



*euronoise*

**Acoustics'08  
Paris**  
June 29-July 4, 2008

[www.acoustics08-paris.org](http://www.acoustics08-paris.org)

## **Resolution, spectral weighting, and integration of information across tonotopically remote cochlear regions: hearing-sensitivity, sensation level, and training effects**

Blas Espinoza-Varas

Communication Sciences & Disorders, OU Health Sciences Ctr., 825 NE 14th St., Oklahoma City, OK 73126-0901, USA  
blas-espinoza-varas@ouhsc.edu

This paper examines how listeners resolve, weight, and integrate redundant sensory information from tonotopically remote cochlear regions, one in the mid (M) the other in the high (H) frequencies. Normal- or impaired-hearing observers listened to two-component complexes and attempted to resolve and integrate simultaneous increments in frequency ( $\Delta F_M$ ) at  $\approx 1000$  Hz and in duration ( $\Delta T_H$ ) at  $\approx 3500$  Hz. Discrimination performance was studied as a function of hearing sensitivity, sensation level (SL) and training. Normal-hearing listeners tend to resolve both  $\Delta F_M$  and  $\Delta T_H$  and integrate the information if the components are equal in SL, but resolve only the louder component increment if the SLs are very unequal; i.e., spectral weighting is biased and integration is limited or nil. A similar pattern obtains when the low SL of H is caused by high-frequency sensorineural hearing loss. Once established, increasing the SL of H (e.g., with amplification) is often ineffective to rectify the weighting bias; doing so requires extensive discrimination training with the component that initially had low SL.

## 1 Introduction

Sensorineural hearing loss appears to interfere with the efficient integration of information conveyed by low-, mid-, and high-frequency regions [e.g., 6, 9, 14]. Inefficient cross-spectral integration would impair the ability to combine information across frequency channels, which is necessary because phoneme recognition depends on information conveyed by multiple acoustic cues distributed in time and across the speech frequency spectrum [e.g., 11, 12, 4]. A spectral integration deficit has been suggested by studies examining speech-recognition accuracy as a function of the number of bands representing the speech bandwidth [6, 10, 8, 14].

The factors accounting for inefficient cross-spectral integration are not fully understood. The available evidence is sparse and comes mainly from recognition tasks using speech stimuli; further research is needed to substantiate and characterize the integration deficits and to unravel their nature. In this study, spectral integration was studied using forced-choice discrimination tasks and well-controlled complex-tone stimuli. The goal was to determine how listeners learn to resolve and integrate redundant sensory information arising simultaneously from two tonotopically remote cochlear regions. The listeners' hearing sensitivity was normal in both regions or normal in one and impaired in the other.

While listening to a two-component complex, listeners with normal hearing or high-frequency sensorineural loss tackled the task of resolving simultaneous, redundant increments in the frequency of one component and in the duration of the other. Component frequencies were chosen so as to

stimulate concurrently a normal- and an impaired-sensitivity cochlear region of listeners with hearing loss. In short, the study focused on the abilities to resolve, integrate, and make efficient use of sensory information from remote cochlear regions having nearly equal or widely unequal sensitivities. These abilities were studied as a function of the sensation level (SL) of the high-frequency component and of the amount of training.

## 2 Method

### 2.1 Stimuli

With complex tones consisting of a mid- and a high-frequency (M and H) component, discrimination thresholds were measured for single increments in the frequency of M, the duration of H, or both simultaneously. Each complex comprised a mid-frequency (M) 1000-Hz, 80-ms sinusoid, and a high-frequency (H) 3127-Hz, 60-ms sinusoid, both gated with 4-ms raised-cosine rise-fall times (see Fig. 1). The components were tonotopically distant, inharmonic, and different in base duration (i.e., 60 versus 80 ms); thus, it was relatively easy to attend selectively to either M or H [13]. In order to minimize potential inharmonicity cues created by the increments in M frequency [1], non-harmonic component frequencies were used in all experiments. In addition, the components were relatively brief so that duration and frequency differences could be perceived as being nearly simultaneous.

### 2.2 Experimental Task

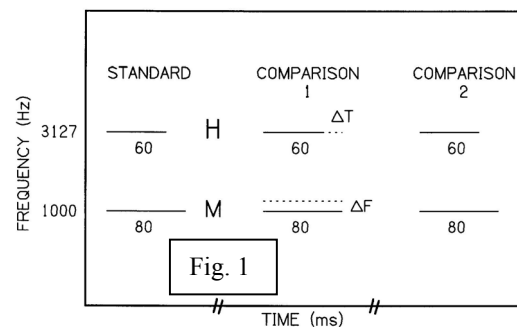


Fig. 1 depicts the structure of the 3I/2AFC trial; the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> observation interval displayed, respectively, a standard complex tone followed by comparisons 1 and 2. Either the 2<sup>nd</sup> or 3<sup>rd</sup> interval displayed a complex tone with increments in the frequency of M ( $\Delta F_M$ ), the duration of H ( $\Delta T_H$ ), or both ( $\Delta F_M - \Delta T_H$ ). Listeners had to report which comparison was different (in any respect) from the standard complex tone; the trial-by-trial feedback on the CRT monitor scored the response as correct or in error.

Within a trial block, eight  $\Delta F_M$  values (0, 1, 2, 4, 5, 6, 8, or 10 Hz) were crossed with eight  $\Delta T_H$  values (0, 1, 3, 5, 7, 9, 11 or 13 ms) yielding a total of 64 combinations of  $\Delta F_M$  and  $\Delta T_H$ . On any given 126-trial block, each of the 64 possible combinations was presented twice, excepting the two trials

in which  $\Delta T_H$  and  $\Delta F_M$  were zero. Since both ranges included zero, a 126-trial block comprised 98 trials (78 percent) with increments in both components,  $\Delta F_M\text{-}\Delta T_H$ , and 28 trials (22 percent) with increments on a single component: 14 with  $\Delta F_M$ , and 14 with  $\Delta T_H$ . That is, the increments were redundant in 78 percent of trials, but the  $\Delta T_H$  and  $\Delta F_M$  values were uncorrelated. The particular  $\Delta T_H\text{-}\Delta F_M$  combination and the observation interval in which it occurred were chosen at random, with two restrictions: each increment combination occurred only twice within a 126-trial block, and had 0.5 probability of being in the 2<sup>nd</sup> or 3<sup>rd</sup> interval. This method encourages normal-hearing listeners to share attention over the two dimensions [2, 3, 5]. Prior to working with combined  $\Delta F_M\text{-}\Delta T_H$  increments, participants underwent training in  $\Delta T_H$ -only or  $\Delta F_M$ -only conditions with the respective component presented in isolation; this training helped to expedite learning.

### 2.3 Equipment

A Macintosh IIfx computer and LabVIEW signal-processing software synthesized (20,000 samples/sec) and shaped the waveforms. The electrical signals generated by separate but synchronous 12-bit D/A converters were low-pass filtered at 8000 Hz, mixed, amplified (Crown D-40), and attenuated. Supraural earphones transduced the electrical signal into a pressure wave. LabVIEW software specified the increments in frequency at 1000 Hz ( $\Delta F_M$ ), and those in duration at 3127 Hz ( $\Delta T_H$ ); the Crown amplifier controlled the overall level of the complex (50 dB SPL). A Quest-155 level meter and a 6-cm<sup>3</sup> coupler calibrated the A-weighted SPL level of individual components and of the two-component complex; a spectrum analyzer and an oscilloscope calibrated the components frequency and duration. The experiments took place in a sound-absorbent room; participants entered responses on a computer keyboard, in accordance with oral and written instructions, while getting trial-by-trial feedback on a CRT monitor. The Macintosh IIfx computer and Lab View software controlled all experimental events.

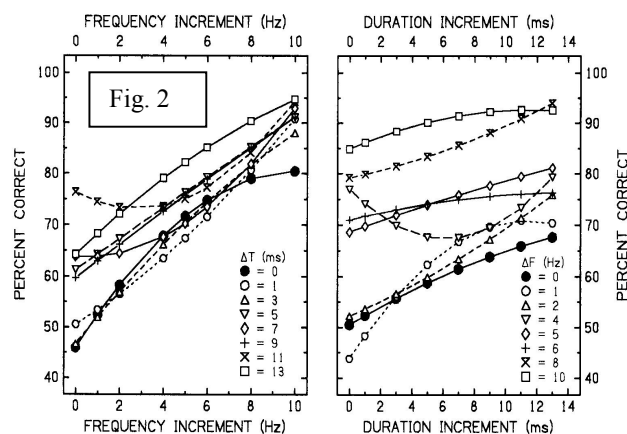
## 3 Experiment I. Equal sensation-level components and normal sensitivity

In normal-hearing listeners, this experiment assessed integration of information conveyed by components presented at optimal sensation level (both at 41 dB). Listeners were trained first to discern single increments in the frequency of M or in the duration of H, with each component in isolation. Training began with duration discrimination in three participants, and with frequency discrimination in the remaining three. Discrimination of frequency-duration increment combinations was studied only after participants achieved 70-80% correct discrimination with increments in a single component presented in isolation. One male and five females ages 23-54 years (mean = 30.8) all with normal hearing sensitivity were paid \$12.00/hr for listening.

## 3.1 Results

**Psychometric functions.** Data were summarized as 8-point psychometric functions relating percent correct discrimination to  $\Delta F_M$  with  $\Delta T_H$  as parameter, or to  $\Delta T_H$  with  $\Delta F_M$  as parameter; each data point is the average of 40-70 trials. Least-RMS-error regression was used to approximate and smooth the raw percent-correct data with second-order polynomials; in choosing these functions, the goal was merely to maximize the goodness of fit. Overall, the approximation was good yielding Pearson correlation coefficients averaging  $\approx 0.91$ .

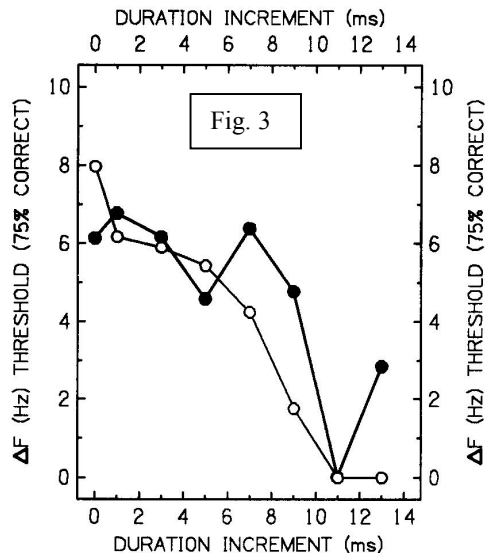
Data from conditions presenting  $\Delta F_M$  and  $\Delta T_H$  combined are shown in Fig. 2 below. For one listener, polynomial functions relating percent-correct discrimination to  $\Delta F_M$  with  $\Delta T_H$  as parameter are shown in Fig. 2 (left panel). The main effects are seen in the slope and Y-axis intercepts: a) in the function for  $\Delta F_M$  alone, the threshold and slope provide estimates of resolving power for M frequency increments; and b) in the functions for  $\Delta F_M\text{-}\Delta T_H$  combined, the parametric elevation estimates the ability to integrate information across the two frequencies (M and H).



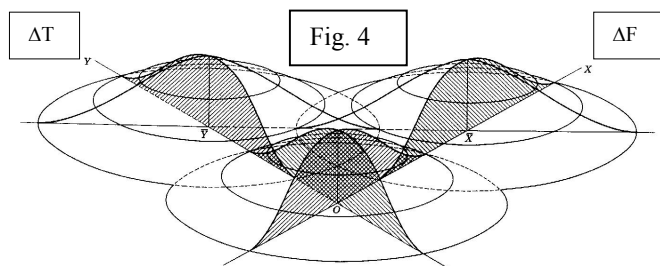
When presented in combination with increments in H duration, the psychometric functions for increments in M frequency showed an orderly elevation proportional to the size of the duration increment. That is, the duration increments improved the discrimination accuracy achieved with increments in frequency alone. Much the same trend obtains when percent correct discrimination is plotted as a function of  $\Delta T_H$  with  $\Delta F_M$  as parameter (Fig. 2, right panel).

**Discrimination Thresholds.** From the polynomial functions (Fig. 2),  $\Delta F_M$  discrimination thresholds were interpolated at the 75% correct level, for the  $\Delta F_M$ -only function and each of the seven functions with  $\Delta F_M$  and  $\Delta T_H$  combined. Fig. 3 shows the 75%-correct  $\Delta F_M$  discrimination thresholds (solid circles) as a function of  $\Delta T_H$ , as well as the thresholds predicted (open circles) by the Signal Detection model [7] described below. Each data point specifies the values of  $\Delta F_M$  and  $\Delta T_H$  that in combination yield 75% correct discrimination (thresholds for  $\Delta F_M$  only are those measured with  $\Delta T_H = 0$  ms). The function relating the  $\Delta F_M$  threshold to  $\Delta T_H$  has negative slope: the steeper the slope the

more efficient the resolution and integration of the duration information conveyed by H. The "trade-off" showing the decrease in  $\Delta F_M$  thresholds with increasing  $\Delta T_H$  suggests spectral integration of information; this trade off obtained in practically all subjects.



### 3.2 Optimal cross-spectral integration



The Signal Detection model [7] can be used to predict discrimination accuracy for simultaneous increments in frequency and duration. This condition would produce probability distributions along two dimensions, one for each stimulus increment ( $X = \Delta F_M$  and  $Y = \Delta T_H$ ). Deciding whether the first or second comparison is different from the standard would depend on optimal combination of the likelihood ratios associated to  $\Delta F_M$  and  $\Delta T_H$ . Discrimination accuracy is proportional to the distance between the distributions means. In the Fig-4 diagram and Eq. (1) below, the correlation between the  $\Delta F_M$  and  $\Delta T_H$  sensory processes is assumed to be zero (i.e., the two processes are independent). The distance between the mean of the sensory-effect distribution for combined differences in frequency and duration, and the mean of the distribution for no stimulus differences would be equal to:

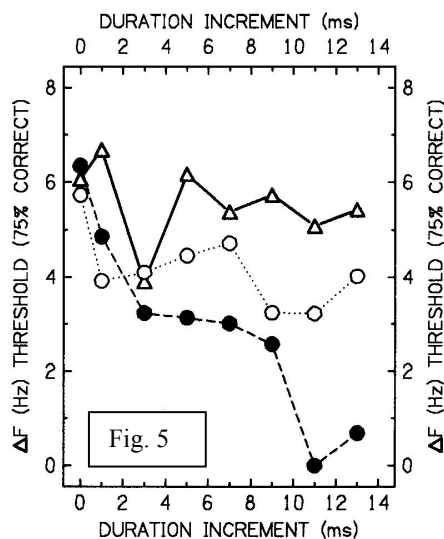
$$d'_{\Delta F, \Delta T} = \sqrt{(d'^2_{\Delta F} + d'^2_{\Delta T})} \quad (1)$$

To compute Eq. (1), the percent correct values associated to  $\Delta F_M$  only and  $\Delta T_H$  only were expressed as probabilities and converted to  $d'$  values using a table of areas under the normal probability distribution. The predicted  $d'$  value for a  $\Delta F_M$ - $\Delta T_H$  combination was converted back to percent correct and used to generate predicted psychometric functions. These functions were approximated with second-order polynomials from which predicted discrimination thresholds were estimated at the 75% correct level. The percent correct values used to compute Eq. (1) were those of the training conditions with frequency or duration increments in components presented in isolation; thus, the prediction is based on the best discrimination performance achieved with increments on each dimension. Fig. 3 shows close agreement between the observed and predicted thresholds.

### 4 Experiment II. Unequal sensation-level components and normal sensitivity

In this experiment, resolution and integration of simultaneous increments in M frequency and H duration were studied as a function of the sensation level (SL) in the high-frequency component (H) and of training. The stimuli and task were as in Exp. I, but at the beginning of the experiment the sensation level was very low for H and nearly optimal for M; thereafter, a condition with optimal SL in both components was studied. This test order attempted to simulate high frequency hearing loss, prior to and following high-frequency amplification. Does the improvement in discrimination accuracy obtain instantly after increasing the H sensation level or does it require extra special training?

The experiment comprised four stages: 1) "Unequal SL," M and H had a sensation level of 41 and 1.5 dB, respectively; that is, audibility was optimal for M and poor for H. 2) "Equal SL," M and H at 41 dB SL; 3) " $\Delta T_H$  discrimination training," aimed at improving  $\Delta T_H$  resolution with the H presented in isolation at 41 dB SL. 4) Retest of stage 2.

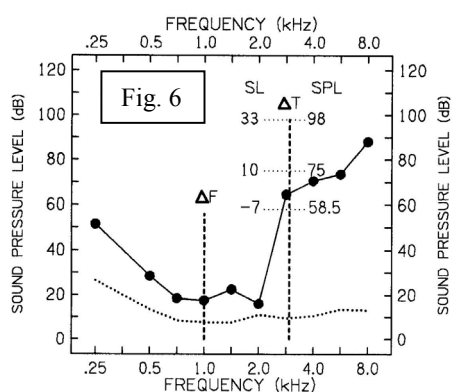


**Discrimination Thresholds.** For one listener, the trade-off functions relating the 75% correct  $\Delta F_M$  threshold to the size of  $\Delta T_H$  are shown in Fig. 5. Thresholds measured in stages 1, 2, and 4 are depicted by triangles, open, and solid circles, respectively. In Stage 1, the 75% correct  $\Delta F_M$  threshold did not decrease with  $\Delta T_H$ , meaning that the listener extracted little or no information from  $\Delta T_H$ . In stage 2, the  $\Delta F_M$  threshold did decrease with  $\Delta T_H$ , but the trade-off function slope was not steep; that is, listeners did extract information from  $\Delta T_H$ , but not with optimal efficiency. In Stage 3, the trade-off function slope is steep, consistent with efficient extraction and use of  $\Delta T_H$  information.

The results shown in Fig. 5 suggest that once the listener learns to resolve  $\Delta F_M$  with unequal-SL complex components, increasing the audibility of H does not necessarily yield efficient extraction and use of  $\Delta T_H$  information. Accomplishing the latter requires training the listener to resolve  $\Delta T_H$  with the H component presented in isolation.

## 5 Experiment III. Unequal sensation-level components and impaired sensitivity

This Exp. employed the same stimuli and task of the Exp. II, but the listeners had moderate-to-severe high-frequency hearing loss, of a sensorineural nature. Exp. III comprised four training stages: 1) "no H amplification," 2) "H amplification alone," 3) " $\Delta T_H$  discrimination training" with H presented in isolation, and 4) retest of stage 2. The 1000-Hz component was 39-48 dB SL in all four stages; the 3127-Hz component was -7 dB SL in stage 1, and 33 dB in stages 2, 3, and 4.



One male and three female college students, 20-44 years old (mean = 30), participated in Exp. III; they had normal tympanograms, air-bone gaps no greater than 10 dB, and medical histories consistent with those of sensorineural hearing loss. On average, their detection thresholds were  $\geq 50$  dB SPL above 2000 Hz and within the normal range below 1500 Hz. For one listener, Fig. 6 depicts the audiogram in SPL units, the line spectra of the stimulus complex, and the levels of M and H in the various experimental conditions. To ensure that central nervous system reorganization had taken place, only post-lingual

hearing-impaired participants experiencing hearing difficulties for at least five years were enrolled in the study. Participants were paid \$ 9.00/hr for listening.

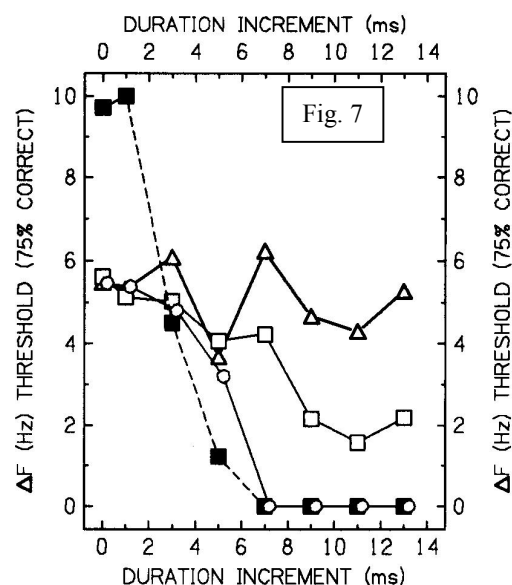


Fig. 7 shows 75%-correct  $\Delta F_M$  discrimination thresholds as a function of  $\Delta T_H$  measured in three conditions: no amplification (triangles), and optimal-SL amplification prior to (open squares) and post  $\Delta T_H$  discrimination training at optimal-SL (solid squares). Also shown are the thresholds predicted by the optimal integration model described above (open circles). Each data point specifies the values of  $\Delta F_M$  and  $\Delta T_H$  that in combination yield 75% correct discrimination (thresholds for  $\Delta F_M$  only are those measured with  $\Delta T_H = 0$  ms). The "trade-off" functions reveal how the  $\Delta F_M$  threshold decreases as the  $\Delta T_H$  resolution improves with amplification and training. In both amplification conditions, a trade off obtained between the  $\Delta F_M$  and  $\Delta T_H$  values that in combination yielded 75% correct. However, the optimal integration predicted by the model was approached only after training the listener to resolve  $\Delta T_H$  with H in isolation; amplification alone is not sufficient to approach the optimum. Post training, an elevation in the  $\Delta F_M$  discrimination threshold obtained, suggesting between channel interference.

## 6 Conclusion

If the components have equal sensation level, normal-hearing listeners tend resolve and integrate the information conveyed by the mid- and high-frequency component. To achieve 75% correct discrimination, a trade off obtains between frequency increments at 1000 Hz and duration increments at 3127 Hz. Observed performance is in good agreement with the Signal Detection Theory prediction. If the components have widely unequal sensation levels, a spectral bias obtains and integration is minimal or absent: normal-hearing listeners resolve the high but not the low sensation-level component; thus, frequency and duration

increments exhibit little or no trade off. Thereafter, making the sensation level optimal for both components is not sufficient to rectify the spectral bias; achieving this requires prolonged duration discrimination training with a high sensation-level, high frequency component presented in isolation. Much the same trends obtain when the unequal sensation levels are caused by sensorineural sensitivity deficits.

## Acknowledgments

This project was supported by research grants from the Presbyterian Health Foundation and from the College of Allied Health, OU Health Sciences Center (PI: B. Espinoza-Varas).

## References

- [1] C. J. Darwin, R. W. Hukin, "Perceptual segregation of a harmonics from a vowel by interaural time difference in conjunction with mistuning and onset asynchrony," *J. Acoust. Soc. Am.* 103, 1080-1084 (1998)
- [2] B. Espinoza-Varas, "Integration of spectral and temporal cues in discrimination of nonspeech sounds: A psychoacoustic analysis." *J. Acoust. Soc. Am.* 74, 1687-1694 (1983)
- [3] B. Espinoza-Varas, D.G. Jamieson, "Integration of spectral and temporal cues separated in time and frequency." *J. Acoust. Soc. Am.* 76, 732-738 (1984)
- [4] B. Espinoza-Varas, "Levels of representation of phonemes and bandwidth of spectral and temporal integration," in *The Psychophysics of Speech Perception*, edited by M.E.H. Schouten (M. Nijhoff, The Netherlands), pp. 80-90 (1987)
- [5] B. Espinoza-Varas, M. Bingabr, P. Loizou, "Amplification and training effects in resolution and cross-spectral integration of redundant cues," *J. Acoust. Soc. Am.* 121, 3184 (2007)
- [6] K.W. Grant, J.B. Tufts, S. Greenberg, "Integration efficiency for speech perception within and across sensory modalities by normal-hearing and hearing-impaired individuals," *J. Acoust. Soc. Am.* 121, 1164-1176 (2007)
- [7] D.M. Green, J.A. Swets, *Signal Detection Theory and Psychophysics*, John Wiley and Sons, New York (1966)
- [8] E.W. Healy, S.P. Bacon, "Across-frequency comparison of temporal speech information by listeners with normal and impaired hearing," *J. Speech Lang. Hear. Res.* 45, 1262-1275 (2002)
- [9] E.W. Healy, A. Kannabiran, S.P. Bacon, "An across-frequency processing deficit in listeners with hearing impairment is supported by acoustic correlation," *J. Speech Lang. Hear. Res.* 48, 1236-1242 (2002)
- [10] F.J. Hill, L.P. McRae, R.P. McClellan, "Speech recognition as a function of channel capacity in a discrete set of channels," *J. Acoust. Soc. Am.* 44, 13-18 (1968)
- [11] D.H. Klatt, "Voice onset time, frication, and aspiration in word initial consonant clusters," *J. Speech Lang. Hear. Res.* 18, 686-706 (1975)
- [12] B.H. Repp, "Phonetic trading relations and context effects: New Experimental evidence for a speech mode of perception," *Psychol. Bull.* 92, 81-110 (1982)
- [13] B. Roberts, S.D. Holmes, "Asynchrony and the grouping of vowel components: Captor tones revisited," *J. Acoust. Soc. Am.* 119, 2905-2918 (2006)
- [14] C.W. Turner, S.L. Chi, S. Flock, "Limiting spectral resolution in speech for listeners with sensorineural hearing loss," *J. Speech Lang. Hear. Res.* 42, 773-784 (1999)