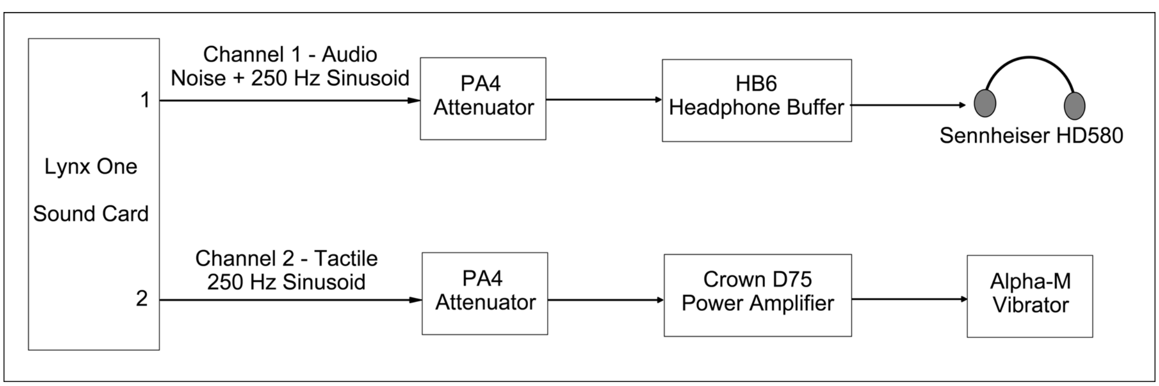

FIG 1: Block Diagram for Delivery of Auditory and Tactile Stimuli (demonstrating use of 250-Hz sinusoid in both modalities).

Test Procedures: Detection thresholds were measured under three basic conditions (Auditory alone, Tactile alone, and Auditory plus Tactile) with the goal of comparing performance on the multisensory conditions with performance measured within each of the unisensory modalities**.** Absolute-detection thresholds for the signals in each of the two unisensory modalities were first approximated using an adaptive 3-interval, 3-alternative, forced-choice procedure whose adaptive rule (1-up, 2-down) estimates the level of the signal required for 70.7% correct detection. These threshold levels were then used in performing fixed-level experiments (conducted with a 2-interval, 2-alternative forced-choice procedure with trial-by-trial correct-answer feedback) with the goal of obtaining performance in the range of 63-77%-correct performance. The resulting levels were defined as 0 dB sensation level (SL). Four intensity combinations were tested in the auditory + tactile (A+T) conditions: (1) A=(0 dB SL)+T=(0 dB SL); (2) A=(2 dB SL)+T=(0 dB SL); (3) A=(0 dB SL)+T=(2 dB SL); and (4) A=(2 dB SL)+T=(2 dB SL); Measurements of %-correct and d**'** were obtained for each of these A+T conditions as well as for the single-modality A and T stimuli that formed a particular combination.

Results and Discussion: The scores for the 2 dB SL signals in each modality were greater than those for the 0 dB SL signals. For the Auditory modality, 0 dB SL=71%-correct and 2 dB SL=81%-correct scores; for Tactile modality, 0 dB SL=72%-correct and 2 dB SL=78%-correct. For each A=T condition, the A+T score was significantly higher than the scores on the corresponding single-modality conditions. Mean percent-correct performance on the four A+T conditions described above, however, was 86.1% (Condition 1), 87.8% (Condition 2), 87.4% (Condition 3), and 90.1% (Condition 4) and showed no significant differences in performance among the conditions. This observation of no effect when the level of A and/or T was increased by 2 dB on the A+T performance was somewhat unexpected and implies a ceiling effect on performance in the A+T condition using the current methodology. We had expected to see a 5 percentage-point increase in performance with every 1-dB increase above threshold in each single modality (based on the psychometric functions measured within each modality for performance at 0 dB SL and 2 dB SL). Performance may have been limited, for example, by the subject's level of attention (i.e., this would imply less attention for more salient stimuli).

Loudness Effects at Suprathreshold Levels

The frequency tuning effects observed in the detection experiments were also examined in experiments employing supra-threshold stimuli using a loudness-matching paradigm (Wilson et al., 2010b). These experiments were motivated by auditory loudness-matching results (Zwicker et al., 1957) that show that two auditory tonal stimuli are louder when they occupy different critical bands than when they lie within the same critical band. Our A+T experiments used a matching paradigm to measure the level of an auditory probe tone as its loudness was compared with either a two-tone auditory complex or a two-tone auditory-tactile complex using tones that were either close in frequency or far apart in frequency.

Subjects: Five subjects (age range of 18 to 39 years) with normal hearing and no-known tactile deficits participated in these experiments.

Test Procedures: All stimuli were 500-ms tones with 20 ms rise/fall times whose frequencies included 200, 250, 300, and 547 Hz for auditory presentations and 20, 250, and 400 Hz for tactile presentations. Auditory masked thresholds (in the presence of a 50-dB SPL diotic broadband noise) and tactile absolute thresholds were measured at the above frequencies using a three-interval, two-alternative forced-choice procedure. A series of loudness matching experiments (in the presence of a 55-dB SPL diotic background noise) were then conducted using the two-interval, adaptive, interleaved paradigm described by Jesteadt (1980) and Silva and Florentine (2006). First, the stimuli used in the adaptive-threshold tests (auditory tones of 250, 300, and 547 Hz and tactile tones of 20, 250, and 400 Hz) were equated in loudness to a 200-Hz auditory tone at a level of 25 dB SL. Then, tones at levels equated in loudness to the 25 dB SL 200-Hz auditory tone were combined into auditory-auditory (A+A) or auditory-tactile (A+T) reference stimulus pairs. The 200-Hz auditory probe stimulus was then matched in loudness to six different pairs of stimuli: A250+A300, A250+A547, A250+T250, A547+T250, and A250+T20.

Results: The thresholds of each of the individual auditory and tactile stimuli, levels of the loudness matching of the tones to a 200-Hz 25 dB SL tone (51.6 dB SPL), and a description of the equal-loudness levels in terms of their individual sensation levels are provided in Table 1 below.

TABLE 1: The first row provides the frequencies of the individual auditory (Aud.) and tactile (Tac.) stimuli employed in the experiment. The second row provides the auditory masked thresholds and tactile absolute thresholds for these tones. The third row provides the levels for auditory-only and tactile-only stimuli when matched to a 200-Hz auditory tone presented at 25 dB SL (51.6 dB SPL). In rows two and three, the auditory measures are given in dB SPL and the tactile measures in dB re 1μm rms displacement; entries are mean levels across subjects and numbers in parentheses represent standard errors. The final row describes the equal-loudness stimuli in terms of their individual sensation levels.

	Aud.	**Aud.**	**Aud.**	**Tac.**	**Tac.**	**Tac.**
Frequency (Hz)	250	300	547	20	250	400
Threshold (dB)	25.7 (0.4)	26.9 (0.8)	27.3 (0.5)	-3.6 (1.2)	-29.2(2.6)	-18.2 (3.1)
Matching Level (dB)	49.7 (0.3)	49.2 (0.3)	47.4 (0.4)	5.1 (0.4)	-12.3 (0.9)	-6.9 (0.8
Sensation Level (dB)	24.0	22.3	20.4	8.7	16.9	11.3

Loudness-matching results averaged across 5 subjects are shown in Fig. 2. For the two auditory complexes, a 3-dB increase in the probe level (compared to the original individual-tone matches for the probe set at 51.6 dB) was required for the closely-spaced tones compared to a 4.5-dB increase for the widely-spaced tones. Presenting an A+T complex led to a 5.2-dB increase in the probe level when the two frequencies were the same but larger differences of 7-8 dB for the unequal-frequency cases. Our A-alone results are consistent with the trends reported by Zwicker et al. (1957) and our A+T results also demonstrate greater loudness effects for the unequal frequency conditions compared to the A250+T250 condition. These results indicate that the frequency-tuning effects demonstrated at near-threshold levels in the detection experiments described by Wilson et al. (2010a) can also be observed for supra-threshold stimuli using loudness-matching techniques. They are also consistent with results in the auditory-only literature concerning critical bands and suggest that the auditory and tactile systems are interacting in a frequency-specific manner similar to the interactions of auditory stimuli within the auditory system.

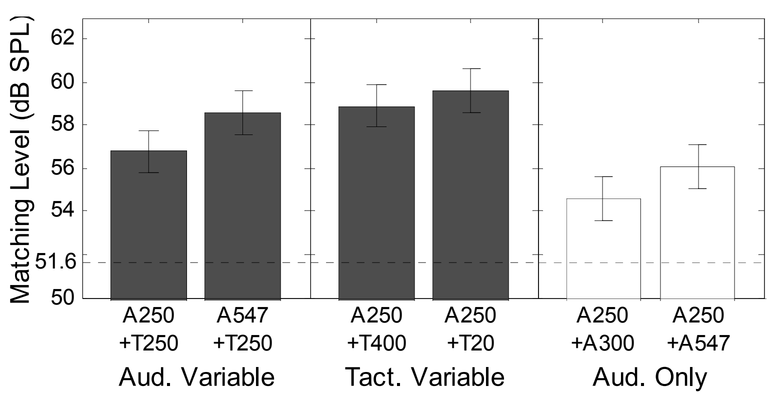

FIG 2: The average levels of the two-tone complexes when matched to the 25 dB SL 200-Hz tone. Error bars are one SEM.

Future Work

Ongoing research is being conducted on several topics in this area, including phase effects as a function of stimulating frequency as well as preliminary investigations of the effects of hearing impairment on auditory-tactile integration.

Evaluation of Vibrotactile Speech-Reception Aids for the Deaf

Over the past year, this research has resulted in a publication (Moallem et al., 2010) concerned with tactile threshold detection and tactile temporal onset and offset order discrimination in both normal-hearing and deaf subjects. In addition, work has begun on perceptual evaluations of a tactile display of consonant voicing information (Moallem, 2010). In the currently implemented display, the digitized speech signal is split into three streams, each of which is band-pass filtered (4th order Butterworth) with one of three sets of cutoff frequencies: 50-400 Hz, 800-2200 Hz, and 3000-8000 Hz. Each band is processed further before being presented to one of three channels of a tactile display that activates the thumb, index finger, and middle finger. The lowest band is passed through an additional fourth-order Butterworth 400-Hz lowpass filter, compressed in amplitude, converted into an analog voltage and presented to the middle finger. The two higher-frequency band-pass filtered signals are further processed through full-wave rectification and division into two sub-streams, one processed with a 20-Hz lowpass filter (to capture the temporal amplitude envelope) and the other with a 400-Hz lowpass filter (to capture information related to vocal fundamental frequencies) followed by amplitude compression. A weighted sum of the two sub-streams is combined within each band and converted into an analog voltage. The output of the middle-frequency band stimulates the index finger and the output of the high-frequency band stimulates the thumb.

Experimental work over the past year includes a preliminary evaluation of the ability of normal-hearing subjects to discriminate and identify speech signals presented through the vibrotactile display described above. These perceptual studies include pair-wise discrimination of pairs of English consonants that differ only in the voicing feature, as well as consonant pairs that contrast features of manner or place of articulation. Tests were conducted through conditions of lipreading alone and lipreading in conjunction with the three-channel tactile display. These results indicated that lipreading performance was close to chance for consonant pairs contrasting voicing and manner and that the addition of the tactile display led to substantial increases in the discriminability of nearly all the pairs tested. For consonants contrasting place of articulation, however, discrimination levels were high and through lipreading alone and did not improve with the addition of the tactile display. Preliminary results were also obtained on identification of 16 consonants using a one-alternative, 16-alternative forced-choice procedure under conditions of lipreading alone and lipreading plus the tactile speech display. Performance through lipreading alone averaged roughly 25% correct compared to roughly 60% correct for the combined display. These results indicate that the tactile display is capable of improving performance on the discrimination and identification of consonant segments over that obtained through lipreading alone.

Future Work

Ongoing research is concerned with the evaluation of two different approaches to training in the use of the tactile display in combination with speechreading. In the traditional approach, subjects are provided with visual correct-answer feedback during the training phase. In an experimental approach, subjects are encouraged to incorporate vocal imitation of the speech utterances presented on each trial during training. Thus, subjects will be able to compare the tactile percepts associated with the experimental stimuli to those derived from their own speech productions.

Publications

Journal Articles, Published

Israr, A., Meckl, P.H., Reed, C.M., and Tan, H.Z. (2009). "Controller Design and Consonantal Contrast Coding Using a Multi-Finger Tactual Display," Journal of the Acoustical Society of America, 125, 3925-3935.

Wilson, E.C., Reed, C.M., and Braida, L.D. (2009). "Integration of auditory and vibrotactile stimuli: Effects of phase and stimulus-onset asynchrony," Journal of the Acoustical Society of America, 126, 1960-1974.

Tan, H.Z., Reed, C.M., Durlach, N.I. (2010). "Optimum information-transfer rates for communication through haptic and other sensory modalities," IEEE Transactions on Haptics, 3, 98-108.

Wilson, E.C., Reed, C.M., and Braida, L.D. (2010a). "Integration of auditory and vibrotactile stimuli: Effects of frequency," Journal of the Acoustical Society of America, 127, 3044-3059.

Wilson, E. C., Braida, L. D., and Reed, C. M. (2010b). "Perceptual interactions in the loudness of combined auditory and vibrotactile stimuli," Journal of the Acoustical Society of America, 127, 3038-3043.

Moallem, T. M., Reed, C. M., and Braida, L. D. (2010). "Measures of tactual detection and temporal order resolution in congenitally deaf and normal-hearing adults," Journal of the Acoustical Society of America, 127, 3696-3709.

Meeting Papers, Presented

Wilson, E.C., Braida, L.D., and Reed, C.M. (2009). "Perceptual Interactions in the Detectability and Loudness of Combined Auditory-Tactile Stimuli," Poster Presentation, International Multisensory Research Forum, 10th Annual Meeting, The City College of New York, New York City, June 29-July 2, 2009.

Theses

Wilson, E. C. (2009). "Interactions between the Auditory and Vibrotactile Senses: A Study of Perceptual Effects," Ph. D. Thesis, Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA.

References

Gescheider, G.A., and Migel, N. (1995). "Some temporal parameters in vibrotactile forward masking," Journal of the Acoustical Society of America, 98, 3195-3199.

Jesteadt, W. (1980). "An adaptive procedure for subjective judgments," Perception and Psychophysics, 28, 85-88.

Moallem, T. M. (2010). "An Integrated Sensorimotor Strategy for Tactual Speech Acquisition." Proposal for Ph.D. Thesis Research, Harvard-MIT Division of Health Science and Technology, Massachusetts Institute of Technology, Cambridge.

Moallem, T. M., Reed, C. M., and Braida, L. D. (2010). "Measures of tactual detection and temporal order resolution in congenitally deaf and normal-hearing adults," Journal of the Acoustical Society of America, 127, 3696-3709.

Plack C.J., and Oxenham A.J. **(**1997**).** "Basilar-membrane nonlinearity and the growth of forward masking," Journal of the Acoustical Society of America, 103, 1598-1608.
Silva, I., & Florentine, M. (2006). "Effect of adaptive psychophysical procedure on loudness matches," Journal of the Acoustical Society of America, 120, 2124-2131.

Soto-Faraco, S., and Deco, G. (2009). "Multisensory contributions to the perception of vibrotactile events," Behavioural Brain Research, 196, 145-154.

Wilson, E. C. (2009). "Interactions between the Auditory and Vibrotactile Senses: A Study of Perceptual Effects," Ph. D. Thesis, Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA.

Wilson, E.C., Reed, C.M., and Braida, L.D. (2009). "Integration of auditory and vibrotactile stimuli: Effects of phase and stimulus-onset asynchrony," Journal of the Acoustical Society of America, 126, 1960-1974.

Wilson, E.C., Reed, C.M., and Braida, L.D. (2010a). "Integration of auditory and vibrotactile stimuli: Effects of frequency," Journal of the Acoustical Society of America, 127, 3044-3059.

Wilson, E. C., Braida, L. D., and Reed, C. M. (2010b). "Perceptual interactions in the loudness of combined auditory and vibrotactile stimuli," Journal of the Acoustical Society of America, 127, 3038-3043.

Zwicker, E., Flottorp, G., and Stevens, S. (1957). "Critical bandwidth in loudness summation," Journal of the Acoustical Society of America, 29, 548-557.