AUDITORY-TACTILE INTEGRATION: EFFECTS OF PHASE OF SINUSOIDAL STIMULATION AT 50 AND 250 Hz

Parivash Ranjbar[1], E. Courtenay Wilson[2], Charlotte M. Reed[2*],

and Louis D. Braida[2]

[1]Audiological Research Centre, University Hospital at Örebro, Sweden,

[2] Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA

ABSTRACT

The perceptual integration of 50- and 250-Hz, 500-ms vibrotactile and auditory tones was studied in detection experiments as a function of the relative phase (0˚, 72˚, 144˚, 216˚, and 288˚) 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 SPL broadband noise. The observers were four adults with normal hearing. The vibrotactile and auditory stimulus levels used each yielded 63-77%-Correct unimodal detection performance in a two-interval two-alternative forced-choice task. Scores for the auditory-alone and tactile-alone conditions averaged roughly 70%-Correct. Mean scores for the auditory plus tactile conditions averaged across different phases were 77.1%-Correct at 50 Hz and 79.6%-Correct at 250 Hz. At 50 Hz, no differences in performance were observed as a function of the relative phase at which the combined auditory and tactile signals were presented. At 250 Hz, significantly higher scores were observed for one phase combination (72˚) compared to two of the other four relative phases. Performance on the auditory plus tactile conditions resulted in significant integrative effects and was generally more consistent with a "Pythagorean Sum" model than with either an "Algebraic Sum" or an “Optimum Single Cannel” model of perceptual integration.

Keywords: auditory-tactile, multisensory, integration, phase, detection

1. INTRODUCTION

Recent perceptual studies indicate that the senses of audition and touch can interact in the formation of unified auditory-tactile percepts [1-10]. These perceptual results are supported by recent neuroanatomical [11-15] and neurophysiological data [16-23] showing strong connections between the auditory and tactile senses. Such multi-sensory convergence has important implications for human communication in both the production of speech (which involves simultaneous auditory and somatosensory feedback) and the perception of speech (which may be enhanced in deaf persons through multimodal auditory-tactile inputs at low frequencies).

Recently we have used rigorous quantitative psychophysical techniques based on detection theory and models of cross-modal integration to demonstrate that certain combinations of auditory and tactile stimuli result in a significant increase in detect ability above performance when the stimuli are presented in isolation [24, 25]. For tonal stimuli, the size of the increase is dependent on the frequencies of stimulation employed within each modality, with larger increases occurring for more closely spaced tones. While these detect ability effects have been found for near-threshold tones, we obtained similar and consistent findings at supra-threshold stimulus levels using loudness matches as the criterion for interaction [26]. Specially, the loudness of a pair of auditory-tactile tones is less if the tones are of the same frequency than if they are widely spaced in frequency.

An additional variable addressed in the detection experiments was the role of the starting phase of the tactile relative to the auditory sine wave. For a 250-Hz stimulus, auditory-tactile integration was independent of the phase at which the stimuli were combined across the two modalities. This lack of a phase effect suggests that auditory-tactile integration likely operates on the envelopes as opposed to the fine structure of the stimuli in each modality.

The current experiment extends the range of frequencies used in the experiment in which we found no effect of auditory-tactile phase on detection of a 250-Hz tone [24]. It is possible, however, that relative phase could play a role in multimodal detection at frequencies lower than 250 Hz. In audition, the relative inter-aural phase of tones contributes to binaural detection only at frequencies below about 1400 Hz [27] but is unimportant above that. Likewise, in the tactual sensory system, it is possible that phase effects are apparent at lower frequencies (e.g., below 60 Hz) where detection is thought to depend on a non-Pacinian tactile channel [28], which was not addressed in our previous study [24]. A different set of relative phase values was also employed in the current experiment. Whereas Wilson, Reed, and Braida [24] examined four phases (0˚, 90˚, 180˚, and 270˚), the current experiment employed a set of five phases with 72˚ spacing between 0˚ and 288˚. This new set of phases was selected to explore further the generality of the previous results.

The experiments reported here used measures of d' and %-Correct for auditory-alone, tactile-alone, and combined auditory-tactile presentations of signals near the threshold of detection. The observed performance in the combined condition was then compared to three predictions of multi-modal performance derived from observed measures of detect ability within each of the two separate sensory modalities.

2. METHODS

2.1 Stimuli

The auditory stimulus was either a 50- or 250-Hz pure tone presented in a background of pulsed 50 dB SPL Gaussian broadband noise (bandwidth: 0.1 to 11.0 kHz). The tactile stimulus was a sinusoidal vibration with a frequency of either 50- or 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 50- or 250-Hz signals in both modalities were generated digitally (using MATLAB™ software; Math Works, Natick, MA) to have a total duration of 500 ms that included 100-ms raised cosine-squared rise/fall times. Thus, the tone consisted of a 100-ms rise, a 300-ms steady-state, and a 100-ms fall.

The digitized signals were played through a D/A 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). A laser accelerometer was used to calibrate the tactile device.

2.2 Subjects

Four subjects ranging in age from 19 to 42 years (three females and one male) participated in this study (which was approved by the MIT Committee on Use of Humans as Experimental Subjects). 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. Audio logical testing was conducted on the first visit to the laboratory to determine that subjects met the criterion of normal audiometric thresholds (20 dB HL or better at frequencies of 125, 250, 500, 1000, 2000, 4000 and 8000 Hz). Two of the subjects (S1 and S2) had participated in previous studies of frequency effects in auditory-tactile integration [25, 26] but not in the previous [24]. The remaining two subjects (S3 and S4) were naive to auditory-tactile integration experiments.

2.3 Experimental Conditions

The experiments examined the perceptual integration of 50- or 250-Hz sinusoidal auditory and vibrotactile signals as a function of relative phase that were each presented near the threshold of detection. Threshold measurements were first obtained under each of the two single-modality conditions (A and T separately). Then the detectability of the combined auditory plus tactile (A+T) signal was measured at threshold levels within each of the two individual modalities. The experimental conditions examined the effects of relative phase of the tactile signal relative to the auditory signal.

The auditory starting phase was always 0˚, while the tactile starting phase took on five different values: 0˚, 72˚, 144˚, 216˚, and 288˚. In each of these five conditions, the auditory and tactile stimuli were temporally synchronous and thus had identical onset and offset times. Measurements were made at two test frequencies (50 and 250 Hz). The order in which these frequencies were tested was selected at random for each subject. Then the order of the five phases (0˚, 72˚, 144˚, 216˚, and 288˚) was randomized for each subject for each replication at each test frequency. All replications were completed at the first test frequency before testing the second frequency. The number of replications collected for each subject at each test frequency is described below.

2.4 Experimental Procedures

For all experiments, subjects were seated in a double-walled sound-treated booth and were presented 50 dB SPL broadband noise diotically via Sennheiser 580 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 (0.9 cm diameter) 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 [29], (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 each of the five A+T conditions and (iv) single-modality fixed-level testing was repeated as in (ii) except with an expanded acceptable performance range of 56-84%-Correct (± 2 standard deviations around 70%-Correct). (Data from the second set of single-modality conditions was not otherwise used.) If the single-modality threshold re-tests at the end of a given session were outside the 56-84%-Correct range the data for that session were discarded.

A test session typically lasted two hours for each subject (including frequent breaks) and consisted of one complete repetition of the experiment at one test frequency. Attention to the combined A+T stimulus was ensured by having subjects count the number of times they perceived a signal, but this count was ignored in the data analysis.

For each subject, four training sessions identical to the experimental sessions (two sessions with each of the two test frequencies) were provided before data were recorded. Test sessions were then initiated and continued until the criteria described above for stable post-test single-modality thresholds were achieved in four complete replications of the experiment at each frequency. Two subjects were tested in the order of 250 Hz followed by 50 Hz (S1 and S2) and two in the opposite order (S3 and S4). The total number of test sessions required to obtain four stable sets of data at 50 Hz was 9 for S1, 5 for S2, 7 for S3 and 8 for S4, and at 250 Hz was 4 for S1 and S2, 6 for S3, and 5 for S4. Thus, the total number of sessions discarded per subject across both test frequencies ranged from 1 (S2) to 5 (S1, S3, and S4).

Adaptive threshold testing was conducted to obtain an initial estimate of the levels for the single-modality conditions using a three-interval three-alternative forced-choice (3I, 3AFC) adaptive procedure with trial-by-trial correct-answer feedback. Other than helping to determine the levels used in subsequent tests, the results of these adaptive tests were discarded.

The single-modality adaptive threshold estimates were then used to measure A-alone and T-alone performance in a two-interval two-alternative forced-choice (2-I, 2-AFC) fixed-level procedure with 75 trials per run. The order in which A-alone and T-alone were tested was selected at random for each subject on each test day. Stimulus levels were adjusted and runs were repeated until scores of 63-77%-Correct were obtained for each modality tested separately. 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 500 ms. Each interval 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 2.5 seconds, 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 after each trial 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.

2.5 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' and bias). The measures of bias were discarded while the measures of d' were averaged across the repetitions of each experimental condition within a given subject. Statistical tests performed on the data included 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.

2.6 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 observer's responses are based on the better of the tactile or auditory input channels. The predicted D'_{OSCM} (see Note 1) for the combined A+T condition is the greater of the tactile (d'_T) or auditory (d'_A),

$$D'_{OSCM} = \mathrm{Max}(d'_T, d'_A) \qquad (1)$$

The PSM assumes that integration occurs across channels e.g., as in audio-visual integration, [30] and that the measure of sensitivity in the combined auditory-tactile condition is the Pythagorean-sum of the d's for the separate channels,

$$D'_{PSM} = \sqrt{d'^2_A + d'^2_T} \qquad (2)$$

The ASM, on the other hand, assumes that integration occurs within a given channel and that the combined measure of sensitivity is the linear sum of the d's for the separate channels,

$$D'_{ASM} = d'_A + d'_T \qquad (3)$$

For example, if the auditory d'_A was 1.00 (69%-Correct) and the tactile d'_T was 0.80 (66%-Correct), the OSCM would predict a D'_{OSCM} of 1.00 (69%-Correct), the PSM would predict a D'_{PSM} of 1.28 (74%-Correct) and the ASM would predict a D'_{ASM} of 1.80 (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' values were determined for each auditory (d'_A) and tactile (d'_T) experiment, on the basis of 75 total trials. Second, predicted d' values were computed for the three models according to the formulas given above. Third, predicted %-Correct scores were computed for each of the models in the following manner:

$$Predicted\ \%\text{-}Correct = 100 * \phi(\frac{D'_{A+T}}{2}) \quad \textbf{(4)}$$

where 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 %-Correct score was computed as

$$Estimated\ \%\text{-}Correct = 100 * \phi(\frac{d'_{A+T}}{2}) \quad (5)$$

This relatively small adjustment **(**0.76 percentage points on average, 6 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 [31]. This analysis allows for errors in both the observed score and the predicted score. Finally, as another means of evaluating the performance of the three models, the ratio of the observed d' to predicted D' was computed for each repetition at each phase for individual subjects.

3. RESULTS

3.1 Signal Levels Employed in Single-Modality Conditions

Single-modality scores were obtained both at the beginning and ending of each individual test session. The data reported here, however, are based solely on the initial measurements and the post-experiment measurements were employed merely for determining threshold stability.

3.1.1 Levels for Auditory-Alone Conditions

The mean signal levels in dB SPL established for performance in the range of 63-77%-Correct for auditory-alone presentation of 50-Hz and 250-Hz signals in 50-dB SPL broadband noise are shown in the left side of Table 1. Mean signal levels and standard deviations (SD) over the four replications of the experiment are shown for each individual subject. At 250 Hz, the mean signal level and SD across replications and subjects was 20.9 ± 0.93 dB SPL. These levels showed little variability either within or across subjects and are consistent with those reported by Wilson et al. [24, 25] for 250 Hz tones in the same level of broadband noise employed here. At 50 Hz, the mean signal level over replications and subjects was 42.2 ± 2.7 dB SPL. Although showing higher variability across subjects than at 250 Hz, the results are in line with those reported by Wilson et al. [25] where 50-Hz tone levels in a 50-dB SPL noise averaged roughly 46 dB SPL. Mean critical ratios were calculated using measured values of the spectrum level of the noise of 7.4 dB/Hz at 250 Hz and 20 dB/Hz at 50 Hz. The resulting critical ratios are 13.4 dB at 250 Hz and 20.2 dB at 50 Hz. These values are consistent with those reported in the literature [32, 33] and indicate that subjects were listening to the auditory tones in noise at levels that were close to masked threshold.

Table 1: *The Mean and SD of A and T Threshold Levels in 50 dB Broadband Noise at 50 and 250 Hz. Auditory-Alone thresholds are in dB SPL and Tactile-Alone thresholds are in dB re: 1 µm peak displacement.*

		Auditory-Alone		Tactile-Alone	
Frequency	Subject	Mean	SD	Mean	SD
50-Hz	S1	43.68	1.04	-6.38	2.43
	S2	44.20	1.47	-1.53	1.02
	S3	42.70	1.78	-0.65	1.64
	S4	38.10	0.20	-2.82	0.62
	Mean	42.19	2.74	-2.85	2.65
250-Hz	S1	20.93	1.02	-25.20	2.88
	S2	21.50	1.00	-30.63	0.96
	S3	20.53	1.00	-25.40	1.16
	S4	21.00	0.82	-16.05	1.75
	Mean	20.99	0.93	-24.31	5.66

3.1.2 Levels for Tactile-Alone Conditions

The mean signal levels established for performance in the range of 63-77%-Correct for 50-Hz and 250-Hz sinusoidal vibrations to the left middle fingertip are shown in the right side of Table I. All threshold measurements were obtained in the presence of a diotic 50 dB SPL broadband noise presented over headphones. Signal levels are given in dB re: 1 µm peak displacement for individual subjects. Mean signal levels and SD over the four replications of the experiment are shown for each individual subject. Average signal levels employed at 250 Hz ranged from -30.6 dB re: 1 µm (S2) to -16.0 dB (S4) with a mean of -24.3 ± 5.7 dB (consistent with both the range over subjects and the mean of -24.2 dB reported by Wilson et al. [25]. At 50 Hz, the average signal levels ranged from -0.68 dB re: 1 µm peak (S3) to -6.38 dB (S1) with a mean of -2.85 ± 2.7 dB (consistent with the range and mean of 1.0 dB reported by Wilson et al. [25]. The largest variability across subjects was observed at 250 Hz where the threshold of the 42-year old subject was higher than the other three subjects (whose mean age was 23 years). The observation of a possible age effect at 250 Hz but not at 50 Hz is consistent with previous studies indicating a greater effect of age in the Pacinian compared to non-Pacinian systems of the tactile sensory system [34]. [See the Appendix of Wilson et al. [24] for a discussion of the unlikely possibility that the tactile stimulus was detected auditorially via bone conduction.]

3.2 Auditory Plus Tactile Performance as a Function of Relative Phase

The individual results of each subject obtained with the 50-Hz and 250-Hz test frequencies are provided in the Appendix. Mean percent-correct scores and SD across the four runs for each subject at each of the seven experimental conditions are provided for the 50-Hz and 250-Hz test frequencies: A-alone, T-alone, and combined A+T with five different values of the starting phase of the tactile stimulus relative to that of the auditory stimulus (0˚, 72˚, 144˚, 216˚, and 288˚). The mean and SD across subjects for each test condition are provided in the final column of the Appendix.

Results obtained with the 50-Hz test frequency by the four subjects are shown on the top of Fig. 1. Percent-correct scores are given for each of the seven experimental conditions: A-alone, T-alone, and combined A+T with five different values of the starting phase of the tactile stimulus relative to that of the auditory stimulus. Results are shown for individual subjects and for means across subjects. Average scores were 70.3 ± 3.8%-Correct for A-alone and 69.7 ± 3.5%-Correct for T-alone. For the five phase conditions, the A+T scores were: 76.8 ± 10.4 (0˚), 78.3 ± 8.2 (72˚), 78.1 ± 9.4 (144˚), 75.8 ± 8.7 (216˚), and 76.8 ± 9.4%-Correct (288˚) for an average of 77.1 ± 9.1%-Correct.

A mixed-model ANOVA with experimental condition as a fixed variable and subject as a random variable showed significant effects for condition [$F(6, 18)= 4.15$, $p<0.01$)] and subject [$F(3, 18) = 7.81$, $p<0.01$] but not for their interaction [$F(18, 84) = 0.96$, $p=0.514$]. The *post hoc* analysis showed no differences between the results of the conditions A-Alone and T-Alone. The A+T conditions with phases of 72˚ and 144˚ had significantly ($p<0.05$) higher scores than the A-alone and T-Alone conditions; no other significant differences between conditions were observed. Thus, none of the ten possible pairs of A+T conditions had scores that were significantly different from each other. The scores of Subject 4 were significantly higher than those of Subjects 1 and 3.

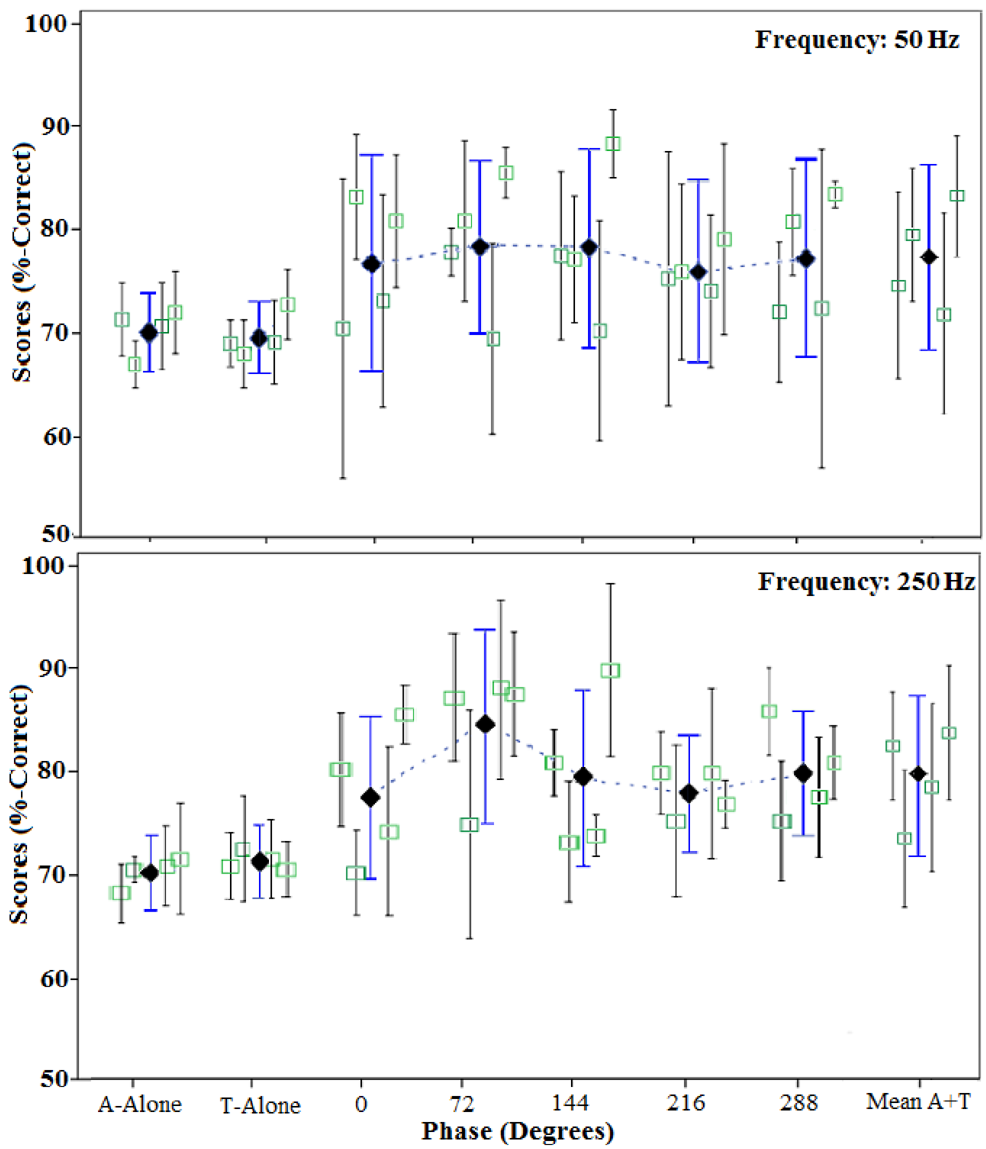

Figure 1. Mean percent-correct scores at 50 Hz (top panel) and 250 Hz (bottom panel) for individual subjects averaged across the four runs and means across subjects at each of the seven conditions (A-alone, T-alone, and 5 A+T conditions with different phases). Error bars are ± 1 SD. For each condition, data are presented in order from left to right for S1, S2, Mean, S3, and S4. Means over the five A+T phase conditions are given at the far right of the plot.

Results obtained with the 250-Hz test frequency by the four subjects are shown in Fig. 1 (bottom). Percent-correct scores are given for each of the seven experimental conditions: A-alone, T-alone, and combined A+T with five different values of the starting phase of the tactile stimulus relative to that of the auditory stimulus (0˚, 72˚, 144˚, 216˚, and 288˚). Results are shown for individual subjects and for means across subjects (these values are provided in the Appendix). Average scores were 70.1± 3.6 %-Correct for A-alone and 71.2 ± 3.5 %-Correct for T-alone. For the five phase conditions, the A+T scores were: 77.3 ± 7.8 (0˚), 84.3 ±9.4 (72˚), 79.3 ±8.5 (144˚), 77.8 ± 5.7 (216˚), and 79.7 ± 6.1%-Correct (288˚) for an average of 79.6 ± 7.8%-Correct.

A mixed-model ANOVA with experimental condition (A-alone, T-alone, and the five phase values for A+T) as a fixed variable and subject as a random variable showed significant effects for condition [$F(6, 18)= 5.06$, p =0.003)] but not for subject [$F(3,18)=3.39$, p =0.04)]. There was a significant [$F(18, 84) = 2.25$, p =0.007) effect of interaction between subject and condition. The *post hoc* analysis showed no differences between the results of the conditions A-Alone and T-Alone. The A+T conditions with each of the five phases had significantly ($p<0.05$) better results than A-alone, while three of the phases (72˚, 144˚, 288˚) had significantly better results than T-alone. The only differences among the ten possible pairs of A+T conditions were significantly higher A+T performance for the 72˚ phase compared to 0˚ and 216˚. A *post hoc* analysis of the subject by condition interaction indicated that no significant differences were observed among the five A+T conditions for S1, S2, and S4; for S3, however, the score on the 72˚ condition was significantly higher than that on the 0˚ and 216˚ conditions.

A+T compared to A-alone and T-alone conditions. Slightly higher scores were observed at 250 Hz compared to 50 Hz across A comparison of results for the two test frequencies can be observed in Fig. 1. Averaged across subjects and runs, a similar pattern of performance was observed for both test frequencies showing higher scores for the conditions.

A repeated-measures ANOVA was conducted using two within-subjects variables of test frequency and experimental condition (with multiple repetitions within frequency/condition/subject). Significant effects were observed for the factor condition [$F(6, 18) = 9.34$, p=0.0001] but not for the factor frequency [$F(1, 3) = 0.99$, $p = 0.39$]. The interaction between condition and frequency was significant [$F(6, 18) = 7.11$, $p = 0.0005$]. The *post hoc* comparisons for condition indicated no significant difference between the A-alone and T-alone conditions and no significant differences among the A+T conditions with the exception of the 72˚ phase yielding higher results than 216˚. The *post hoc* comparisons for the condition by frequency interaction effect indicated that at 50 Hz there were no significant differences among scores across the five phase values but that at 250 Hz, results were higher at 72˚ than at 0˚ and 216˚. The main effect of subject [$F(3, 168) = 9.72$, p=0.0002] and the interaction between subject and frequency [$F(3, 168) = 11.5$, $p<0.0001$] were also significant. No other significant interactions were observed. *Post hoc* comparisons indicated that the main effect of subject arose from higher performance by S4 compared to the other three subjects. The frequency by subject effect arose from the fact that for two of the subjects (S2 and S4) the scores at 50-Hz and 250-Hz scores were not different from each other, whereas for the other two subjects (S1 and S3), scores at 250 Hz were significantly higher than at 50 Hz.

3.3 Comparisons to Model Predictions

Chi-Squared tests were performed in order to examine which of the three models (OSCM, PSM, or ASM) best fit the measured %-Correct scores (Sec. 2.6 of Methods). The number of observations in agreement with predictions for each subject and frequency is summarized in Table 2.

Table 2: *The Number of Observations in Agreement with Predictions. The total number of observations was 20 for each subject at each test frequency. The number in bold represents the model with the most observations that passed the Chi-Square test criteria (i.e., values less than 3.841).*

Frequency	Subject	OSCM	PSM	ASM
50 Hz	S1	**16**	15	7
	S2	8	13	**14**
	S3	**15**	13	4
	S4	10	**15**	13
	Sum	49	**56**	38
250 Hz	S1	11	**17**	15
	S2	**16**	13	5
	S3	**13**	11	9
	S4	12	**13**	12
	Sum	52	**54**	41

Table 2 shows that at 50 Hz the model with the highest proportion of observations in agreement with predictions was the OSCM for S1 and S3, the ASM for S2, and the PSM for S4. At 250 Hz, the best-fitting models were the PSM for S1 and S4 and the OSCM for S2 and S3.

The number of observations in agreement with predictions for each phase and frequency is summarized in Table 3.

Table 3: *The Number of Observations in Agreement with Predictions The total number of observations was 16 for each phase at each test frequency. The number in bold represents the model with the most observations that passed the Chi-Square test criteria (i.e., value less than 3.841).*

Frequency	Phase	OSCM	PSM	ASM
50 Hz	0	9	**12**	6
	72	9	**12**	9
	144	7	**10**	**10**
	216	**14**	10	5
	288	10	**12**	8
	Sum	49	**56**	38
250 Hz	0	**12**	9	8
	72	5	**8**	**8**
	144	**12**	10	6
	216	12	**14**	7
	288	11	**13**	12
	Sum	52	**54**	41

Table 3 shows that at 50 Hz the model with the highest proportion of observations in agreement with predictions was the PSM for 0˚, 72˚, and 288˚, a tie between PSM and ASM for 144˚, and the OSCM for 216˚. At 250 Hz, the best-fitting models were the OSCM for 0˚ and 144˚, the PSM for 72˚, 216˚, and 288˚ and the ASM for 72˚. In general, both at 50 Hz and 250 Hz the model with the highest proportion of observations in agreement with predictions was the PSM with the OSCM a close second and the ASM providing the least number of fits to the data.

The performance of the three models was also assessed by examining the ratio of the observed d' to the predicted D' for each test frequency. Mean ratios for individual subjects (averaged across phases and repetitions) are shown in Fig. 2 and mean ratios for each of the five phases (averaged over subjects and repetitions) are shown in Fig. 3. In these figures, a ratio of 1.0 indicates equal values of d' and D'; values >1 indicate that the model under-predicts the data; and values <1 indicate that the model over-predicts the data.

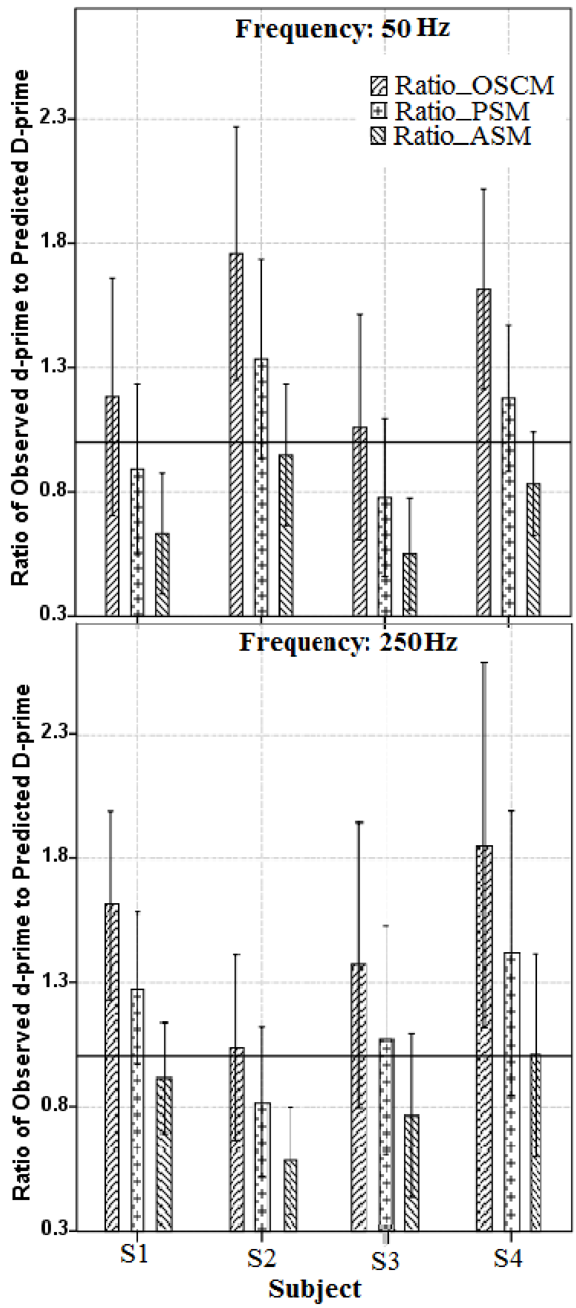

Figure 2. Ratio of observed d' to predicted D'. Mean ± 1 SD of ratios for individual subjects averaged over phases and repetitions are shown for 50-Hz (left panel) and 250-Hz (right panel) stimuli.

Individual subject results at 50 Hz (top of Fig. 2) indicate a similar pattern for S2 and S4 showing that the best fit is provided by the ASM model. The data are well-matched by the PSM for S1 and the OSCM for S3. At 250 Hz (bottom of Fig. 2), ratios closest to 1 are seen for the ASM for S1 and S4, the OSCM for S2, and the PSM for S3.

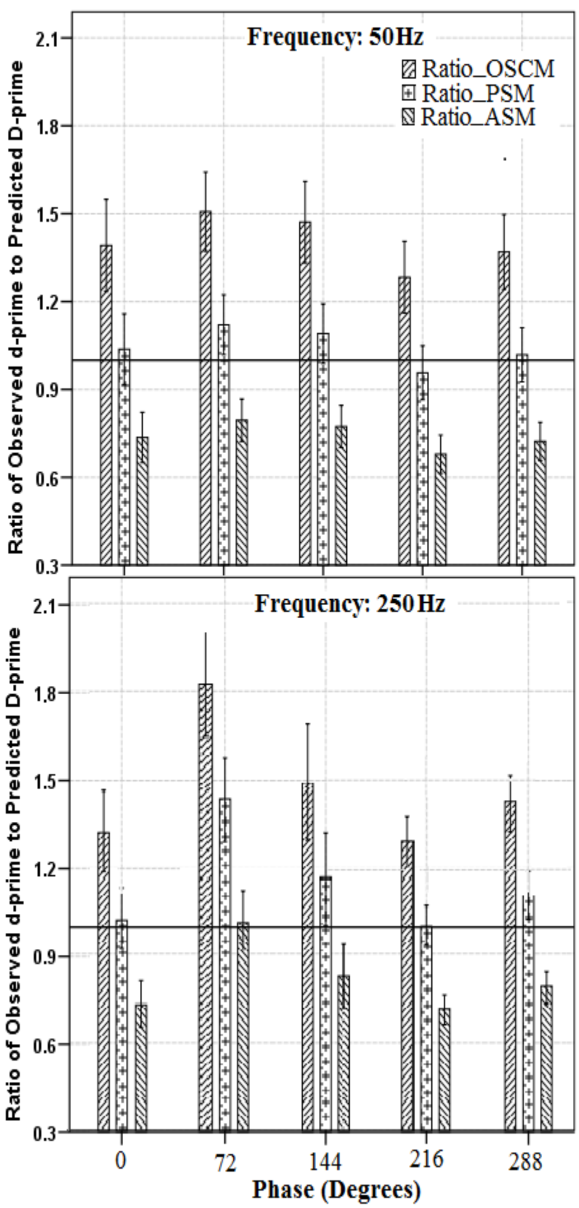

Figure 3. Ratio of observed d' to predicted D'. Mean ± 1 SD of ratios for each of the five phase conditions averaged over subjects and repetitions are shown for 50-Hz (top panel) and 250-Hz (bottom panel) stimuli.

For results shown as a function of phase, at 50-Hz (top of Fig. 3) the PSM provides the best fit for all five phases (0˚, 72˚, 144˚, 216˚, and 288˚). At 250-Hz (bottom of Fig. 3), the PSM provides the best fit at four phases (0˚, 144˚, 216˚, and 288˚) while the ASM provides the best fit at 72˚.

4. DISCUSSION

The current study extended the work of Wilson et al. [24] who found no significant effects of the relative phase of 250-Hz auditory and tactile sinusoidal signals on auditory-tactile integration using four phase values with 90˚-separation (0˚, 90˚, 180˚, and 270˚). In addition to retesting the Pacinian-range frequency of 250 Hz with a new group of subjects and a different set of phase differences, the current study also examined phase effects for a lower test frequency of 50 Hz which is outside the Pacinian range of the tactile sensory system. Although it was conjectured that phase effects may possibly be present at the lower test frequency (thus comparable to the behavior of the auditory system), no evidence was found for stronger phase effects at 50 Hz. Instead, the current results at 250 Hz showed some effects of relative phase contrary to the earlier results of Wilson et al. [24].

At 50 Hz, the *post hoc* statistical comparisons showed no significant differences in performance among the five relative phases employed in the A+T conditions, where the A+T scores averaged77.1%-Correct across phases with a range of 75.8-78.2%-Correct. At 250 Hz, the mean A+T performance averaged over the five phase values was 79.6%-Correct with a range of 77.3-84.2%-Correct. The A+T condition with a relative phase of 72˚ yielded the highest score which was statistically different in *post hoc* testing from two of the other A+T phase conditions: 0˚ and 216˚. The magnitude of this effect was relatively small: the difference between the score at 72˚ and the mean at the remaining phases was less than five percentage points. Furthermore, this effect appears to have been driven in large part by the performance of S3 who, in *post-hoc* testing of the significant interaction between subjects and conditions at 250 Hz, was the only subject for whom a significant difference was observed among the five A+T conditions. These observations suggest that the phase effect observed here at 250 Hz should be regarded with caution. There is no obvious explanation for the presence of a phase effect at 72˚ in the current set of data but not at a similar value of 90˚ in the Wilson et al. [24] data, for example. Further experiments are needed to determine if the result observed here can be repeated in a new group of subjects or instead has its basis in random fluctuations in performance.

Performance on the A+T conditions was compared to the predictions of three different models of multi-modal performance derived from observed measures of d' for A-alone and T-alone. The three models were evaluated and compared using Chi-Squared goodness-of-fit calculations as well as the ratio of the observed d' to the predicted D'. When Chi-Squared fits are examined in terms of phase at each test frequency (Table 3), the most immediate observation is that the ASM model provided a worse fit to the data than either the OSCM or the PSM. Thus, it appears that the A+T integration observed here is generally less than that predicted by an arithmetic sum of the single-channel d' values. In terms of the total number of fits, the OCSM resulted in slightly fewer fits overall than the PSM. The PSM had the most results in agreement (or was tied for most with another model) for four of the five phases at 50 Hz and three of the five at 250 Hz. In terms of d' ratios (Fig. 3), the PSM produced the best fits for four of the five phases at 50 Hz and four phases at 250 Hz. In Fig. 3, it can be seen that the OSCM typically under-predicts while the ASM over-predicts the observed A+T performance. These observations lead us to conclude that the PSM provides the best fit to the data at both test frequencies and implies that integration occurs across two independent sensory channels.

Results for individual subjects indicate that they may have employed different strategies for processing the simultaneous presentation of threshold-level auditory and tactile cues. For example, S4 appears to have integrated the cues from the two sensory modalities to achieve either Pythagorean or additive summation of cues from the two modalities at each of the two test frequencies. On the other hand, the A+T performance of S3 can generally be predicted by assuming that it arises from the use of the single-modality cue with the higher rate of detection (rather than on integrating both cues). The other two subjects appear to have employed different strategies of integration at the two different test frequencies. In particular, S1 was operating with the OSCM at 50 Hz but with the PSM at 250 Hz while S2 used ASM at 50 Hz and OSCM at 250 Hz. Across subjects, all three of the models were in evidence at 50 Hz and both OSCM and PSM at 250 Hz. Of course, it is possible that a given subject could have employed different integration strategies at different times. The generally larger SDs observed for the A+T phase conditions compared to SDs of the A-alone and T-alone conditions for each of the subjects at each frequency is also consistent with the possible use of different types of integration across runs.

5. SUMMARY AND CONCLUSION

In summary, the results at 50 Hz indicate similar A+T scores regardless of the relative phase of the signals in the two modalities. At 250 Hz, a significant effect was observed in that performance with a relative phase of 72° was higher than that at two of the other four phase values. In light of the relatively small size of the effect (less than five percentage points) and that it derives primarily from the performance of one subject, we regard this result with caution before conducting further tests. In general, A+T performance was best described as the Pythagorean sum of the single-modality scores. The phase-independence of the results at 50 Hz and for three of the four subjects at 250 Hz provides further support for integration of the auditory and tactual stimuli on the basis of their envelopes rather than fine-structure. Finally, the use of Pythagorean summation indicates that the signals from each modality are processed independently before being added at some more central location of the neural system.

ACKNOWLEDGMENTS

Research was supported by grant R01-DC00117 from the National Institutes of Health. The authors would like to thank Joseph Desloge and Timothy Villabona for developing the programs used to control the experiments.

NOTES

1. We denote measures of sensitivity that can be estimated directly from the data using lower case d' and those that are predicted by models by upper case letters D'.

REFERENCES

[1] Caclin, A., Soto-Faraco, S., Kingstone, A. and Spence, C. (2002) Tactile "capture" of audition. *Percept. Psychophys.* **64,** 616-630.

[2] Gillmeister, H. and Eimer, M. (2007) Tactile enhancement of auditory detection and perceived loudness. *Brain Res.* **1160,** 58-68.

[3] Ley, I., Haggard, P. and Yarrow, K. (2009) Optimal integration of auditory and vibrotactile information for judgments of temporal order. *J. Exp. Psychol. Hum. Percept. Perform.* **35,** 1005-1019.

[4] Occelli, V., Spence, C. and Zampini, M. (2009) The effect of sound intensity on the audiotactile crossmodal dynamic capture effect. *Exp. Brain Res.* **193,** 409-419.

[5] 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.

[6] Schnupp, J. W., Dawe, K. L. and Pollack, G. L. (2005) The detection of multisensory stimuli in an orthogonal sensory space. *Exp. Brain Res.* **162,** 181-190.

[7] 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.

[8] 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.

[9] Yarrow, K., Haggard, P. and Rothwell, J. C. (2008) Vibrotactile--auditory interactions are post-perceptual. *Perception* **37,** 1114-1130.

[10] Yau, J. M., Olenczak, J. B., Dammann, J. F. and Bensmaia, S. J. (2009) Temporal frequency channels are linked across audition and touch. *Curr. Biol.* **19,** 561-566.

[11] 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.

[12] 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.

[13] Shore, S. E., Vass, Z., Wys, N. L. and Altschuler, R. A. (2000) Trigeminal ganglion innervates the auditory brainstem. *J. Comp. Neurol.* **419,** 271-285.

[14] Smiley, J. F., Hackett, T. A., Ulbert, I., Karmas, 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.

[15] 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.

[16] Caetano, G. and Jousmaki, V. (2006) Evidence of vibrotactile input to human auditory cortex. *Neuroimage* **29,** 15-28

[17] 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.* **88,** 540-543.

[18] Fu, K. M., 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.

[19] Kayser, C., Petkov, C. I., Augath, M. and Logothetis, N. K. (2005) Integration of touch and sound in auditory cortex. *Neuron* **48,** 373-384.

[20] 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.

[21] Schroeder, C. E., Lindsley, 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.

[22] Schurmann, M., Caetano, G., Hlushchuk, Y., Jousmaki, V. and Hari, R. (2006) Touch activates human auditory cortex. *Neuroimage* **30,** 1325-1331.

[23] Shore, S. E., Koehler, S., Oldakowski, M., Hughes, L. F. and Syed, S. (2008) Dorsal cochlear nucleus responses to somatosensory stimulation are enhanced after noise-induced hearing loss. *Eur. J. Neurosci.* **27,** 155-168.

[24] Wilson, E. C., Reed, C. M. and Braida, L. D. (2009) Integration of auditory and vibrotactile stimuli: Effects of phase and stimulus-onset asynchrony. *J. Acoust. Soc. Am.* **126,** 1960-1974.

[25] Wilson, E. C., Reed, C. M. and Braida, L. D. (2010b) Integration of auditory and vibrotactile stimuli: Effects of frequency. *J. Acoust. Soc. Am.* **127,** 3044-3059.

[26] Wilson, E. C., Braida, L. D. and Reed, C. M. (2010a) Perceptual interactions in the loudness of combined auditory and vibrotactile stimuli. *J. Acoust. Soc. Am.* **127,** 3038-3043.

[27] Durlach, N. I. (1963) Equalization and cancellation theory of binaural masking-level differences. *J. Acoust. Soc. Am.* **35,** 1206-1218.

[28] Bolanowski, S. J., Gescheider, G. A.,

Verrillo, R. T. and Checkosky, C. M. (1988) Four channels mediate the mechanical aspects of touch. *J. Acoust. Soc. Am.* **84,** 1680-1694.

[29] Levitt, H. (1971) Transformed up-down methods in psychoacoustics *J. Acoust. Soc. Am.* **49,** 467-477.

[30] Braida, L. D. (1991) Crossmodal integration in the identification of consonant segments. *Q. J. Exp. Psychol. A.* **43,** 647-677.

[31] Neville, A. M. and Kennedy, J. B. (1966) Basic Statistical Methods for Engineers and Scientists. *(International Textbook, Scranton, PA)***,** p 133.

[32] Houtsma, A. J. M. (2004) Hawkins and Stevens revisited with insert earphones. *J. Acoust. Soc. Am.* **115,** 967-970.

[33] Reed, C. M. and Bilger, R. C. (1973) Acomparative study of S/No and E/No. *J. Acoust. Soc. Am.* **53,** 1039-1044.

[34] Gescheider, G. A., Bolanowski, S. J., Hall, K. L., Hoffman, K. E. and Verrillo, R. T. (1994) The effects of aging on information-processing channels in the sense of touch: I. Absolute sensitivity. *Somatosens. Mot. Res.* **11,** 345-357.

Appendix: *Summary of the results, Mean and SD, given for the four subjects, seven conditions, and two frequencies. Final column shows Mean and SD across subjects.*

Frequency	Condition	S1		S2		S3		S4		*Across subjects*	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
50 Hz	A	71.3	3.5	67.0	2.3	70.7	4.2	72.0	3.9	70.3	3.8
	T	69.0	2.3	68.0	3.3	69.0	4.0	72.7	3.4	69.7	3.5
	0	70.3	14.4	83.0	6.0	73.0	10.2	80.7	6.4	76.8	10.4
	72	77.7	2.3	80.7	7.7	69.3	9.2	85.3	2.4	78.3	8.2
	144	77.3	8.1	77.0	6.1	67.0	10.6	88.0	3.3	78.1	9.4
	216	75.0	12.2	75.7	8.5	73.7	7.3	78.7	9.2	75.8	8.7
	288	71.7	6.7	80.3	5.2	72.0	15.3	83.0	1.3	76.8	9.4
	Mean over 5 phases	74.4	9.2	79.3	6.6	71.6	9.8	83.1	5.8	77.1	9.1
250 Hz	A	68.0	2.9	70.3	1.3	70.7	3.9	71.4	5.4	70.1	3.6
	T	70.7	3.3	72.3	5.2	71.4	3.8	70.4	2.7	71.2	3.5
	0	80.0	5.5	70.0	4.2	74.0	8.2	85.3	2.9	77.3	7.8
	72	87.0	6.2	74.7	11.1	88.0	8.9	87.3	6.0	84.3	9.4
	144	80.7	3.2	73.0	5.9	73.7	2.0	89.7	8.4	79.3	8.5
	216	79.7	4.0	75.0	7.3	79.7	8.3	76.7	2.3	77.8	5.7
	288	85.7	4.3	75.0	5.8	77.3	5.8	80.7	3.5	79.7	6.1
	Mean over 5 phases	82.6	5.3	73.5	6.7	78.5	8.3	83.9	6.6	79.6	7.8