Detection of Acoustic-Phonetic Landmarks in Mismatched Conditions Using a Biomimetic Model of Human Auditory Processing

Sarah King *Mark hasegawa – Johnson*
University of Illinois, Urbana, Illinois
sborys@illinois.edu,jhasegaw@illinois.edu

ABSTRACT
Acoustic-phonetic landmarks provide robust cues for speech recognition and are relatively invariant between speakers, speaking styles, noise conditions and sampling rates. The ability to detect acoustic-phonetic landmarks as a front-end for speech recognition has been shown to improve recognition accuracy. Biomimetic inter-spike intervals and average signal level have been shown to accurately convey information about acoustic-phonetic landmarks. This paper explores the use of inter-spike interval and average signal level as input features for landmark detectors trained and tested on mismatched conditions. These detectors are designed to serve as a front-end for speech recognition systems. Results indicate that landmark detectors trained using inter-spike intervals and signal level are relatively robust to both additive channel noise and changes in sampling rate.

KEYWORDS: Auditory Modeling, Acoustic-Phonetic Landmark, Mismatched Conditions.

Proceedings of COLING 2012: Posters, pages 589–598,
COLING 2012, Mumbai, December 2012.

1 Introduction

Mismatched conditions — differences in channel noise between training audio and testing audio — are problematic for computer speech recognition systems. Signal enhancement, mismatch-resistant acoustic features, and architectural compensation within the recognizer are common solutions (Gong, 1995). The human auditory system implements all three of these solutions by 1.) enhancing the speech signal via the filtering of the head, outer ear, and basilar membrane, 2.) extracting prominent, noise-resistant information from the speech signal, and 3.) implementing dereverberation and noise reduction mechanisms within the cellular architecture of the brain.

Commercial speech recognizers must inevitably deal with mismatched conditions. Such mismatches may include additive channel noise or loss of frequency information. Both of these events occur in the telephone channel. Telephone-band speech recognition (8 KHz) is a difficult task (Bourlard, 1996; Karray and Martin, 2003). Both Gaussian systems (Chigier, 1991; Moreno and Stern, 1994) and non-Gaussian systems (Hasegawa-Johnson et al., 2004) trained on telephone-band speech are not as accurate as systems trained on wide band speech (16 KHz) (Halberstadt and Glass, 1998). This may indicate that a speech recognition system should compensate for channel anomalies before the decoding phase.

The distinctive features [silence, continuant, sonorant, syllabic, and consonantal] are binary valued phonetic descriptors (Stevens, 1999). For example, a sound can either be produced in the nucleus of a syllable ([+syllabic]) or not ([−syllabic]). The vowel /æ/ is a [+syllabic] sound and the consonant /d/ is a [−syllabic] sound. A transition between the two sounds, as in the word "add," is a [+−syllabic] landmark. Detection and recognition of acoustic-phonetic landmarks as a front-end to an HMM-based speech recognition system improves both phone and word recognition accuracy on telephone-band speech (Borys and Hasegawa-Johnson, 2005; Borys, 2008). Landmark-based systems generalize accurately to noisy and mismatched conditions (Kirchhoff, 1999; Juneja and Espy-Wilson, 2004).

Models of the auditory periphery have been used for denoising/enhancing speech (Hunt and Lefebvre, 1989; Virag, 1999), speech recognition in clean (Cohen, 1989; Hunt and Lefebvre, 1989; Ghitza, 1994; Ying et al., 2012) and noisy conditions (Kim et al., 1999), and emotion recognition (Ying et al., 2012). When applied as a front-end, models of the auditory periphery improve speech recognition accuracy (Cohen, 1989; Hunt and Lefebvre, 1989; Ghitza, 1994; Virag, 1999), however, such systems fail to achieve human performance. Current auditory models primarily mimic the cochlea and auditory nerve, both ignoring the effects of head-related filtering and failing to account for neural processing in the brainstem. Neurologists have proposed that the processing in auditory brainstem nuclei, such as the cochlear nucleus and lateral lemniscus, may improve the robustness of human speech recognition to changes in environment (Ehret and Romand, 1997; Winer and Schreiner, 2005; Schnupp et al., 2011).

Both landmark detection and auditory modeling improve recognition accuracy when used as front-ends for speech recognition systems operating in mismatched conditions. This paper proposes an approach that unifies the two methods.

2 Data For Mismatched Speech Recognition

The TIMIT corpus (Garofolo et al., 1993) contains 6300 phonetically rich sentences collected from 630 different male and female speakers from 8 different dialect regions of the United

States. Each utterance is sampled at a rate of 16 KHz. NTIMIT (Jankowski et al., 1990) was constructed by filtering the original TIMIT utterances through a telephone channel and then downsampling to 8 KHz. TIMIT contains detailed, orthographic phonetic transcriptions. NTIMIT is time aligned with TIMIT such that the original transcriptions describe the NTIMIT utterances.

3 The Auditory Model

A diagram of the complete binaural auditory model is shown in Figure 1. Relevant parts of the model are highlighted.

The head, outer, and middle ear are modeled using Tidemann's head-related transfer function (HRTF) measurements (Tidemann, 2011). The output of the HRTF is a set of two acoustic signals — the sound as it is heard at both the left and right ears. HRTF output inputs to the basilar membrane (BM) model.

The BM is modeled using a bank of 2760 parallel gammatone filters. The design of the gammatone filters mimics that in (Patterson and Holdsworth, 1996). The central frequencies of the filters are spaced according to the ERB (equivalent rectangular bandwidth) (Moore and Glasberg, 1983) scale, with 100 filters per ERB, arranged at center frequencies between $f_1 = 60$Hz and $f_{2760} = 8000$Hz. The filter outputs at each time t can be placed side by side to form a topographic map of BM motion from time $t = 0$ to time $t = T$. Examples of such maps for the vowel /i/ are shown in Figure 2.

The location on the BM with the maximal vertical displacement corresponds to a maximum in acoustic pressure (Geisler, 1998) which in turn corresponds to the frequency of the acoustic stimulus (Bekesy, 1960). In Figure 2, a spatio-temporal maximum is equivalent to a pressure maximum. The biomimetic model assumes that processes of lateral inhibition (e.g., as in (Greenberg, 1988)) prevent auditory nerve (AN) fibers from responding to sub-maximal inputs, so that any individual nerve fiber fires only when it is aligned with a spatio-temporal pressure maximum. This aspect of the model is not physiologically accurate, but this approximation is extremely useful for controlling the computational complexity of the biomimetic model. A minimum displacement is required for the inner hair cells (IHCs) to fire. Figure 3 shows which IHCs fire in response to the BM filter outputs of Figure 2. In the figure, an "x" indicates that the IHC corresponding to a given frequency fired at time t. Intensity information is not shown in the figure. A spectrogram of the same audio data used to create Figures 2 and 3 is shown in Figure 4.

The intensity level can be calculated directly from the displacement of the BM model.

$$I(t, f_m) = 20 \log_{10} \frac{y_m(t)}{Y_{ref}} \tag{1}$$

Here, $I(t, f_m)$ is the intensity level in decibels in the m^{th} frequency band at time t, $y_m(t)$ is the observed output at time t from the m^{th} filter given that a maximum has been found, and Y_{ref} is the threshold of hearing of the BM model.

Level, frequency, and timing information for the duration of the acoustic signal are stored as a sparse binary third order tensor A. An individual entry in A is referenced by its time, frequency, and intensity level values and $A(t, f, i) \in \{0, 1\}$. When $A(t, f, i) = 1$, the neuron has fired at time t for an auditory signal at frequency f (in Hz) having level i dB. A value of 0 indicates that

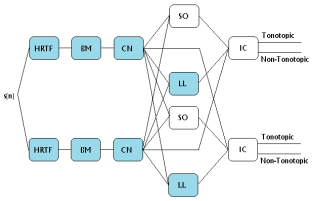

Figure 1: The complete auditory model. The model filters the signal $s[n]$ through a head-related transfer function (HRTF). The HRTF produces a two-channel audio signal that is filtered by the basilar membrane (BM) model. The BM model innervates a cochlear nucleus (CN) model. The CN model innervates a superior olive (SO) model, a lateral lemniscus (LL) model, and a model of the inferior colliculus (IC). The SO model inputs to the LL and the IC. The LL model inputs to the IC. The IC outputs tonotopic and non-tonotopic acoustic features. Only the HRTF, BM, and parts of the CN and LL are described in this paper.

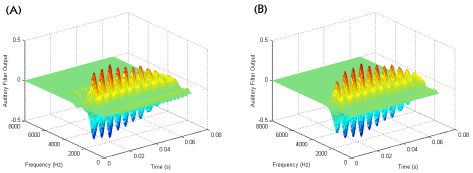

Figure 2: **(A.)** A topographic map of basilar membrane (BM) displacement as a function of time and frequency for the vowel /i/ (as in the word "she") for speaker MMDB0 from the TIMIT corpus. **(B.)** A topographic map of basilar membrane (BM) displacement as a function of time and frequency for the vowel /i/ for speaker MMDB0 from the NTIMIT corpus. For both **(A.)** and **(B.)**, the x-axis is the time in seconds. The y-axis is frequency in Hertz. The z-axis is the amplitude of the the auditory filter outputs.

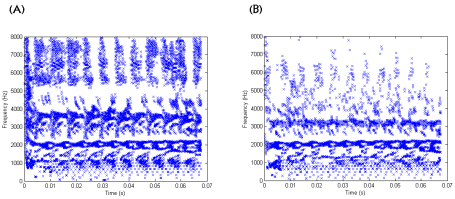

Figure 3: **(A.)** Inner hair cell (IHC) activation as derived from the tonotopic map of the vowel /i/ produced by the speaker MMDB0 from the TIMIT corpus shown in Figure 2A. **(B.)** Inner hair cell (IHC) activation as derived from the tonotopic map of the vowel /i/ produced by the speaker MMDB0 from the NTIMIT corpus shown in Figure 2B. In both **(A.)** and **(B.)**, an "x" indicates that the IHC corresponding to a given frequency (y-axis) has fired at time t (x-axis). The y-axis is spaced according to the ERB scale.

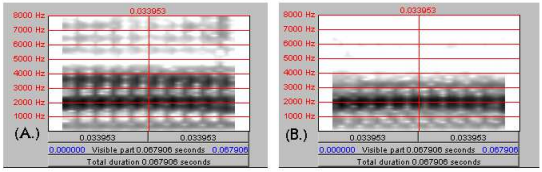

Figure 4: **(A.)** A spectrogram of the vowel /i/ produced by speaker MMDB0 from the TIMIT corpus. **(B.)** A spectrogram of the vowel /i/ produced by speaker MMDB0 from the NTIMIT corpus. corresponding inner hair cell (IHC) maps are shown in Figures 3A and 3B, respectively.

the neuron has not fired at time t. The majority of entries in A will be equal to 0 for any given speech utterance. The left and right ears of the model each produce their own independent tensors (A^l and A^r, respectively), though, for current experiments only the left channel is used. The tensor A is analogous to the information transmitted via the AN.

The octopus cells of the cochlear nucleus (CN) detect synchrony among AN fibers. The tensor A is a more sparse representation than the representation transmitted by the physiological AN. While it is known that octopus fibers are innervated by many AN fibers and that the organization of these fibers is tonotopic, the exact organization of the innervating fibers is unknown. To combat this problem, synchrony is detected using the logical union ($\cup_i$) of the binary variables $A(t,f,i)$ over different values of i, and summing over a time-frequency window of duration T_w and over F_w frequency bands. In other words,

$$S_w(t,f) = \sum_{\tau=0}^{T_w-1} \sum_{\phi=0}^{F_w-1} \cup_{i=I_{min}}^{I_{max}} A(t-\tau, f-\phi, i) \tag{2}$$

and

$$O_w(t,f) = \begin{cases} 1 & S_w(t,f) > \rho \\ 0 & \text{otherwise} \end{cases}$$

where ρ is the minimum number of active neurons in a window w required for the octopus cell to fire. The optimum window size was determined experimentally to be 3 ms by 60 neural inputs (0.6 ERB) with an optimum firing threshold of $\rho = 2$. The frequency step is 0.2 ERB. Each octopus cell overlaps 0.4 ERB with its neighbor. The time step is 1 ms.

The lateral lemniscus (LL) model determines the rate $R_{O_w(t,f)}$ at which the octopus cells corresponding to frequency f fire at time t. The rate (inverse inter-spike interval) is determined as follows

$$R_{O_w(t,f)} = \frac{1}{\tau(O_w(t_m,f)) - \tau(O_w(t_n,f))} \tag{3}$$

where $\tau(O_w(t,f)) = t$, and $O_w(t_m,f)$ and $O_w(t_n,f)$ are two chronologically ordered, nonzero instances of octopus cell activation, i.e., $t_m > t_n$.

The multipolar neurons of the CN calculate the average spectral level of synchronous frequencies. The average spectral level is calculated as follows

$$M_w(t,f) = \frac{\sum_{\tau=0}^{T_w-1} \sum_{\phi=0}^{F_w-1} I(t-\tau, f-\phi)}{T_w F_w} \tag{4}$$

Level is summed from all active nerve fibers in a time-frequency window of duration T_w and over F_w frequency bands. The multipolar cells use the same window as the octopus cells (3 ms by 0.6 ERB) to calculate the average level. The average level feature $M_w(t,f)$ differs from the Mel-frequency spectral coefficients (MFSCs) in at least two ways, and may therefore encode complementary information: 1.) $M_w(t,f)$ averages the log magnitude, whereas MFSC are the logarithm of an averaged magnitude, and 2.) $M_w(t,f)$ averages only detected peaks, whereas MFSC averages all signal components. These properties of the auditory model may make it more resistant to additive noise and channel-dependent filtering.

4 Experiments

Ten support vector machines (SVMs) were trained using the TIMIT corpus to detect the landmarks listed in Table 1. The training set consisted of the SX TIMIT audio files. The two separate test sets contained the SI files from TIMIT and the SI files from NTIMIT. A total of 10000 training tokens (5000 landmark tokens and 5000 non-landmark tokens) were extracted from the TIMIT training data. A total of 8000 tokens (4000 landmark tokens and 4000 non-landmark tokens) were extracted for each of the test sets. No tokens overlap between the training and testing sets.

For the training set and each of the test sets, Mel-frequency cepstral coefficients (MFCCs), the neural firing rate and level features described in this paper (NRL), and a combination set of the MFCCs and NRLs (MFCCNRL) were calculated. The MFCCs were calculated using a 25 ms window with a time-step of 5 ms. NRL features are calculated every millisecond to match the maximal firing rate of the octopus cells of the CN. The NRL feature vector is a 276 dimensional vector composed of 138 instances of $O_w(t,f)$ and 138 instances of $M_w(t,f)$ for the window w at time t.

The training and testing data for each landmark detector SVM in Table 1 consist of feature vectors $\vec{x}_t$ containing 11 concatenated acoustic feature frames. The first frame in $\vec{x}_t$ was sampled at 50 ms before the landmark, the 6th frame was sampled at the landmark time t, and the 11th frame was sampled at 50 ms after the landmark; i.e., $\vec{x}_t \equiv [\vec{y}_{t-50}, \ldots, \vec{y}_t, \ldots, \vec{y}_{t+50}]$ where $\vec{y}_t$ included either MFCCs, NRLs, or a combination of MFCCs and NRL features (MFCCNRL). In other words, $\vec{x}_t$ is created by concatenating n acoustic feature frames on both sides of the landmark frame $\vec{y}_t$, where the time step between frames is 10 ms and the total number of concatenated frames in $\vec{x}_t$ is $2n+1$.

Radial basis function (RBF) SVMs (Burges, 1998) were trained on TIMIT to detect acoustic landmarks. The TIMIT landmark detectors were tested on TIMIT and NTIMIT. No adaption algorithms were implemented. Results are shown in Table 1. SVM training and testing was performed using LibSVM (Chang and Lin, 2001).

5 Results

Landmark detection results are shown in Table 1. When training and testing conditions are matched, NRL and MFCCNRL-based detectors outperform the MFCC baseline. Detectors trained on MFCCNRLs are not as accurate as those trained on the NRLs alone. When training and testing conditions are mismatched, (i.e., added noise and downsampling), the overall landmark detection accuracy degrades. In mismatched conditions, NRL and MFCCNRL-based landmark detectors generally either outperform the MFCC baseline or do not produce results significantly different from the baseline. SVM landmark detectors trained on the MFCCNRL do not perform as well as the SVMs trained on NRL. Significance is calculated using the binomial

	TIMIT/TIMIT			TIMIT/NTIMIT		
	MFCC	NRL	MFCCNRL	MFCC	NRL	MFCCNRL
−+silence	92.6	**94.7**	93.4	84.3	**86.1**	85.8
+−silence	87.2	**94.9**	91.6	82.3	**87.6**	86.8
−+continuant	81.2	**89.1**	86.4	70.8	**79.4**	**79.4**
+−continuant	92.3	**92.7**	91.5	85.5	**86.5**	85.3
−+sonorant	81.9	**88.8**	86.5	74.0	73.5	**77.9**
+−sonorant	**93.9**	92.6	93.1	86.2	79.1	**88.0**
−+syllabic	85.6	**89.5**	85.5	77.0	**87.1**	80.5
+−syllabic	85.8	**88.1**	84.5	83.0	**86.5**	79.5
−+consonantal	90.8	**92.0**	90.0	86.5	83.2	**87.3**
+−consonantal	80.4	**85.4**	82.2	71.2	71.6	**74.9**

Table 1: Support vector machine (SVM) landmark detection results for SVMs trained and tested on TIMIT (TIMIT/TIMIT), and for SVMs trained on TIMIT and tested on NTIMIT (TIMIT/NTIMIT). SVMs are trained using a Mel-frequency cepstral coefficients (MFCCs), auditory neural rate and level (NRL), and a combination of MFCCs and NRLs (MFCCNRL). Chance is 50%. Underlined values show a significant difference in accuracy from the MFCC baseline for $p = 0.05/20 = 0.0025$. The factor of 20 is necessary because for any given test set, the table above shows the results of 20 simultaneous significance tests.

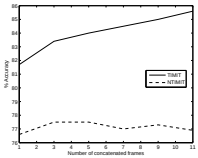

Figure 5: Detection accuracy as a function of the number of concatenated acoustic feature frames in $\vec{x}_t$ for the [−+syllabic] MFCC-based landmark detection SVM. The [−+syllabic] landmark detector was tested on TIMIT (solid line) and NTIMIT (dashed line).

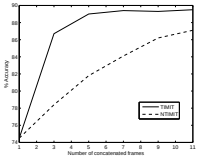

Figure 6: Detection accuracy as a function the number of concatenated acoustic feature frames in $\vec{x}_t$ for the [−+syllabic] rate and level-based landmark detection SVM. The [−+syllabic] landmark detector was tested on TIMIT (solid line) and NTIMIT (dashed line).

test described in (Gillick and Cox, 1989).

Figure 5 shows a plot of detection accuracy vs number of concatenated frames in $\vec{x}_t$ for the MFCC-based [−+syllabic] landmark detector for both test corpora. For MFCC-based SVMs, no significant increase in detection accuracy is observed as a function of the number of concate-

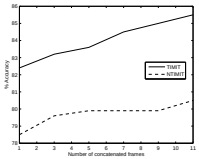

Figure 7: Detection accuracy as a function of the number of concatenated acoustic feature frames in $\vec{x}_t$ for the [−+syllabic] MFCC/rate and level-based landmark detection SVM. The [−+syllabic] landmark detector was tested on TIMIT (solid line) and NTIMIT (dashed line).

nated frames when the training and testing conditions are mismatched. Detection accuracy increases as a function of the number of concatenated frames when the training and testing conditions are matched.

Figure 6 shows a plot of detection accuracy vs number of concatenated frames in $\vec{x}_t$ for the NRL-based [−+syllabic] landmark detector for both test corpora. The detection accuracy of the NRL-based SVMs increases as a function of the number of concatenated frames regardless of whether the training and testing conditions are matched or mismatched.

Figure 7 shows a plot of detection accuracy vs the number of concatenated frames in $\vec{x}_t$ for the MFCCNRL-based [−+syllabic] landmark detector. There is a slight increase in landmark detection accuracy as a function of the number of concatenated frames for both matched and mismatched conditions.

Conclusion

This paper explores the use of octopus cell neural firing rate and average spectral level as acoustic features, and presents an auditory model that can be used to create these features. Neural firing rate and average spectral level accurately represent acoustic-phonetic landmarks in both matched and mismatched conditions.

The current system exploits only the left channel of the model. In the brainstem, input to both ears is essential for signal denoising. Future work will explore methods to combine both channels to increase landmark detection accuracy in mismatched conditions.

Accurate landmark detection may be essential for accurate phonetic segmentation — a process that is essential for speech recognition. The current system provides a building block for an automatic speech segmentation system designed to be integrated with a speech recognizer. Implementation of these systems is the focus of future research.

Acknowledgments

This work was supported in part by a grant from the National Science Foundation (CCF 0807329) and in part by a grant from the Qatar National Research Foundation (NPRP 09-410-1-069). The findings and opinions expressed in this article are those of the authors, and are not endorsed by QNRF or the NSF.

References

Bekesy, G. v. (1960). *Experiments in Hearing*. McGraw-Hill, New York, NY.

Borys, S. (2008). An SVM front end landmark speech recognition system. Master's thesis, University of Illinois, Urbana-Champaign.

Borys, S. and Hasegawa-Johnson, M. (2005). Distinctive feature based discriminant features for improvements to phone recognition on telephone band speech. In *Eurospeech*, pages 679–700.

Bourlard, H. (1996). A new ASR approach based on independent processing and recombination of partial frequency bands. *ICSLP*, 1.

Burges, C. (1998). A tutorial on support vector machines. *Data Mining and Knowledge Recovery*, 2(2).

Chang, C.-C. and Lin, C.-J. (2001). *LIBSVM: a library for support vector machines*. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

Chigier, B. (1991). Phonetic classification on wide-band and telephone quality speech. In *International Conference on Acoustics, Speech, and Signal Processing*.

Cohen, J. (1989). Application of an auditory model to speech recognition. *JASA*, 85(6).

Ehret, G. and Romand, R., editors (1997). *The Central Auditory System*. Oxford University Press, New York, NY.

Garofolo, J., Lamel, L., Fisher, W., Fiscus, J., Pallett, D., and Dahlgren, N. (1993). The DARPA TIMIT acoustic phonetic speech corpus. Technical report, National Institute of Standards and Technology, Gaithersburg, MD.

Geisler, C. D. (1998). *From Sound To Synapse: Physiology Of The Mammilian Ear*. Oxford University Press, Oxford, NY.

Ghitza, O. (1994). Auditory models and human performance to tasks related to speech coding and speech recognition. *IEEE Transactions on Speech and Audio Processing*, 2(1).

Gillick, L. and Cox, S. (1989). Some statistical issues in the comparison of speech recognition algorithms. In *ICASSP*.

Gong, Y. (1995). Speech recognition in noisy environments: A survey. *Speech Communication*, 16.

Greenberg, S. (1988). The ear as a speech analyzer. *Journal of Phonetics*.

Halberstadt, A. K. and Glass, J. R. (1998). Heterogeneous measurements and multiple classifiers for speech recognition. In *International Conference on Speech and Language Processing*, Sydney, Australia.

Hasegawa-Johnson, M., Baker, J., Greenberg, S., Kirchhoff, K., Muller, J., Sommez, K., Borys, S., Chen, K., Juneja, A., Livescu, K., Mohan, S., Coogan, E., and Wong, T. (2004). Landmark-Based speech recognition: Report of the 2004 Johns Hopkins summer workshop. Technical report, Johns Hopkins University, Center for Speech and Language Processing, Baltimore, MD.

Hunt, M. and Lefebvre, C. (1989). A comparison of several acoustic representations for speech recognition for degraded and undegraded speech. In *ICASSP*, volume 1.

Jankowski, C., Kalyanswamy, J., Basson, S., and Spritz, J. (1990). NTIMIT: A phonetically blaanced, continuous speech, telephone bandwidth speech database. In *International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 109–112.

Juneja, A. and Espy-Wilson, C. (2004). Significance of invariant acoustic cues in a probabilistic framework for landmark-based speech recognition. In *From Sound to Sense: 50+ Years of Discoveries in Speech Communication*, pages C–151–C–156. MIT, Cambridge, MA.

Karray, L. and Martin, A. (2003). Towards improving speech detection robustness for speech recognition in adverse conditions. *Speech Communication*, 40.

Kim, D., Lee, S., and Kil, R. (1999). Auditory processing of speech signals for robust speech recognition in real-world noisy environments. *IEEE Transactions on Speech and Audio Processing*, 7(1).

Kirchhoff, K. (1999). *Robust speech recognition using articulatory information*. PhD thesis, University of Bielefield, Germany.

Moore, B. and Glasberg, B. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *JASA*, 74(3).

Moreno, P. and Stern, R. (1994). Sources of degradation of speech recognition in the telephone network. In *IEEE Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 109–112.

Patterson, R. and Holdsworth, J. (1996). A functional model of neural activity patterns and auditory images. *Advances in Speech, Hearing, and Language Processing*, 3.

Schnupp, J., Nelken, I., and King, A. (2011). *Auditory Neuroscience: Making Sense of Sound*. MIT Press, Cambridge, MA.

Stevens, K. (1999). *Acoustic Phonetics*. MIT Press, Cambridge, MA.

Tidemann, J. (2011). Characterization of the head-related transfer function using chirp and maximum length excitation signals. Master's thesis, University of Illinois.

Virag, N. (1999). Single channel speech enhancement based on masking properties of the human auditory system. *IEEE Transactions on Speech and Audio Processing*, 7(2).

Winer, J. A. and Schreiner, C. E. (2005). *The central auditory system: A functional analysis*, chapter 1. Springer Science+Business Media, Inc., New York, NY.

Ying, S., Werner, V., and Xue-ying, Z. (2012). A robust feature approach based on an auditory model for classification of speech and expressiveness. *J. Cent. South Univ.*, 19.