5 Discussion and Future Work

5.1 Problem statement

Younger adults with normal hearing (YNH) have a remarkable ability to understand a single talker in the presence of competing talkers in the real world listening environments. This ability is greatly reduced in listeners with the acoustic variations (stimulus sound level or duration) as well as with the auditory variations due to increased age and hearing loss. People who wear hearing aids often complain that they can hear sound but have major difficulty in speech recognition. One of the fundamental limitations to the success of auditory prostheses is the current lack of understanding of the neural mechanism underlying the speech recognition scores. To address this issue, a physiologically inspired signal processing model was employed in this dissertation to understand the effects of acoustic and auditory changes on neural mechanisms that contribute to recognition of concurrent vowel scores. The long-term goal of this multi-disciplinary modeling approach is to understand the neural mechanisms underlying any speech signals and then transfer this knowledge to improve the signal processing algorithms in hearing-aids.

5.2 Significance of the dissertation work

Concurrent vowel identification is an experimental task performed for understanding how differences in F0 can be used for segregating multiple talkers prior to identifying the target speech. The behavioral studies have investigated the effect of F0 differences on the percent-correct identification of concurrent vowels in YNH subjects (varying sound levels or durations), normal-aging subjects and hearing-impaired subjects with aging. The level-dependent studies have shown that concurrent vowel scores were reduced at low and high-levels across different F0 conditions (Chintanpalli et al., 2014). In concurrent-vowel data, the F0 benefit increased with increasing vowel level (25-50 dB SPL) and then remained fairly constant from 50 dB SPL onwards. For shorter vowel duration, the concurrent vowel scores were reduced and had across subject-variability in the F0-benefit (Assmann and Summerfield, 1990, 1994; Culling and Darwin, 1993; McKeown and Patterson, 1995). With increase in age (Arehart et al., 2011; Snyder and Alain, 2005; Vongpaisal and Pichora-Fuller, 2007) and hearing loss (Arehart et al., 1997, 2005; Summers and Leek, 1998), the concurrent vowel scores were reduced across F0 differences. Chintanpalli et al. (2016) had further shown that the scores were lowest for older adults with hearing loss, when compared with the aged (normal-hearing) listeners. The investigations of the neural representations due to the variations of these identification scores are generally studied using a physiologically based computational modeling. More specifically, there are modeling studies on concurrent vowels that had qualitatively predicted the scores only for YNH listeners either at a fixed duration or sound level (Chintanpalli and Heinz, 2013; Meddis and Hewitt, 1992). To address the above-mentioned problems related to acoustic and auditory changes, the current dissertation had employed a similar modeling framework to predict (1) level and duration effects on concurrent vowel scores across YNH subjects (through objectives I and II) and (2) decline in concurrent vowel scores due to increased age and hearing loss (objectives III).

Chapter 2 developed a computational model that had predicted the level-dependent changes in concurrent vowel identification for both the 0 and 26 Hz F0 difference between the vowels for YNH listeners. This model was developed by cascading a physiologically realistic AN model of Zilany et al. (2014) and a modified version of the Meddis and Hewitt (1992) segregation algorithm to predict the concurrent vowel scores. The model was qualitatively successful in capturing the level-dependent changes of concurrent vowel scores, as observed in the Chintanpalli et al. (2014) behavioral data. The model predictions suggest that changes in phase-locking of AN fibers to F0s and formants of concurrent vowels contribute to changes in vowel segregation and thereby resulting in level-dependent changes in the scores. Chapter 3 developed a computational model to understand the possible underlying cause for reduced concurrent vowel scores for shorter duration (e.g., 50 ms), relative to the scores obtained for longer duration (e.g., 200 ms) in YNH subjects. It was hypothesized that a limited use of F0-guided segregation cue might contribute to the reduced vowel scores for shorter (50 ms) duration. A similar computational model that was utilized in Chapter 2 was used to study the duration effects. The limited ability to use the F0-guided segregation cue in the model resulted in a lowered concurrent vowel scores at 50 ms duration, and this provides an explanation to the Assmann and Summerfield (1994) behavioral data. Additionally, the differential reduction in the ability to avail this cue had resulted in a variable F0 benefit that were observed across various behavioral studies (Assmann and Summerfield, 1990, 1994; Culling and Darwin, 1993; McKeown and Patterson, 1995).

To assess the age and hearing loss effects chapter 4 developed a computational model to predict the concurrent vowel scores across F0 differences for three different listening groups. The YNH model was developed by cascading a recent version of a physiologically realistic AN model (Bruce et al., 2018) with a modified version of Meddis and Hewitt (1992) F0-guided segregation algorithm. For the ONH computational model, EP reduction and CS were incorporated. Whereas, for the OHI model, hair-cell damage and CS (with a larger effect than ONH) were incorporated. The concurrent vowel scores for the three computational models were validated using the Chintanpalli et al. (2016) behavioral data. The reduced vowel segregation in both the ONH and OHI models suggest a limited use of the F0-difference segregation cue with increased age and hearing loss. Additionally, all three models successfully captured the pattern of concurrent vowel scores with increasing F0 difference and F0 benefit, as observed in the concurrent vowel data. These model predictions confirm our hypotheses and suggest that peripheral changes (largely by CS) due to increased age and hearing loss could contribute to reduced concurrent vowel scores for ONH and OHI subjects.

5.3 Limitation of the model

In the current dissertation, the time-varying DRs of the AN model were passed through the F0-guided segregation algorithm for the prediction of concurrent vowel scores. These DRs are the temporal cues of AN fibers and are robust across sound levels. However, the model scores may be able to improve at low-and-mid levels (section 2.4) by incorporating the rate place cues of AN fibers (e.g., Young and Sachs, 1979), along with the existing temporal cues. One way to include both the rate-place and temporal cues is to use a neural network framework for vowel identification. There has been limited work to predict the concurrent vowel scores using deep neural networks. A single-layer perceptron was used to predict the concurrent vowel scores from the temporal responses of a simple AN model (Culling and Darwin, 1993). Instead of directly computing the percent correct score, the response probability associated with each concurrent vowel pair was computed. Recently, there have been attempts to develop a deep neural networks based AN models using a time-delay neural network (Nagathil et al., 2021). Instead of the Meddis and Hewitt (1992) F0-segregation based algorithm, a time-delay neural network along with multi-task learning can be developed to predict the concurrent vowel scores across F0 differences.

There is a variability in listeners' responses in any of the behavioral study. However, the model scores are deterministic for a fixed parameter set. To include the response variability in the model, a probabilistic approach needs to be incorporated while computing the model scores.

5.4 Future Work

5.4.1 Motivated Behavioral studies

The current modeling study can be extended to predict the effects of age and hearing loss on the concurrent vowel scores with the acoustic changes (sound level and duration variations). The ONH and OHI models (developed in Chapter 4) can be used to understand the effects of age and hearing loss on concurrent vowel scores due to changes in acoustics (across levels and durations), using the same parameters of the F0-guided segregation algorithm that were used in the Chapter 2 and 3, respectively. These predictions may help us design an experimental paradigm (e.g., a suitable duration and sound level) for aged or hearing-impaired listeners, for which behavioral data is currently unavailable in the literature.

5.4.2 Auditory Nerve Models

The Zilany et al. (2014) and Bruce et al. (2018) are phenomenological models which were tested extensively against neurophysiological responses from cats for both normal and impaired ears for a wide range of input stimuli. However, it is a computationally intensive model and it takes time to run simulations for the larger number of CFs and stimulus set to draw the conclusions regarding the concurrent vowel scores. An investigation of proposing a less computationally intensive model, may widen the potential uses of model predictions in real time measurements. As an initial step, the auditory filters (i.e., a filter-bank with varying Q_{10} values) can be implemented on the digital signal processing (DSP) kit with an aim of reducing the processing speed and memory.

5.4.3 Clinical Applications

In the hearing clinics, currently, it is impossible to determine the extent of configuration of hair-cell damage (OHC and IHC loss) and CS in hearing impaired listeners. The current modeling approach of Chapter 4 can be used as a diagnostic tool for identifying the extent of hair-cell damage and CS. In the computational model, we can alter the extent of hair-cell damage and CS such that the model scores across F0 difference were qualitatively similar to scores of the specific hearing impaired patient. Hence, the extent of hearing loss and CS can be speculated for an individual patient.