

Integration of auditory and vibrotactile stimuli: Effects of phase and stimulus-onset asynchrony

E. Courtenay Wilson, Charlotte M. Reed, and Louis D. Braida

Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 and Speech and Hearing Bioscience and Technology Program, Harvard-MIT, Division of Health Sciences and Technology, Cambridge, Massachusetts 02139

(Received 5 February 2009; revised 22 July 2009; accepted 22 July 2009)

The perceptual integration of 250 Hz, 500 ms vibrotactile and auditory tones was studied in detection experiments as a function of (1) relative phase and (2) temporal asynchrony of the tone pulses. Vibrotactile stimuli were delivered through a single-channel vibrator to the left middle fingertip and auditory stimuli were presented diotically through headphones in a background of 50 dB sound pressure level broadband noise. The vibrotactile and auditory stimulus levels used each yielded 63%–77%-correct unimodal detection performance in a 2-I, 2-AFC task. Results for combined vibrotactile and auditory detection indicated that (1) performance improved for synchronous presentation, (2) performance was not affected by the relative phase of the auditory and tactile sinusoidal stimuli, and (3) performance for non-overlapping stimuli improved only if the tactile stimulus preceded the auditory. The results are generally more consistent with a “Pythagorean Sum” model than with either an “Algebraic Sum” or an “Optimal Single-Channel” Model of perceptual integration. Thus, certain combinations of auditory and tactile signals result in significant integrative effects. The lack of phase effect suggests an envelope rather than fine-structure operation for integration. The effects of asynchronous presentation of the auditory and tactile stimuli are consistent with time constants deduced from single-modality masking experiments.

© 2009 Acoustical Society of America. [DOI: 10.1121/1.3204305]

PACS number(s): 43.66.Lj, 43.66.Wv [RLF]

Pages: 1960–1974

I. INTRODUCTION

Multisensory interactions commonly arise in everyday exploration of the environment, and numerous examples can be cited to demonstrate the influence of one sensory modality over another. For example, the presence of an auditory signal can alter judgments regarding the intensity, numerosity, and motion of visual signals (Stein *et al.*, 1996; Bhattacharya *et al.*, 2002; Sekuler *et al.*, 1997), and the location of a visual stimulus can modify the perceived location of an auditory signal (as in the ventriloquism effect; Woods and Recanzone, 2004). In the area of speech perception, for example, the McGurk effect (McGurk and MacDonald, 1976) provides a powerful demonstration of the ability of visual cues derived from lip-reading to influence the perception of auditory speech cues. The current research is concerned with exploring perceptual interactions between the senses of hearing and touch and is motivated by recent results from anatomical and physiological studies demonstrating significant interactions between these two senses.

In anatomical research, recent studies indicate that areas of the central nervous system that have traditionally been thought to receive auditory-only inputs may also receive inputs from the somatosensory system. For example, in the brainstem, the trigeminal nerve sends somatosensory input to the cochlear nucleus of the guinea pig (Zhou and Shore, 2004), while in the thalamus, somatosensory projections are sent to non-primary areas of the auditory cortex of the macaque monkey (Hackett *et al.*, 2007). Projections within the cortex have been found from the secondary somatosensory cortex to the primary auditory cortex of the marmoset

monkey (Cappe and Barone, 2005) as well as to non-primary auditory cortical areas of the macaque monkey (Smiley *et al.*, 2007). Additionally, recent physiological studies in humans (using non-invasive imaging) as well as in non-human primates (using electrophysiology) suggest that the auditory cortex is an active multisensory area, responding to somatosensory input alone as well as to combined auditory and tactile stimuli in a manner that is different from responses to auditory-only stimulation (Schroeder *et al.*, 2001; Foxe *et al.*, 2002; Fu *et al.*, 2003; Caetano and Jousmaki, 2006; Kayser *et al.*, 2005; Schurmann *et al.*, 2006; Lakatos *et al.*, 2007).

Although there is increasing anatomical and physiological evidence that tactile and auditory stimuli interact, there is less direct perceptual evidence for this interaction. Previous perceptual studies of auditory and tactile interactions can be organized into two broad categories: the influence of tactile stimulation on auditory perception and the influence of auditory stimulation on tactile perception. In the first category, experiments have shown that tactile stimuli can influence auditory localization (Caclin *et al.*, 2002) and auditory motion (Soto-Faraco *et al.*, 2004). Other perceptual studies have examined the effects of tactile stimulation on the perceived loudness or discriminability of auditory stimuli (Schurmann *et al.*, 2004; Schnupp *et al.*, 2005; Gillmeister and Eimer, 2007; Yarrow *et al.*, 2008). These studies employed a variety of experimental procedures (i.e., loudness matching, signal detectability, and signal discriminability) and, under certain

experimental conditions, have shown increased loudness or discriminability for paired auditory-tactile stimuli compared with the single-modality stimulus.

In the second category, auditory stimuli have been effective in influencing tactile perception, including such examples as changes in tactile threshold or tactile magnitude when paired with an auditory stimulus (Gescheider *et al.*, 1969; Gescheider *et al.*, 1974; Ro *et al.*, 2009). Other studies have shown that changing the high-frequency components of the auditory stimulus on a tactile task can affect the roughness judgment of the tactile stimulus (Jousmaki and Hari, 1998; Guest *et al.*, 2002) and that judgments of tactile numerosity can be affected by the presence of competing auditory signals (Bresciani *et al.*, 2005). In several of these studies (Soto-Faraco *et al.*, 2004; Bresciani *et al.*, 2005; Gillmeister and Eimer, 2007), temporal synchrony between the auditory and tactile stimuli was an important factor in eliciting interactive effects.

Further systematic and objective studies exploring the perceptual characteristics of the auditory and tactile systems are necessary for understanding the interactions between these sensory systems. In addition, perceptual studies will aid in interpreting the anatomical and neurophysiological studies which demonstrate significant interactions between the auditory and tactile sensory systems. The goal of the current research was to obtain objective measurements of auditory-tactile integration for near-threshold signals through psychophysical experiments conducted within the framework of signal-detection theory using d' (and percent correct) as a measure of detectability. Our hypothesis [derived from a general model proposed by Green (1958)] states that if the auditory and tactile systems do integrate into a common neural pathway, then the detectability of the two sensory stimuli presented simultaneously will be significantly greater than the detectability of the individual sensory stimuli. Specifically, if the stimuli are judged independently of one another, the resulting d' should equal the root-squared sum of the individual sensory d' values. If, on the other hand, the stimuli are integrated into a single percept before being processed, the resulting d' should equal the sum of the individual d' values.

The experiments reported here explore the perceptual integration between auditory pure tones and vibrotactile sinusoidal stimuli as a function of (1) phase and (2) stimulus-onset asynchrony (SOA). Manipulations of the relative phase of the tactile and auditory tonal stimuli were conducted as a means of exploring whether the interaction of the stimuli occurs at the level of the fine structure or envelope of the signals from the two separate sensory modalities. Manipulations of SOA between the tactile and auditory signals were conducted to explore the time course over which cross-modal interactions may occur. Measurements of d' (and percent correct) were obtained for auditory-alone, tactile-alone, and combined auditory-tactile presentations. The observed performance in the combined condition was then compared to predictions of multi-modal performance derived from observed measures of detectability within each of the two separate sensory modalities.

II. METHODS

A. Stimuli

The auditory stimulus employed in all experimental conditions was a 250-Hz pure tone presented in a background of pulsed 50 dB SPL (sound pressure level) Gaussian broadband noise (bandwidth of 0.1–11.0 kHz). The tactile stimulus employed in all experimental conditions was a sinusoidal vibration with a frequency of 250 Hz. The background noise was utilized to mask possible auditory cues arising from the tactile device and was present in all auditory (A), tactile (T), and combined auditory plus tactile (A+T) test conditions. The 250-Hz signals in both modalities were generated digitally (using MATLAB 7.1 software) to have a total duration of 500 ms that included 20-ms raised cosine-squared rise/fall times.

The digitized signals were played through a digital-analog sound card (Lynx Studio Lynx One) with a sampling frequency of 24 kHz and 24-bit resolution. The auditory signal was sent through channel 1 of the sound card to an attenuator (TDT PA4) and headphone buffer (TDT HB6) before being presented diotically through headphones (Sennheiser HD 580). The tactile signal was passed through channel 2 of the sound card to an attenuator (TDT PA4) and amplifier (Crown D-75) before being delivered to an electromagnetic vibrator (Alpha-M Corporation model A V-6). The subject's left middle fingertip made contact with the vibrator (0.9 cm diameter). A laser accelerometer was used to calibrate the tactile device.

B. Subjects

Eleven subjects ranging in age from 18 to 48 years (five females) participated in this study. Audiological testing was conducted on the first visit to the laboratory. Only those subjects who met the criterion of normal audiometric thresholds (20 dB hearing level or better at frequencies of 125, 250, 500, 1000, 2000, 4000, and 8000 Hz) were included in the studies. All subjects were paid an hourly wage for their participation in the experiments and signed an informed-consent document prior to entry into the study. Six subjects participated in experiment 1 (S₁, S₂, S₃, S₄, S₆, and S₇), four in experiment 2A (S₆, S₈, S₉, and S₁₀), four in experiment 2B (S₂, S₄, S₇, and S₈), and four in experiment 2C (S₈, S₁₀, S₂₁, and S₂₄). Six of the subjects participated in multiple experiments (S₂, S₄, S₆, S₇, S₈, and S₁₀).

An additional 11 subjects passed the audiometric criteria and began participation in the study but were terminated from the experiments on the basis of instability in their threshold measurements over the course of the 2-h test sessions. Further details of the criteria that were used for disqualifying a subject from continued participation in the study are provided below in Sec. II D.¹

C. Experimental conditions

The experiments examined the perceptual integration of 250-Hz sinusoidal auditory and vibrotactile signals that were each presented near the threshold of detection. Threshold measurements were first obtained under each of the two

TABLE I. Description of experimental conditions studied in Experiments 1, 2A, 2B, and 2C. In all experiments, the frequency of the auditory and tactile stimulus was always 250 Hz with duration of 500 ms. SOA was defined as $SOA = \text{Onset Time}_{\text{tactile}} - \text{Onset Time}_{\text{auditory}}$.

Condition	Starting phase (deg)		SOA (ms)	Subjects	No. of repetitions
	Auditory stimulus	Tactile stimulus			
Experiment 1: Variable studied: phase					
1-1	0	0	0	S ₁ , S ₂ , S ₃ , S ₄ , S ₆ , S ₇	6
1-2	0	90	0	S ₁ , S ₂ , S ₃ , S ₄ , S ₆ , S ₇	6
1-3	0	180	0	S ₁ , S ₂ , S ₃ , S ₄ , S ₆ , S ₇	6
1-4	0	270	0	S ₁ , S ₂ , S ₃ , S ₄ , S ₆ , S ₇	6
Experiment 2: Variable studied: SOA					
Experiment 2A: Auditory stimulus precedes tactile stimulus					
2A-1	0	0	0	S ₆ , S ₈ , S ₉ , S ₁₀	≥4
2A-2	0	0	500	S ₆ , S ₈ , S ₉ , S ₁₀	4
2A-3	0	0	550	S ₆ , S ₈ , S ₉ , S ₁₀	4
2A-4	0	0	600	S ₆ , S ₈ , S ₉ , S ₁₀	4
2A-5	0	0	650	S ₆ , S ₈ , S ₉ , S ₁₀	4
2A-6	0	0	700	S ₆ , S ₈ , S ₉ , S ₁₀	4
2A-7	0	0	750	S ₆ , S ₈ , S ₉ , S ₁₀	4
Experiment 2B: Tactile stimulus precedes auditory stimulus, no temporal overlap					
2B-1	0	0	0	S ₂ , S ₄ , S ₇ , S ₈	≥4
2B-2	0	0	-500	S ₂ , S ₄ , S ₇ , S ₈	4
2B-3	0	0	-550	S ₂ , S ₄ , S ₇ , S ₈	4
2B-4	0	0	-600	S ₂ , S ₄ , S ₇ , S ₈	4
2B-5	0	0	-650	S ₂ , S ₄ , S ₇ , S ₈	4
2B-6	0	0	-700	S ₂ , S ₄ , S ₇ , S ₈	4
2B-7	0	0	-750	S ₂ , S ₄ , S ₇ , S ₈	4
Experiment 2C: Tactile stimulus precedes auditory stimulus, with temporal overlap condition					
2C-1	0	0	0	S ₈ , S ₁₀ , S ₂₁ , S ₂₄	≥4
2C-2	0	0	-250	S ₈ , S ₁₀ , S ₂₁ , S ₂₄	4
2C-3	0	0	-500	S ₈ , S ₁₀ , S ₂₁ , S ₂₄	4
2C-4	0	0	-550	S ₈ , S ₁₀ , S ₂₁ , S ₂₄	4
2C-5	0	0	-600	S ₈ , S ₁₀ , S ₂₁ , S ₂₄	4
2C-6	0	0	-650	S ₈ , S ₁₀ , S ₂₁ , S ₂₄	4
2C-7	0	0	-700	S ₈ , S ₁₀ , S ₂₁ , S ₂₄	4
2C-8	0	0	-750	S ₈ , S ₁₀ , S ₂₁ , S ₂₄	4

single-modality conditions (A and T separately). Then the detectability of the combined auditory plus tactile (A+T) signal was measured at levels established for threshold within each of the two individual modalities. The experimental conditions examined the effects of relative phase (Experiment 1) and SOA (Experiments 2A, 2B, and 2C) of the tactile signal relative to the auditory signal.

A summary of the conditions employed in the two experiments is provided in Table I. Throughout the experiments, the stimuli were 250-Hz sinusoids of 500-ms duration (including 20-ms rise/fall times). The stimulus parameters are described in terms of the starting phase of the auditory (column 2) and tactile (column 3) stimuli and SOA (column 4). Specifically, we define SOA to be: $\text{Onset Time}_{\text{tactile}} - \text{Onset Time}_{\text{auditory}}$. Thus, the SOA is positive when the auditory stimulus precedes the tactile, 0 when the two stimuli have simultaneous onsets, and negative when the tactile stimulus precedes the auditory. Information concerning the

subjects and the number of repetitions of each experimental condition is provided in the final two columns of Table I.

Baseline condition. A baseline condition employing 0-ms SOA and starting phase of 0° for both auditory and tactile stimuli was included in each of the experiments (conditions 1-1, 2A-1, 2B-1, and 2C-1 in Table I). Performance on this baseline condition was generally measured as the first A+T condition in each test session for each subject under each of the four experiments.

Experiment 1. Experiment 1 examined the effect of the starting phase of the tactile relative to the auditory stimuli and is described in Table I (conditions 1-1 through 1-4). The auditory starting phase was always 0°, while the tactile starting phase took on four different values: 0°, 90°, 180°, and 270°. In each of these four conditions, the auditory and tactile stimuli were temporally synchronous (0-ms SOA) and thus had identical onset and offset times. This experiment was conducted on six subjects; each completed six repeti-

tions of each condition in six or seven test sessions. The order of the four experimental conditions was randomized within each replication for each subject.

Experiment 2. Experiment 2 examined the effect of asynchronous presentation of the auditory and tactile stimuli and is described in Table I (conditions 2A, 2B, and 2C). The starting phase of the auditory and tactile sinusoids was 0° throughout all of the conditions of Experiment 2. The presentation order of the experimental conditions in Experiments 2A, 2B, and 2C was randomized across sessions for each of the subjects.

In Experiment 2A (Table I, conditions 2A-1 through 2A-7), the auditory stimulus preceded the tactile stimulus with six values of SOA in the range of 500–750 ms (i.e., there was never any temporal overlap between the two stimuli). Thresholds in the baseline condition (0-ms SOA) were also measured for a total of seven conditions. Four subjects completed four replications of each of the non-zero SOA conditions, while the 0-ms SOA condition was measured at the start of each session (resulting in more than four measurements of this condition for some subjects). The number of test sessions required to complete the experiment ranged from 4 to 9 across subjects.

In Experiments 2B and 2C (Table I, conditions 2B-1 through 2B-7 and 2C-1 through 2C-8), the tactile stimulus preceded the auditory stimulus. In Experiment 2B, six values of SOA were studied in the range of –500 to –750 ms (there was no temporal overlap between the two stimuli), in addition to the baseline (0-ms SOA) condition. Four subjects^{2,3} completed four replications of each of the six non-zero SOA conditions, requiring four to nine test sessions. In Experiment 2C, in addition to the conditions described above for Experiment 2B, an SOA of –250 ms was included in order to examine the effect of partial temporal overlap between the two stimuli. Four subjects each completed four replications of the seven non-zero SOA conditions. In Experiment 2C, one subject from Experiment 2B (S_8) returned to complete four repetitions of condition 2C-2 (–250-ms SOA) and a partial subset of the remaining SOA values. Three additional subjects (S_{10} , S_{21} , and S_{24}) completed four replications of the eight experimental conditions in five to nine sessions.

D. Experimental procedures

For all experimental conditions, subjects were seated in a sound-treated booth and were presented 50-dB SPL broadband noise diotically via headphones. For testing in conditions that involved presentation of the tactile stimulus (T and A+T), the subject placed the left middle finger on a vibrator which was housed inside a wooden box for visual shielding and sound attenuation. A heating pad was placed inside the box in order to keep the box and tactile device at a constant temperature.

The following protocol was employed for testing within each experimental session: (i) Thresholds for each single-modality condition (A and T) were estimated adaptively (Levitt, 1971). (ii) Fixed-level testing was conducted for A and T separately to establish a signal level for single-

modality performance in the range of 63%–77%-correct. (iii) Fixed-level performance was measured for the baseline A+T condition (0-ms SOA, 0° phase). (iv) Fixed-level performance was measured in the experimental A+T conditions. (v) Single-modality fixed-level testing was repeated as in (ii) except with an expanded acceptable performance range of 56%–84%-correct. (Data from the second set of single-modality conditions were not otherwise used.) The number of experimental A+T conditions that could be completed within a given test session was dependent on the time required to establish signal levels that met the single-modality performance criterion.

A test session typically lasted 2 h, during which performance was measured in fixed-level experiments for the A and T conditions and A+T conditions associated with a given experiment. For each subject, three training sessions identical to the experimental sessions were provided before data were recorded. If a subject participated in multiple experiments, the three training sessions were provided only prior to the first experiment (i.e., subjects S_2 , S_4 , S_6 , S_7 , S_8 , and S_{10} underwent only three training sessions even though they participated in multiple experiments). Attention to the combined A+T stimulus was ensured by having subjects count the number of times they perceived a signal. Each experimental session lasted no more than 2 h on any given day, and subjects took frequent breaks throughout the session.

If the single-modality threshold re-tests at the end of a given session were less than 56%-correct or greater than 84%-correct (± 2 standard deviations assuming an original score of 70%-correct), the data for that session were discarded. The number of sessions discarded per subject ranged from 0 to 3 in Experiments 1 and 2A, 0–2 in Experiment 2B, and 0 in Experiment 2C. Subjects were terminated from the experiment if their scores shifted by more than 2 standard deviations in three non-training sessions, resulting in the disqualification of 11 subjects from participation in the study. Typically, disqualification resulted from increased variability in tactile threshold measurements. On average, the difference between scores measured at the beginning and end of a test session was 10.8 percentage points in the disqualified subjects compared to –0.6 percentage points in the retained subjects (with the differences in absolute values being 16.2 and 8.7 percentage points, respectively). Differences between disqualified and retained subjects were not as great for auditory scores: The corresponding differences were 4.8 and 3.0 percentage points (with the differences in absolute values being 7.5 and 6.6 percentage points, respectively).

2-I, 2-AFC fixed-level tests. The adaptive threshold estimates under the single-modality A and T conditions were employed in a 2-I, 2-AFC fixed-level procedure with 75 trials per run. Stimulus levels were adjusted and runs were repeated until scores of 63%–77%-correct were obtained. These stimulus levels were then used in testing the combined A+T conditions with the fixed-level 2-I, 2-AFC procedure.

On each presentation, the tone (auditory, tactile, or auditory-tactile) was presented with equal *a priori* probability in one of the two intervals. The interval duration was 1.15 s for Experiment 1 and 1.25 s for Experiment 2. Each inter-

val was cued by visually highlighting a push-button on the computer screen located in front of the subject. Noise was presented diotically over headphones starting 500 ms before the first interval and played continuously throughout a trial (including the durations of the two intervals and the 500-ms duration between intervals) before being turned off 500 ms after the end of the second interval. Each trial had a fixed duration of 3.8 s (Experiment 1) or 4 s (Experiment 2), plus the time it took subjects to respond. The onset of the stimulus (A, T or combined A+T) was always coincident with the onset of the observation interval in which it appeared. Subjects responded between trials by selecting the interval in which they thought the stimulus was presented (using either a mouse or keyboard) and were provided with visual correct-answer feedback.

E. Data analysis

A two-by-two stimulus-response confusion matrix was constructed for each 75-trial experimental run and was used to determine percent-correct scores and signal-detection measures of sensitivity (d'). These measures were averaged across the repetitions of each experimental condition within a given subject. Statistical tests performed on the data included analyses of variance (ANOVAs) on the arcsine transformed percent-correct scores, with statistical significance level defined for probability (p -values) less than or equal to 0.01. For statistically significant effects, a *post hoc* Tukey–Kramer analysis was performed with $\alpha=0.05$.

F. Models of integration

The results of the experiments were compared with three different models of integration: the Optimal single-channel model (OSCM), the Pythagorean sum model (PSM), and the algebraic sum model (ASM). The OSCM assumes that the observers' responses are based on the better of the tactile or auditory input channels. The predicted D'_{OSCM} ⁴ for the combined A+T condition is the greater of the tactile (d'_T) or auditory (d'_A), $D'_{OSCM}=\max(d'_T, d'_A)$. The PSM assumes that integration occurs across channels (e.g., as in audio-visual integration, [Braidá, 1991](#)) and that the d' in the combined auditory-tactile condition is the Pythagorean sum of the d -primes for the separate channels, $D'_{PSM}=\sqrt{d'^2_A+d'^2_T}$. The ASM, on the other hand, assumes that integration occurs within a given channel and that the combined d' is the linear sum of the d -primes for the separate channels, $D'_{ASM}=d'_A+d'_T$. For example, if the auditory d'_A was 1.0 (69%-correct) and the tactile d'_T was 0.8 (66%-correct), the OSCM would predict a D'_{OSCM} of 1.0 (69%-correct), the PSM would predict a D'_{PSM} of 1.28 (74%-correct), and the ASM would predict a D'_{ASM} of 1.8 (82%-correct). The OSCM prediction is never greater than the PSM prediction, which in turn is never greater than the prediction of the ASM.

Chi-squared goodness-of-fit calculations were employed to compare observed with predicted values from each of the three models. The predictions of the models were evaluated as follows: First, d -prime values were determined for each auditory (d'_A) and tactile (d'_T) experiment on the basis of 75 total trials. Second, predicted d -prime values were computed

for the three models according to the formulas given above. Third, predicted percent correct scores were computed for each of the models in the following manner: Percent Correct= $100\phi(D'_{A+T}/2)$, where ϕ is the cumulative of the Gaussian distribution function and D'_{A+T} is the predicted D' . Fourth, the observed A+T confusion matrix was analyzed to estimate d'_{A+T} and the “no bias” estimate of percent correct score was computed as $\%_{A+T}=100\phi(d'_{A+T}/2)$. This relatively small adjustment (1.6 percentage points on average, 13 points maximum) was necessary because the predictions of the models assumed that the observer is not biased. Predictions (D'_{OSCM} , D'_{PSM} , and D'_{ASM} or $\%_{OSCM}$, $\%_{PSM}$, and $\%_{ASM}$) were compared with observations (d'_{A+T} or $\%_{A+T}$). The proportion of the observations that agreed with predictions was judged by having a chi-squared value less than 3.841 (the 95% criterion) between predicted and observed scores (corrected as discussed above) using a contingency table analysis ([Neville and Kennedy, 1964](#)). This analysis allows for errors in both the observed score and the predicted score.

III. RESULTS

A. Signal levels employed in single-modality conditions

Single-modality auditory and tactile thresholds were obtained both at the beginning and at the ending of each individual test session. The data reported here, however, are based solely on the initial measurements. Analyses that used the average of the beginning and ending single-modality measurements were not significantly different from these. Thus, we used the post-experiment measurements merely as a tool for determining threshold stability.

Levels for auditory-alone conditions. The mean signal levels in dB SPL established for performance in the range of 63%–77%-correct for a 250-Hz tone in 50-dB SPL broadband noise are shown in the upper panel of [Fig. 1](#). Mean levels of the tone are plotted for each individual subject in each of the four experiments. Each data point depicted in the plot is based on an average of at least 4 and as many as 11 measurements in the fixed-level 2-I, 2-AFC procedure (each of which yielded performance in the range of 63%–77%-correct). Ten of the 11 subjects had average auditory masked thresholds within a 2.1-dB range of 22.3–24.4 dB SPL. The remaining subject (S_{10}) had a value of 27.8 dB SPL, measured consistently across multiple sessions. Within a given subject, tonal levels were highly stable for measurements made within a given experiment and across experiments. Values of ± 2 standard error of the mean (SEM) (accounting for 96% of the measurements) ranged from 0.095 to 1.1 dB across subjects and experiments.

These results are consistent with those obtained in previous studies of tonal detection in broadband noise. Critical ratios were calculated for the tone-in-noise levels shown in [Fig. 1](#) by subtracting the spectrum level of the noise at 250 Hz (which was 7.4 dB/Hz) from the presentation levels of the 250-Hz tone. Across subjects and experiments, mean critical ratios ranged from 14.9 to 20.4 dB and are consistent with the critical ratio value of 16.5 dB at 250 Hz reported by

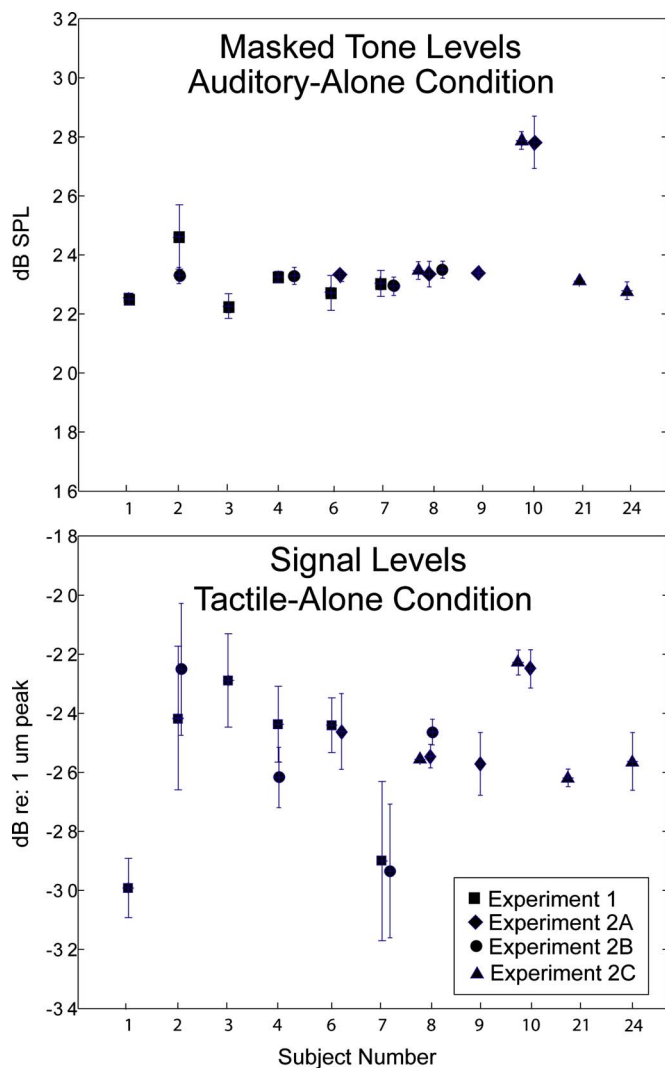


FIG. 1. (Color online) Single-modality signal levels employed for individual subjects tested in each of the four experiments. Auditory levels are for detection of a 250-Hz pure tone in 50-dB SPL broadband noise. Tactile levels are for detection of a 250-Hz sinusoidal vibration presented to the fingertip. Different symbols represent results obtained in different experiments. Some subjects participated in more than one experiment. Error bars are 2 SEM.

Hawkins and Stevens (1950). Thus, these results indicate that subjects were listening to the auditory tones in noise at levels that were close to masked threshold.

Levels for tactile-alone conditions. The mean signal levels established for performance in the range of 63%–77% correct for a 250-Hz sinusoidal vibration to the left middle fingertip are shown in the lower panel of Fig. 1. All threshold measurements were obtained in the presence of a diotic 50-dB SPL broadband noise presented over headphones. Signal levels are plotted in decibel *re* 1 μm peak displacement for individual subjects who participated in each of the four experiments. Each mean level is based on 4–11 measurements across individual subjects and experiments. Average signal levels employed in the tactile-alone conditions ranged from -30 to -22 dB *re* 1 μm peak. Within-subject values of ± 2 SEM (accounting for 96% of the measurements) ranged from 0 to 2.4 dB across subjects and experiments. The Appendix discusses the unlikely possibility that

the tactile stimulus was detected auditorally via bone conduction.

The signal levels employed for the tactile-alone conditions are generally consistent with previous results in the literature for vibrotactile thresholds at 250 Hz obtained using vibrators with contactor areas similar to that of the device employed in the present study (roughly 80 mm^2). Investigators using contactor areas in the range of 28–150 mm^2 have reported mean thresholds in the range of -21 to -32 dB *re* 1 μm peak (Verrillo *et al.*, 1983; Lamore *et al.*, 1986; Rabinowitz *et al.*, 1987).

B. Baseline experiment

Results from the baseline experiment are shown for individual subjects in Experiments 1, 2A, 2B, and 2C in the four panels of Fig. 2. The mean percent correct scores with error bars depicting ± 2 SEM are plotted for the three conditions of A-alone, T-alone, and A+T (SOA=0 ms, phase = 0° ; see Table I, experimental conditions 1-1, 2A-1, 2B-1, and 2C-1) for each subject within each experiment. Averages across subjects are provided as the rightmost data bars within each panel. Across the four experiments, there is a substantial increase in the percent correct score when the auditory and tactile stimuli are presented simultaneously compared with the A- and T-alone conditions. Averaged over subjects, the results indicate that scores for the A-alone and T-alone condition were similar (ranging from 67.8% to 74.9% correct across experiments) and lower than the average scores in the A+T conditions (which ranged from 75.2% to 88.8% correct across the four experiments). Variability was generally low, with values of ± 2 SEM ranging from 0.6 to 15.1 percentage points across subjects and experiments with all but one subject less than 7 percentage points.

A two-way ANOVA was performed on the results of the baseline experiment to examine the main effects of Condition (A, T, A+T) and Subject (11 different subjects across experiments). These results indicate a significant main effect for Condition [$F(2, 257) = 91.44, p < 0.01$] but not Subject [$F(10, 257) = 1.00, p = 0.035$], and a significant effect for their interaction [$F(20, 257) = 2.8, p < 0.01$]. A *post hoc* analysis of the main effect of Condition showed that scores on the A+T condition were significantly greater than on the A-alone and T-alone conditions and that the A-alone and T-alone conditions were not significantly different from one another. A *post hoc* analysis of the Condition by Subject effect indicated that all subjects were similar on the A-alone and T-alone conditions but different on the A+T condition. Specifically, of the 11 subjects tested, 8 had a significantly higher A+T score compared with the A-alone and T-alone scores; two subjects showed no significant increase in score (S_8 and S_{24}); and one subject (S_{21}) had significantly greater A+T scores compared to either A-alone or T-alone, but not to both.

C. Experiment 1: Effects of relative auditory-tactile phase

The results of Experiment 1 are shown in Fig. 3. Percent-correct scores averaged across six subjects and six repetitions per condition are shown for each of the six ex-

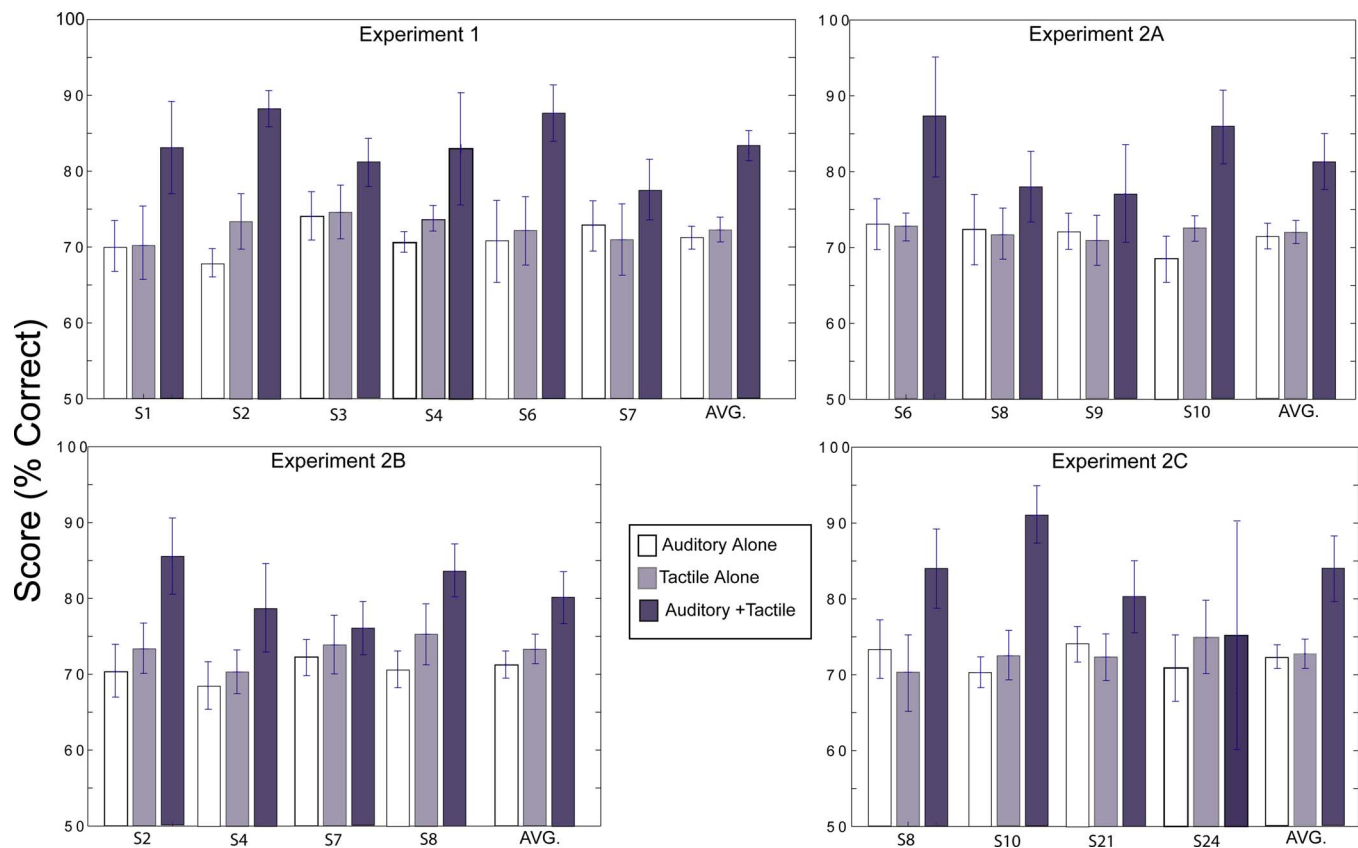


FIG. 2. (Color online) Summary of results for baseline condition in Experiments 1, 2A, 2B, and 2C. Percent correct scores for the individual subjects in each experiment are averaged across multiple repetitions per condition; number of repetitions varies by subject and is equal to or greater than 4 per subject. AVG is an average across subjects and repetitions in each experiment. White bars represent A-alone conditions, gray bars represent T-alone conditions, and black bars represent the A+T baseline condition with SOA=0 ms and phase=0°. Error bars are 2 SEM.

perimental conditions: A-alone, T-alone, and combined A+T with four different values of the starting phase of the tactile stimulus relative to that of the auditory stimulus (0°, 90°, 180°, and 270°). Average scores were 71.2%-correct for A-alone, 72.2%-correct for T-alone, and ranged from 83.2%-to 84.6%-correct across the four combined A+T conditions.

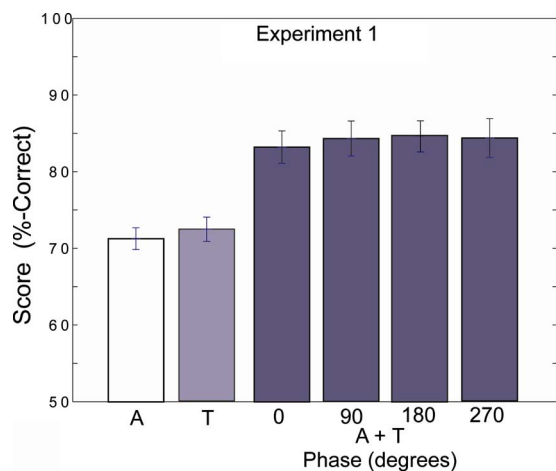


FIG. 3. (Color online) Summary of results for experiment 1. Percent correct scores are averaged across six subjects with six sessions per condition. Scores are shown for A-alone (white bar), T-alone (light gray bar), and combined A+T condition (dark gray bars) as a function of starting phase (in degrees) of the tactile stimulus relative to the auditory stimulus. Error bars are 2 SEM.

The Appendix discusses the unlikely possibility that this variation was caused by a bone-conducted interaction between the tactile and auditory stimuli. Variability in terms of ± 2 SEM ranged from 2.0 to 2.5 percentage points across the four phase conditions.

A two-way ANOVA was performed with main factors of Condition (A, T, A+T: tactile phase) and Subject. The results of the ANOVA indicate a significant main effect for factors of Condition [$F(5, 192)=44.93, p<0.01$] and Subject [$F(5, 192)=4.01, p<0.01$] but not for their interaction [$F(25, 192)=1.61, p=0.04$]. The *post hoc* analysis on Condition indicated that scores on the A-alone and T-alone conditions were not significantly different from one another, that scores for the four A+T combined conditions were not significantly different from one another, and that the scores for each of the four A+T conditions were significantly greater than the A- and T-alone scores. The *post hoc* analysis on subject indicated that the A+T scores for S_6 were significantly greater than those of S_1 and S_7 , and that the scores for S_2 were significantly greater than those of S_7 .

D. Experiment 2: Effects of SOA

Experiment 2 explored the effect of SOA between the auditory and tactile stimuli in three different experiments. Experiment 2A tested conditions in which the auditory stimulus preceded the tactile stimulus, and Experiments 2B and 2C tested conditions in which the tactile stimulus pre-

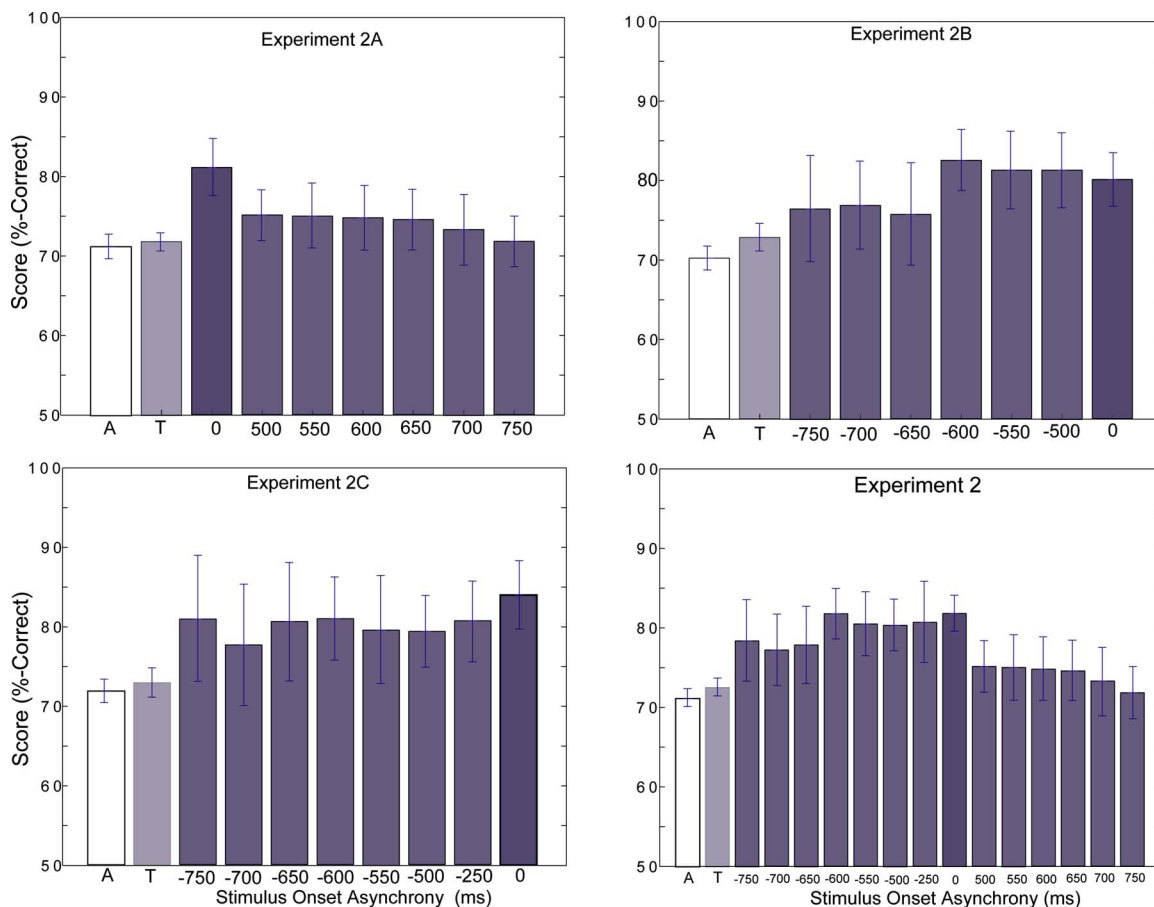


FIG. 4. (Color online) Summary of results for Experiment 2. In all panels, scores are shown for A-alone (white bars), T-alone (light gray bars), and combined A+T condition (dark gray bars). In the upper left panel (Experiment 2A: auditory precedes tactile), percent correct scores are averaged across four subjects with four sessions per condition (SOA=0 ms has more than four repetitions). In the upper right panel (Experiment 2B: tactile precedes auditory), percent correct scores are averaged across four subjects with four sessions per condition (SOA=0 ms has more than four repetitions). In the lower left panel (Experiment 2C: tactile precedes auditory, with temporal overlap), percent correct scores are averaged across four subjects with four sessions per condition (SOA=0 ms has more than four repetitions). The lower right panel (Experiment 2) provides a composite summary of percent correct scores averaged across all subjects and repetitions from Experiments 2A, 2B, and 2C. In all panels, error bars are 2 SEM.

ceded the auditory stimulus. Percent-correct scores averaged across four subjects and four repetitions of each non-zero SOA condition in each of these experiments are shown in Fig. 4. Error bars represent ± 2 SEM.

In Experiment 2A (Fig. 4, upper left panel), scores for the A-alone and T-alone conditions averaged 71.1%- and 71.8%-correct, respectively. For the combined A+T conditions, average scores of the non-zero SOA conditions ranged from 71.8%-correct (SOA=500 ms) to 75.1%-correct (SOA=750 ms). Variability, in terms of ± 2 SEM, ranged from 3.2 percentage points (SOA=500 ms) to 4.3 percentage points (SOA=700 ms). A two-way ANOVA was conducted using main factors of Condition (A, T, and the seven combined A+T conditions with different values of SOA) and Subject. The results of the ANOVA indicate that both main factors (Condition: [$F(8, 156)=6.16, p<0.01$]; subject: [$F(3, 156)=19.32, p<0.01$]), as well as their interaction [$F(24, 156)=2.3, p<0.01$], were significant. The *post hoc* analysis revealed that only one A+T combined condition, that of SOA=0 ms (i.e., the baseline condition), produced a score that was significantly greater than the A-alone or T-alone score. The scores for the remaining SOA conditions were not significantly greater than the scores in the A-alone

or T-alone conditions. The *post hoc* analysis of the subject effect indicated that the scores for S_{10} were significantly different from those of the other three subjects. For the interaction effect, S_{10} showed significantly greater A+T scores at all SOA's except 750 ms compared with A- and T-alone, while none of the other subjects showed a significant difference between non-zero SOA and A-alone and T-alone scores.

In Experiment 2B (Fig. 4, upper right panel), scores for the A-alone and T-alone conditions averaged 70.5%- and 73.3%-correct, respectively. For the combined A+T conditions, averaged scores of the non-zero SOA conditions ranged from 75.7%-correct (SOA=-650 ms) to 82.5%-correct (SOA=-600 ms). Variability in terms of ± 2 SEM ranged from 3.8 percentage points (SOA=-600 ms) to 6.7 points (SOA=-750 ms). The results of a two-way ANOVA indicated that the two main effects of condition and subject were both significant (Condition: [$F(8, 139)=6.6, p<0.01$]; Subject: [$F(3, 139)=14.76, p<0.01$]), but not their interaction [$F(24, 139)=1.77, p=0.02$]. A *post hoc* analysis indicated that scores on the combined A+T conditions with SOA values of 0, -500, -550, and -600 ms were significantly greater than scores on the A-alone and T-alone conditions. Scores on the combined A+T conditions with SOA

TABLE II. Chi-squared tests: predicted vs observed. This table enumerates the number of observations that have passed/failed the chi-squared goodness-of-fit test for each of the three models (i.e., optimal single channel, Pythagorean sum, and algebraic sum).

Experiment	Condition	Total	Optimal single channel			Pythagorean sum			Algebraic sum		
			Pass	Fail	Under-predict, fail	Pass	Fail	Under-predict, fail	Pass	Fail	Under-predict, fail
Baseline											
1	Phase=0°	40	26	14	14	33	7	7	27	13	3
2A	SOA=0 ms	26	17	9	7	20	6	3	14	12	2
2B & 2C	SOA=0 ms	37	20	17	15	29	8	6	24	13	3
Totals		103	63 (61%)	40	36 (35%)	82 (80%)	21	16 (16%)	65 (63%)	38	8 (8%)
Phase											
1	0°	40	26	14	14	33	7	7	27	13	3
1	90°	36	24	12	12	29	7	7	31	5	2
1	180°	37	25	12	12	33	4	4	29	8	1
1	270°	35	21	14	14	30	5	5	32	3	2
Totals		148	96 (65%)	52	52 (35%)	125 (84%)	24	23 (16%)	119 (80%)	29	7 (5%)
SOA											
2A	500 ms	19	18	1	1	18	1	1	10	9	1
2A	550 ms	18	15	3	3	15	3	0	9	9	0
2A	600 ms	19	15	4	1	15	4	0	10	9	0
2A	650 ms	18	17	1	1	15	3	1	7	11	1
2A	700 ms	18	14	4	2	11	7	0	8	10	0
2A	750 ms	18	16	2	1	15	3	0	6	12	0
Totals		110	95 (86%)	15	9 (8%)	89 (81%)	21	2 (2%)	50 (45%)	60	2 (2%)
SOA											
2B & 2C	-250 ms	17	13	4	4	11	6	3	11	6	2
2B & 2C	-500 ms	33	20	13	12	21	12	9	22	11	2
2B & 2C	-550 ms	31	21	10	9	19	12	7	19	12	3
2B & 2C	-600 ms	36	23	13	12	25	11	7	23	13	4
2B & 2C	-650 ms	28	16	12	8	17	11	6	18	10	1
2B & 2C	-700 ms	29	14	15	11	18	11	4	15	14	1
2B & 2C	-750 ms	28	15	13	9	16	12	6	16	12	2
Totals		202	122 (60%)	80	65 (32%)	127 (63%)	75	42 (21%)	124 (61%)	78	15 (7%)

values of -650 , -700 , and -750 ms, on the other hand, were not significantly different from A-alone and T-alone scores. A *post hoc* analysis of the subject effect indicated that three of the four subjects demonstrated the main trends for condition described above.

The results of Experiment 2C (Fig. 4, lower left panel) were similar to those found in Experiment 2B. Average scores for the A-alone and T-alone conditions were 71.9%- and 72.7%-correct, respectively. Average scores on the combined A+T conditions ranged from 77%-correct (SOA = -700 ms) to 81%-correct (SOA = -600 and -750 ms). Variability in terms of ± 2 SEM ranged from 4.5 percentage points (SOA = -500 ms) to 8 points (SOA = -750 ms). A two-way ANOVA with main factors of condition and subject indicated significant effects for both (Condition: $[F(9, 102) = 10.6, p < 0.01]$; Subject: $[F(2, 102) = 91.57, p < 0.01]$), as well as for their interaction $[F(18, 102) = 4.69, p < 0.01]$. A *post hoc* analysis of the condition effect indicated that the scores in the combined A+T conditions for every value of SOA were significantly higher than scores on the A-alone and T-alone conditions. A *post hoc* analysis of the subject effect showed that scores from all subjects tested were significantly different from one another. The response pattern

for S_{10} as a function of condition differed from that of the other three subjects.

E. Comparisons to model predictions

Chi-squared goodness-of-fit tests were performed in order to examine which model, the OSCM, the PSM or the ASM, best fits the measured percent correct scores (Sec. II F). The proportion of observations in agreement with predictions, i.e., having a chi-squared value less than 3.841, is summarized in Table II and also shown in Fig. 5.

The baseline condition (synchronous presentation, 0° tactile-auditory phase; Fig. 5, top row) was included in all testing sessions and involved 103 comparisons. Of these, 63 (61%) of the predictions agreed with the OSCM, 82 (80%) with the PSM, and 65 (63%) with the ASM. All three models failed a simple binomial test for symmetry of error.

The results of the four phases of Experiment 1 had similar proportions in agreement with the predictions of the PSM and ASM, indicating again that relative auditory-tactile phase had no effect on integration. The middle three panels of Fig. 5 show the predicted vs observed for all four phases grouped together. Out of a total of 148 observations, 96

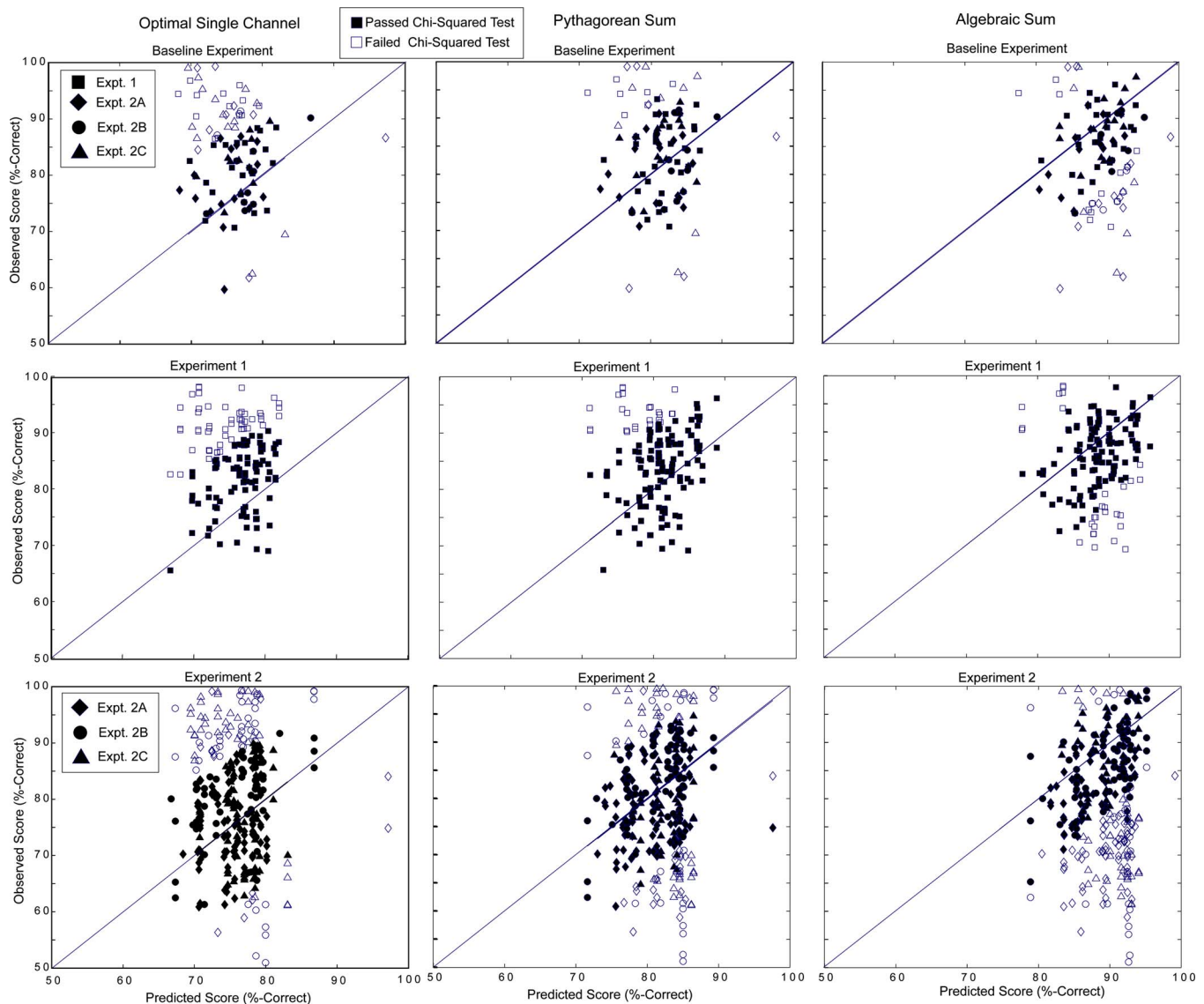


FIG. 5. (Color online) Predicted vs Observed values for the three models of integration: OSCM (far left column), PSM (middle column), and ASM (right column). The first row shows all values from the baseline experiment (SOA=0 ms, phase=0°); data from each experiment are designated by a different shape (see legend). The second row shows values from all phases in experiment 1 (relative phase). The third row shows all non-zero SOA values from experiment 2; each sub-experiment delineated by shape (see legend). Open symbols indicate that the observed value failed the chi-squared test, and filled symbols indicate that the observed value passed the chi-squared test.

(65%) agreed with the OSCM; 125 (85%) agreed with the PSM; and 119 (80%) agreed with the ASM. It can be seen that most of the data points that do not satisfy the chi-squared test are higher than the predictions of the OSCM and PSM (middle left and center panels, respectively) and lower than those of the ASM (middle right panel). The OSCM failed the symmetry test for all four phases, the PSM passed only 0 and 90°, and the ASM passed only 90°.

Discussion of the results of Experiment 2 (SOA, Fig. 5, bottom row) will be restricted to non-zero SOA because the case of zero SOA was considered above (baseline). The bottom three panels in Fig. 5 compare observed and predicted scores in Experiment 2, segregated by sub-experiment (i.e., diamond symbols represent Experiment 2A, circles are experiment 2B, and triangles are Experiment 2C). The OSCM (lower left panel) tends to under-predict the observed scores, the PSM (lower center panel) tends to over- and under-

predict to a roughly equal degree, while the ASM (lower right panel) tends to over-predict scores. Table II enumerates the results of Experiment 2A separately and groups the results of Experiments 2B and C together.

For Experiment 2A, the symmetry test was performed for each model and on all of the non-zero SOA values. The OSCM passed all six non-zero SOA values; the PSM passed all non-zero SOA values except 750 ms; and the ASM failed all non-zero SOA values. The results of a chi-squared test showed that the observed and predicted scores agreed 95 out of 110 times (86%) for the OSCM, and 89 (81%) for the PSM, while only 50 (45%) agreed with the ASM. Of the cases that did not pass the chi-squared test, the OSCM produced roughly an equal number of under- (9) and over-predictions (6), while nearly all errors were over-predictions for the PSM and ASM models.

In Experiments 2B and 2C, the OSCM passed the sym-

metry test for SOA values = -750, -700, and -650 ms; the PSM passed the test for all non-zero SOA values; and the ASM passed the test for only SOA = -500 ms. The results of a chi-squared test showed that out of 202 observations (across all non-zero SOA values), 122 (60%) agreed with predictions of the OSCM, 127 (63%) with the PSM, and 124 (61%) with the ASM. However, there was a change in proportion of observations in agreement with model predictions as a function of SOA. In the case of the OSCM, for SOA values -600 or less, observations agreed with predictions 64%–76% of the time, while for SOA values greater than -600 ms this fell to less than 58% of the time. In the case of the PSM, observations agreed with predictions for all SOA values except -750 ms (all between 61% and 69%) with the lowest agreement with predictions for SOA of -750 ms (57%). In terms of the ASM, SOA values of -250, -500, -550, -600, and -650 agreed with predictions 62%–67% of the time, while SOA values of -700 and -750 ms agreed 52%–57% of the time.

These results could be due to within- or across-subject factors. Confining attention to within-subject factors, it appears that the PSM predicted the results of 4 of 11 observers in the baseline condition and two of the four observers in experiment 1. The OSCM and PSM each made correct predictions for one observer in Experiment 2A and for two observers each in Experiments 2B and 2C (-500 to -600 and -650 to -750 ms SOA). The ASM made no correct predictions for any subjects in Experiments 2A and 2B and made correct predictions for one subject in Experiment 2C (-500 to -600 ms SOA range).

Across subjects, the PSM predicted 80% of the results in the baseline condition, while in Experiment 1 the PSM and ASM predicted 85% and 80% of the results, respectively. For experiment 2A, the OSCM predicted 86% of the results, the PSM 81%, and the ASM 45%. The results for Experiments 2B and 2C did not differentiate among models, each model predicting roughly 60% of the results. When applied to results from groups of observers, none of the models considered gave an accurate statistical description of all the data (i.e., greater than 95% of measurements agreeing with the predictions of a particular model). Failures to satisfy the predictions of the models are of two types: over- and under-prediction. Over-predictions relative to the OSCM accounted for only roughly 5% of the failures for the baseline condition, Experiments 1, and 2A, and only 7% for Experiments 2B and 2C. Under-predictions relative to the ASM were 2% and 5% for experiments 1 and 2A, and 8% and 7% for the baseline condition and Experiments 2B and 2C, respectively. The cause of the over-prediction failures may be the observer's use of the sub-optimal channel or simple inattention. The cause of the under-prediction failures may be simple inattention in the single-channel presentation conditions.

IV. DISCUSSION

A. Phase and temporal asynchrony effects

Our finding of phase insensitivity leads to several important interpretations regarding the facilitative effects found in the A+T conditions. First, the lack of a phase effect on the

combined-modality scores strongly suggests that the auditory background noise present in all testing was sufficient to mask any possible acoustic artifacts arising from the sinusoidal vibrations produced at the tactile device. If this had not been the case, then the relative phase of the two signals would have resulted in addition and cancellation effects, which would improve or decrease their detection. A second possibility that is ruled out by the present results is that of fine-structure operations at the neural level.⁵ Instead, the similar A+T scores, independent of the relative phase of the auditory and tactile stimuli, suggest that the integration may operate on the envelopes of these stimuli rather than their fine structure. The response pattern measured in the current experiment is consistent with an envelope interaction effect: i.e., an overall increase in response but no change that is correlated with changing relative auditory-tactile phase.

The asymmetry in response patterns for the auditory-leading conditions compared to the tactile-leading conditions found in Experiment 2 (see Fig. 4, lower right panel) is consistent with differences in time constants between the auditory and tactile systems. The auditory-first condition suggests an integration window of no more than 50 ms, while the tactile-first condition suggests a window of up to 150–200 ms.

These implications of a short auditory time constant are consistent with results obtained in studies of auditory forward masking (e.g., [Robinson and Pollack, 1973](#); [Vogten, 1978](#); [Kidd and Feth, 1981](#); [Jesteadt et al., 1982](#); [Moore and Glasberg, 1983](#); [Moore et al., 1988](#); [Plack and Oxenham, 1998](#)), which indicate time constants less than 50 ms. The results reported in the current study suggest that the preceding auditory stimulus was not effective in interacting with the tactile stimulus at any SOA. In the single-modality case of auditory forward masking, however, there is significant interaction between the probe and masker at small time delays. The relatively long (500 ms) signal durations of both the auditory (“masker”) and the tactile (“probe”) stimuli may be partially responsible for the shorter auditory time constant observed here. Auditory studies typically employ brief (tens of milliseconds) probes and strong effects have been demonstrated for an increase in the amount of forward masking with an increase in masker duration ([Fastl, 1977](#); [Kidd and Feth, 1981](#)).

Our finding of a relatively long time constant for tactile stimulation is consistent with results obtained in studies of tactile-on-tactile forward masking (e.g., [Hamer et al., 1983](#); [Gescheider et al., 1989](#); [Gescheider et al., 1994](#); [Gescheider and Migel, 1995](#)). Using tactile maskers with durations on the order of hundreds of milliseconds and tactile probes with durations on the order of tens of milliseconds, previous investigators have reported significant amounts of threshold shift for time delays between masker offset and probe onset on the order of 150–200 ms. Such results suggest that the tactile system maintains a persistent neural response even after cessation of the stimulus (see [Craig and Evans, 1987](#)). Our results are consistent with the sensory effect of the tactile stimulus persisting for at least 150–200 ms following its offset and that this effect is capable of interacting with the subsequent auditory stimulus to facilitate detection. For tac-

tile offset times longer than 200 ms, the facilitatory effect declined and performance on the A+T condition was similar to that in the unimodal conditions.

B. Comparisons with previous multisensory work

The facilitatory effects obtained for simultaneous presentation of A+T signals in our baseline experiment, as well as the effects of temporal asynchrony of the auditory and tactile stimuli, are generally consistent with previous reports in the literature. Facilitative interactions for synchronously presented auditory and tactile stimuli were reported by Schnupp *et al.* (2005) using objective techniques to measure the discriminability of visual (V), auditory (A), and tactile (T) stimuli in VA, VT, and AT combinations. Auditory stimuli were 100-ms bursts of broadband noise presented at a background reference (sound) level of 51 dB SPL. Tactile stimuli were 100-ms bursts of 150-Hz sinusoidal vibrations presented at background reference (force) levels of 16.2–48.5 N. The stimulus on a given trial was a simultaneous pair of either VA, VT, or AT bursts that ranged from 0% to 14% (V and A) or from 0% to 35% (T) in 2% or 5% increments of intensity relative to the background reference level. Observers were instructed to respond whether the background level or an incremented level was presented. Data were analyzed in terms of analogs of both the PSM and ASM. While 2 of 5 AT data sets could be adequately accounted for by the ASM (Schnupp, 2009, personal communication with L.D. Braidà) and 5 of 17 data sets overall, all 17 could be accounted for by the PSM.

Ro *et al.* (2009) measured the effect of presenting a relatively intense (59 dB) 500-Hz, 200-ms tone on the detection of a near-threshold 0.3 ms square-wave electrocutaneous stimulus that felt like a faint tap. They found that the presentation of the auditory stimulus increased d -prime from 2.4 to 2.8. This result was interpreted as evidence that “a task-irrelevant sound can enhance somatosensory perception.”

Facilitative interactions have also been observed using subjective techniques such as loudness matching (Schurmann *et al.*, 2004; Yarrow *et al.*, 2008) and loudness magnitude estimation (Gillmeister and Eimer, 2007). In the two loudness-matching studies, the average intensity required to produce equal loudness of an auditory reference tone was 12%–13% (roughly 0.5 dB) lower under the combined auditory-tactile condition compared with the auditory-alone condition, thus suggesting a facilitative interaction between the auditory and tactile stimuli. Gillmeister and Eimer (2007) found that magnitude estimates of an auditory tone presented in a background of white noise were increased by simultaneous presentation of a tactile stimulus for near-threshold auditory tones, but no loudness increase was observed either for higher intensity tones or for non-simultaneous presentation of the tactile and auditory stimuli. It should be noted, however, that based on the results of other experiments, Yarrow *et al.* (2008) attribute the increase in loudness to a bias effect. They conclude that the tactile stimulus “does not affect auditory judgments in the same manner as a real tone.”

Other previous studies of auditory-tactile integration have measured effects of temporal asynchrony between the

two stimuli and have also demonstrated dependence on the order of stimulus presentation: Gescheider and Niblette (1967) for inter-sensory masking and temporal-order judgments and Bresciani *et al.* (2005) for judgments of auditory numerosity. Consistent with the results of the current study, higher levels of interaction between the two senses were obtained for conditions in which the tactile stimulus is delivered before the auditory stimulus. One exception to this pattern is found in the results of Gillmeister and Eimer (2007). While demonstrating effects of temporal synchrony on the detectability of an auditory tone in the presence of a vibratory pulse, they found no effects of stimulus order. Their detectability results, however, are consistent with the results of their loudness-estimation study.

While the experimental conditions used in these studies differ from one another, they all suggest that temporal synchrony is an important factor in showing facilitative auditory-tactile interaction. The current study has shown in greater detail the asymmetry in the temporal window involved in auditory-tactile detection, such that when the tactile stimulus precedes the auditory by up to 200 ms, a facilitative interaction significantly greater than the unimodal levels is measured. This level of response is not seen when the auditory stimulus precedes the tactile, however, as bimodal responses at all asynchronous time periods are not different from unimodal levels.

C. Implications of model results

The amount of integration measured in this study was quantified by comparing performance with the predictions of three models of the integration process: the OSCM, the PSM, and the ASM. It should be noted that these models are not mutually exclusive in the sense that observers need not base their decisions exclusively on one model in all experiments. If the auditory stimulus is presented before the tactile, it is unlikely that the ASM would apply, while it might apply when there is temporal overlap. Also, the predictions of more than one model may fit the data equally well. For example, in the hypothetical case considered in Sec. II F, based on 75 trials the score of 75%-correct would be within 2 standard deviations of the predictions of all three models. Many of the two-frequency results of Marril (1956) can be accounted for by two of these three models. It is only possible to distinguish among the three models based on more than one experimental result, i.e., several results from one observer or the results of multiple observers. When performance exceeds the predictions of the OSCM, this implies at least partial integration of cues, and when performance exceeds predictions of the PSM, this implies at least partial within-channel integration.

In this study, the results show that measurements are more often successfully modeled by the PSM approach than by the OSCM or ASM approaches and are consistent with those found previously in auditory-alone studies (Green, 1958), tactile-alone studies (Bensmaïa *et al.*, 2005), and in multisensory studies (audio-visual and audio-visual-tactile: Braidà, 1991; visual-tactile: Ernst and Banks, 2002; discrimination of pairs of visual-auditory, visual-tactile, and audio-

tactile stimuli: Schnupp *et al.*, 2005). Although most of these studies did not attempt to model the observations with an ASM, Schnupp (2009), personal communication with L.D. Braidá, found that two of five audio-tactile discrimination data sets could be fit by an ASM (all five were fit by a PSM), and we found in experiment 1 that nearly the same number of experiments were accounted for by the ASM as by the PSM.

Thus, we found, in accord with Schnupp *et al.* (2005), that overall the PSM best accounts for the improvement in detectability when auditory and tactile stimuli are combined. There are significant differences, however: The OSCM provides a slightly better account when auditory stimuli precede tactile stimuli, and the ASM provides nearly as good an account of the (non-)effects of varying relative auditory-tactile phase. One problem with this interpretation is that the different models make predictions of detectability that are always ordered: $OSCM \leq PSM \leq ASM$. Thus, for example, if an observer behaves in accord with the ASM but makes a few responses due to inattention, the PSM will tend to be favored. While we discarded data sets for which there were indications that unimodal observer detection had decreased during the course of a single session, it is likely that some reduction in bimodal detection may have occurred as well. Because Schnupp *et al.* (2005) collected data over two or three sessions, it is also possible that criterion shifts may have reduced apparent performance, thus favoring the PSM over the ASM.

It is also possible that the PSM provides a better description of the data than the ASM when qualitatively different stimuli are detected or discriminated. The traditional explanation for the two-frequency detection results of Marrill (1956) and Green (1958) is that the PSM provides a good account of the detection of pairs of tones whose frequencies lie in distinct critical bands while the ASM is appropriate for tones whose frequencies lie in the same critical band. Wilson *et al.* (2008), who tested the detection of auditory and tactile tones of varying frequency, found that performance generally declined as the frequency difference increased. It is possible that Schnupp *et al.* (2005) found that a PSM-like model applied to discrimination of auditory noise and a tactile tone for this reason.

Stein and Meredith (1993) suggested that additive and super-additive responses are a way of measuring facilitative multisensory responses. The different models suggest different mechanisms for integration, with the Pythagorean sum modeling two independent pathways integrating the different stimuli after each has been processed by its own sensory system and the algebraic sum modeling stimuli that are integrated before being processed, leading to a greater level of integration overall. It is possible that both the results of Schnupp *et al.* (2005) and our results, which show that subjects can utilize both Pythagorean and algebraic approaches to integration, suggest that the auditory and tactile sensory systems are capable of integrating in both manners, and both mechanisms are being employed during our experiment.

D. Relationship to neuroanatomy

One potential anatomical pathway for Pythagorean integration may be the ascending somatosensory inputs to the somatosensory cortex, which then project to the auditory cortex. Thus, two independent pathways are operating on input from each of the modalities, and the multisensory stimuli are processed only after the single-modality operations have taken place. A different anatomical pathway that may account for algebraic integration comes from the ascending somatosensory inputs that target early auditory centers (i.e., in the brainstem and thalamus) and thereby affect changes in auditory-tactile integration before the combined signal reaches the auditory cortex. The fact that we see observed responses that are greater than the prediction of PSM suggests that the auditory and somatosensory systems are working together in one multisensory area to process the stimuli.

V. CONCLUDING REMARKS

Our study has shown that certain combinations of auditory and tactile signals result in a significant increase in detectability above the levels when the stimuli are presented in isolation. This is not due to changes in response bias (e.g., Yarrow *et al.*, 2008), as indicated by a detection theory analysis. Specifically, 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. 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. These differences are consistent with the neural mechanics of auditory-on-auditory masking and tactile-on-tactile masking.

Our results were compared with three models of integration. While it is not always possible to differentiate among the models on the basis of a single experimental outcome, the models sort themselves out if one combines results across sessions and/or observers. If one assumes that all observers use a single model in all experiments, then the PSM gives a better fit to the data than the OSCM or the ASM.

Further research is being conducted to examine the effects of other stimulus parameters (including frequency and intensity) on the perceptual aspects of auditory-tactile integration.

ACKNOWLEDGMENTS

This research was supported by grants from the National Institutes of Health (Grant Nos. 5T32-DC000038, R01-DC000117, and R01-DC00126) and by a Hertz Foundation Fellowship (E.C.W.). The authors would like to thank D.M. Green for helpful discussions, J.W.H. Schnupp for sharing data, and Richard Freyman, Fan-Gang Zeng, and the anonymous reviewer for helpful comments on earlier versions of this manuscript.

APPENDIX: ESTIMATIONS OF BONE-CONDUCTED SOUND LEVELS ARISING FROM VIBROTACTILE STIMULATION AT 250 HZ

We consider two possibilities and show that they are unlikely to be responsible for our results: (1) In baseline conditions the vibratory stimulus is detected through the auditory sense. (2) In Experiment 1, the phase dependent combination of vibratory and acoustic stimuli is responsible for the phase dependence of our results. Note that masking noise was used in an attempt to ensure that the task is performed solely through the sense of touch without spurious auditory cues.

Consider first the possibility that bone-conducted sound from vibratory stimulation was responsible for detection of the tactile stimulus. In the measurements of Dirks *et al.* (1976), bone-conduction thresholds for normal listeners in force and acceleration units indicate that the 250-Hz bone-conduction threshold, when measured with a vibrator placed on the mastoid, is 10 dB *re* 1 cm/s² (acceleration units). The maximum displacement of our 250-Hz signal (roughly 5 dB SL) corresponds to a peak displacement of -20 dB *re* 1 μm peak and an acceleration of roughly 5 dB *re* 1 cm/s², roughly 5 dB less than the bone-conduction threshold for mastoid stimulation. It is fairly safe to assume that stimulation of the middle finger results in a highly attenuated bone-conducted signal compared to stimulation of the mastoid. The bone-conducted threshold at the forehead is 12 dB higher than at the mastoid. The impedance mismatches created by tissue and bone junctions from the fingertip to the skull would lead to even higher thresholds, perhaps by 13 dB, than for the forehead. Thus, the highest signal reaching the ear through bone-conducted sound at 250 Hz would be -20 dB *re* 1 cm/s². The bone-conducted threshold at 250 Hz is 10 dB cm/s²; thus, our bone-conducted stimulus would be roughly -30 dB SL, that is, roughly 30 dB below the air-conducted threshold of 18 dB SPL at 250 Hz (Houtsma, 2004) or equivalent to an acoustic stimulus of -12 dB SPL. Such bone-conducted sound would be undetectable.

Assuming a critical ratio of 17.5 dB at 250 Hz and a noise spectral level of 7.4 dB/Hz, the level of the acoustic tone is roughly 25 dB SPL, and (as noted above) the equivalent vibratory stimulus is -12 dB SPL, 37 dB below the level of the acoustic tone. This would cause the 25 dB SPL tone to vary at most from 24.9 to 25.1 dB SPL as the phase is changed. To understand the effect of this phase change, we make use of some unpublished data on the detection of auditory stimuli of different amplitudes: 25 and 27 dB, which correspond to detection rates in the 50 dB SPL noise of 70.9%- and 79.9%-correct, respectively, or about 4.5 percentage points per decibel. Thus, the combination of the bone- and the air-conducted sound would cause the detection rate to change from 70.3% to 71.5%. This is contrary to the results of Experiment 1, which indicate that in the A+T condition in scores varied between 83.2%- and 84.6%-correct with standard errors of less than 1.3 percentage points. This indicates that the effect of combining the vibratory and acoustic stimuli cannot be accounted for by bone conduction alone.

¹Data collected from an additional two subjects (S_5 in experiment 1 and S_7 in experiment 2C) were discarded on the basis of abnormally low values of thresholds for the tactile stimuli that were inconsistent with those of the other subjects and with results in the literature.

²Three subjects (S_2 , S_4 , and S_7) were also tested in two additional conditions in experiment 1 (phase=0°, SOA=+600 ms, and SOA=-600 ms) in addition to the four phase conditions. These subjects later participated in experiment 2B, and SOA values of ±600 ms were not repeated.

³Due to experimenter error, performance on the combined A+T (SOA=0 ms) condition was not measured in several of the experiment 2B test sessions for three subjects (S_2 , S_4 , and S_7), although performance on A-alone and T-alone conditions was always established at the beginning of each session.

⁴We denote *d*-primes that can be estimated directly from the data using lower case (*d*'), *d*-primes that are predicted by models in upper case letters (*D*').

⁵Although both types of interactions might occur simultaneously and cancel, we regard this possibility as unlikely.

- Bensmaia, S. J., Hollins, M., and Yau, J. (2005). "Vibrotactile intensity and frequency information in the Pacinian system," *Percept. Psychophys.* **67**, 828–841.
- Bhattacharya, J., Shams, L., and Shimojo, S. (2002). "Sound-induced illusory flash perception: Role of gamma band responses," *NeuroReport* **13**, 1727–1730.
- Braida, L. D. (1991). "Crossmodal integration in the identification of consonant segments," *Q. J. Exp. Psychol.* **43A**, 647–677.
- Bresciani, J. P., Ernst, M. O., Drewing, K., Bouyer, G., Maury, V., and Kheddar, A. (2005). "Feeling what you hear: Auditory signals can modulate tactile tap perception," *Exp. Brain Res.* **162**, 172–180.
- Caclin, A., Soto-Faraco, S., Kingstone, A., and Spence, C. (2002). "Tactile 'capture' of audition," *Percept. Psychophys.* **64**, 616–630.
- Caetano, G., and Jousmaki, V. (2006). "Evidence of vibrotactile input to human auditory cortex," *Neuroimage* **29**, 15–28.
- Cappe, C., and Barone, P. (2005). "Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey," *Eur. J. Neurosci.* **22**, 2886–2902.
- Craig, J. C., and Evans, P. M. (1987). "Vibrotactile masking and the persistence of tactual features," *Percept. Psychophys.* **42**, 309–317.
- Dirks, D. D., Kamm, C., and Gilman, S. (1976). "Bone-conduction thresholds for normal listeners in force and acceleration units," *J. Speech Hear. Res.* **19**, 181–186.
- Ernst, M. O., and Banks, M. S. (2002). "Humans integrate visual and haptic information in a statistically optimal fashion," *Nature (London)* **415**, 429–433.
- Fastl, H. (1977). "Subjective duration and temporal masking patterns of broadband noise impulses," *J. Acoust. Soc. Am.* **61**, 162–168.
- Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Javitt, D. C., Guilfoyle, D., Ritter, W., and Murray, M. M. (2002). "Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study," *J. Neurophysiol.* **8**, 540–543.
- Fu, K. M. G., Johnston, T. A., Shah, A. S., Arnold, L., Smiley, J., Hackett, T. A., Garraghty, P. E., and Schroeder, C. E. (2003). "Auditory cortical neurons respond to somatosensory stimulation," *J. Neurosci.* **23**, 7510–7515.
- Gescheider, G. A., Barton, W. G., Bruce, M. R., Goldberg, J. H., and Greenspan, M. J. (1969). "Effects of simultaneous auditory stimulation on the detection of tactile stimuli," *J. Exp. Psychol.* **81**, 120–125.
- Gescheider, G. A., Bolanowski, S. J., and Verrillo, R. T. (1989). "Vibrotactile masking: Effects of stimulus onset asynchrony and stimulus frequency," *J. Acoust. Soc. Am.* **85**, 2059–2064.
- Gescheider, G. A., Hoffman, K. E., Harrison, M. A., and Travis, M. L. (1994). "The effects of masking on vibrotactile temporal summation in the detection of sinusoidal and noise signals," *J. Acoust. Soc. Am.* **95**, 1006–1016.
- Gescheider, G. A., Kane, M. J., and Sager, L. C. (1974). "The effect of auditory stimulation on responses to tactile stimuli," *Bull. Psychon. Soc.* **3**, 204–206.
- Gescheider, G. A., and Migel, N. (1995). "Some temporal parameters in vibrotactile forward masking," *J. Acoust. Soc. Am.* **98**, 3195–3199.
- Gescheider, G. A., and Niblette, R. K. (1967). "Cross-modality masking for touch and hearing," *J. Exp. Psychol.* **74**, 313–320.
- Gillmeister, H., and Eimer, M. (2007). "Tactile enhancement of auditory detection and perceived loudness," *Brain Res.* **1160**, 58–68.

- Green, D. M. (1958). "Detection of multiple component signals in noise," *J. Acoust. Soc. Am.* **30**, 904–911.
- Guest, S., Catmur, C., Lloyd, D., and Spence, C. (2002). "Audiotactile interactions in roughness perception," *Exp. Brain Res.* **146**, 161–171.
- Hackett, T. A., de la Mothe, L. A., Ulbert, I., Karmos, G., Smiley, J., and Schroeder, C. E. (2007). "Multisensory convergence in auditory cortex, II. Thalamocortical connections of the caudal superior temporal plane," *J. Comp. Neurol.* **502**, 924–952.
- Hamer, R. D., Verrillo, R. T., and Zwislocki, J. J. (1983). "Vibrotactile masking of pacinian and non-pacinian channels," *J. Acoust. Soc. Am.* **73**, 1293–1303.
- Hawkins, J. E., and Stevens, S. S. (1950). "The masking of pure tones and of speech by white noise," *J. Acoust. Soc. Am.* **22**, 6–13.
- Houtsma, A. J. (2004). "Hawkins and Stevens revisited with insert ear-phones," *J. Acoust. Soc. Am.* **115**, 967–970.
- Jesteadt, W., Bacon, S. P., and Lehman, J. R. (1982). "Forward masking as a function of frequency, masker level, and signal delay," *J. Acoust. Soc. Am.* **71**, 950–962.
- Jousmaki, V., and Hari, R. (1998). "Parchment-skin illusion: Sound-biased touch," *Curr. Biol.* **8**, R190–191.
- Kayser, C., Petkov, C. I., Augath, M., and Logothetis, N. K. (2005). "Integration of touch and sound in auditory cortex," *Neuron* **48**, 373–384.
- Kidd, G., Jr., and Feth, L. L. (1981). "Patterns of residual masking," *Hear. Res.* **5**, 49–67.
- Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., and Schroeder, C. E. (2007). "Neuronal oscillations and multisensory interaction in primary auditory cortex," *Neuron* **53**, 279–292.
- Lamore, P. J., Muijser, H., and Keemijik, C. J. (1986). "Envelope detection of amplitude-modulated high-frequency sinusoidal signals by skin mechanoreceptors," *J. Acoust. Soc. Am.* **79**, 1082–1085.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Marrill, T. (1956). "Detection theory and psychophysics." Technical Report No. 319, Research Laboratory of Electronics, MIT, Cambridge, MA.
- McGurk, H., and MacDonald, J. (1976). "Hearing lips and seeing voices," *Nature (London)* **264**, 746–748.
- Moore, B. C. J., and Glasberg, B. R. (1983). "Growth of forward masking for sinusoidal and noise maskers as a function of signal delay; implications for suppression in noise," *J. Acoust. Soc. Am.* **73**, 1249–1259.
- Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (1988). "The shape of the ear's temporal window," *J. Acoust. Soc. Am.* **83**, 1102–1116.
- Neville, A. M., and Kennedy, J. B. (1964). *Basic Statistical Methods for Engineers and Scientists* (International Textbook, Scranton, PA), p. 133.
- Plack, C. J., and Oxenham, A. J. (1998). "Basilar-membrane nonlinearity and the growth of forward masking," *J. Acoust. Soc. Am.* **103**, 1598–1608.
- Rabinowitz, W. M., Houtsma, A. J. M., Durlach, N. I., and Delhorne, L. A. (1987). "Multidimensional tactile displays: Identification of vibratory intensity, frequency and contactor area," *J. Acoust. Soc. Am.* **82**, 1243–1252.
- Ro, T., Hsu, J., Yasar, N. E., Elmore, L. C., and Beauchamp, M. S. (2009). "Sound enhances touch perception," *Exp. Brain Res.* **195**, 135–143.
- Robinson, C. E., and Pollack, I. (1973). "Interaction between forward and backward masking: A measure of the integrating period of the auditory system," *J. Acoust. Soc. Am.* **53**, 1313–1316.
- Schnupp, J. W. H., Dawe, K. L., and Pollack, G. (2005). "The detection of multisensory stimuli in an orthogonal sensory space," *Exp. Brain Res.* **162**, 181–190.
- Schroeder, C. E., Lindskey, R. W., Specht, C., Marcovici, A., Smiley, J. F., and Javitt, D. C. (2001). "Somatosensory input to auditory association cortex in the macaque monkey," *J. Neurophysiol.* **85**, 1322–1327.
- Schurmann, M., Caetano, G., Hlushchuk, Y., Jousmaki, V., and Hari, R. (2006). "Touch activates human auditory cortex," *Neuroimage* **30**, 1325–1331.
- Schurmann, M., Caetano, G., Jousmaki, V., and Hari, R. (2004). "Hands help hearing: Facilitatory audiotactile interaction at low sound-intensity levels," *J. Acoust. Soc. Am.* **115**, 830–832.
- Sekuler, R., Sekuler, A. B., and Lau, R. (1997). "Sound alters visual motion perception," *Nature (London)* **385**, 308.
- Smiley, J. F., Hackett, T. A., Ulbert, I., Karmos, G., Lakatos, P., Javitt, D. C., and Schroeder, C. E. (2007). "Multisensory convergence in auditory cortex, I. Cortical connections of the caudal superior temporal plane in macaque monkeys," *J. Comp. Neurol.* **502**, 894–923.
- Soto-Faraco, S., Spence, C., and Kingstone, A. (2004). "Congruency effects between auditory and tactile motion: Extending the phenomenon of cross-modal dynamic capture," *Cogn. Affect. Behav. Neurosci.* **4**, 208–217.
- Stein, B. E., London, N., Wilkinson, L. K., and Price, D. D. (1996). "Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis," *J. Cogn. Neurosci.* **8**, 497–506.
- Stein, B. E., and Meredith, M. A. (1993). *The Merging of the Senses* (MIT, Cambridge, MA).
- Verrillo, R. T., Gescheider, G. A., Calman, B. G., and Van Doren, C. L. (1983). "Vibrotactile masking: Effect of one- and two-site stimulation," *Percept. Psychophys.* **33**, 379–387.
- Vogten, L. L. M. (1978). "Simultaneous pure-tone masking: The dependence of masking asymmetries on intensity," *J. Acoust. Soc. Am.* **63**, 1509–1519.
- Wilson, E. C., Reed, C. M., and Braida, L. D. (2008). "Perceptual interactions between vibrotactile and auditory stimuli: Effects of frequency," in *International Multisensory Research Forum Conference*, Hamburg, Germany, July.
- Woods, T. M., and Recanzone, G. H. (2004). "Cross-modal interactions evidenced by the ventriloquism effect in humans and monkeys," in *Handbook of Multisensory Processes*, edited by G. Calvert, C. Spence, and B. E. Stein (MIT, Cambridge, MA), pp. 35–48.
- Yarrow, K., Haggard, P., and Rothwell, J. C. (2008). "Vibrotactile-auditory interactions are post-perceptual," *Perception* **37**, 1114–1130.
- Zhou, J., and Shore, S. (2004). "Projections from the trigeminal nuclear complex to the cochlear nuclei: A retrograde and anterograde tracing study in the guinea pig," *J. Neurosci. Res.* **78**, 901–907.