

TABLE OF CONTENTS

Table of Contents.....	Pg. 1
Day 1	Pg. 2
Day 2	Pg. 7
Day 3	Pg. 11

DAY 1

Eric Young

The Neural Representation of Successive Segments in Speech in Normal and Acoustically-Traumatized Ears

Studies of neural encoding of speech have usually focused on representation of formants. However, for many consonants the formant transitions are only a part of the signal and other sounds, associated with aspiration or frication, are also important. Also, real speech consists of a rapid sequence of sounds with different spectro-temporal characteristics. Here, responses of auditory nerve fibers, in both normal and acoustically-traumatized ears, were studied using the stimulus "Five women played basketball", a natural utterance. For analysis, the stimulus was segmented into sections corresponding to phonemes or pairs of phonemes, with approximately constant spectral shape (50-70 ms in duration). The population response (rate versus BF) was computed for each segment, in BF bins representing equal steps along the basilar membrane. Similar phonemes cluster together in this analysis. Confusion matrices show accurate identification of the 19 segments in the sequence in normal ears. In impaired ears, the second formant representation was missing, resulting in moving together of the clusters and degraded identification performance. This way of quantifying performance degradation provides a direct way of evaluating hearing-aid algorithms. Supported by NIH grant DC00109

Roy Patterson

The Temporal Form of Size Information in Speech, Music and Animal Sounds

There is size information in animal communication sounds; if two animals differ only in size, the resonances of the larger animal will be lower in frequency and expanded in time. For animals, size information is important for assessing the sex of an individual and their ability to defend a territory. In human speech, vocal tract length is a major component of variability and the correct parsing of the variability is essential in speech recognition both by humans and machines. Recently, there have been important developments with respect to size in communication sounds. Fitch and colleagues have shown that there is a strong correlation between vocal tract length and body size in most mammalian species (e.g. Fitch and Giedd, 1999), and that mammals have mechanisms for exaggerating their size. Cohen (1993) has described an affine transformation that illustrates the form of size information in sounds, and Irino and Patterson (2002) have described a time-domain model of size processing in the human auditory system. They argue that it provides the basis for vowel normalization in speech communication. There will also be a special session at the ASA meeting in Vancouver on size information in speech and animal calls. This paper will describe a time-domain model of size processing in the auditory pathway, and review recent developments.

Terry Picton

Timing and the Auditory Evoked Potentials

Many different human auditory evoked potentials can be elicited by temporal changes in sounds – alterations in their relative timing, variations in the correlation between them, or temporal modulations in their amplitude or frequency. Three studies will be briefly reviewed: 1) Late transient responses can be evoked by shifts in the relative timing of a signal presented to the two ears. Earlier transient responses are not clearly recognized, suggesting a continuous comparison between the ears that is only intermittently checked for change. 2) A stimulus that regularly changes its correlation between the ears evokes large steady-state responses at low frequencies and smaller responses at more rapid frequencies. These data can be fit with a model that postulates a rapidly reactive brainstem process and a much more sluggishly reactive cortical response. 3) By sweeping the modulation frequency of a noise or tone, we can physiologically track the temporal modulation transfer function. Physiological recordings provide measurements of amplitude and latency that are not available to psychophysical evaluation. These measurements again indicate that the brainstem can follow modulations up to several hundred Hz whereas the cortex is mainly limited to frequencies of less than 70 Hz.

Willy Wong

Discussion of Physiological Models

The engineering of auditory interfaces must be predicated on a proper understanding of the temporal and spatial characteristics of audition. In this talk, we revisit a classic problem of psychoacoustics related to the time-dependent characteristics of auditory thresholds. We offer new insight to this area and suggest how the concept of entropy can help elucidate the process underlying auditory thresholds. We also compare similar results obtained in vision, results which have played a vital role in the design of visual information displays.

John Grose

Detecting and Discriminating Gaps: Stimulus Factors and Listener Factors

Gap detection gauges a listener's sensitivity to the occurrence of an interruption in an otherwise continuous sound. Gap duration discrimination, on the other hand, measures the just-noticeable-difference in the length of a perceptible silent interval. In both of these paradigms, a listener's performance can be affected by the physical characteristics of the sounds that bound the gap, such as their level or bandwidth. However, the acoustic properties of the gap markers can have less straight-forward effects. For example, the processing of silent intervals is generally more difficult if the gap markers have non-overlapping or asymmetric spectral contents. This highlights the importance of perceptual discontinuities, in addition to physical discontinuities, in the processing of gaps. Finally, characteristics of the listeners themselves may affect performance, such as their age or hearing loss status. The purpose of this presentation is to review some of the stimulus and listener factors that affect the temporal processing of gaps.

Kathy Pichora-Fuller

Age-Related Differences in Temporal Coding of Gaps in Speech and Non-Speech Signals

Auditory temporal processing likely contributes to the difficulty of older adults in understanding speech in noise. In our research we have focused on two aspects of auditory temporal processing: gap detection and synchrony coding and gap detection. The presentation will describe a series of experiments concerning gap detection for various non-speech and speech markers. Ability to detect gaps in speech and non-speech stimuli was measured in young and old adults with good audiograms and also in children, young adults and older adults. Young adults were also tested in conditions of simulated auditory aging. Auditory aging was simulated using temporal jittering to disrupt the periodicity of the signal. The markers surrounding the gap varied in duration (40 vs 250 msec) and in spectral symmetry. In spectrally symmetrical conditions, the leading and lagging markers were the same: the vowel [u] in speech conditions and a 500-Hz tone in non-speech conditions. In asymmetrical speech conditions, the lagging marker was the same as in the symmetrical conditions, but the leading marker was the consonant [s] in the speech conditions and a broadband noise (1 to 6 kHz) in the non-speech conditions. For the intact stimuli, the gap detection thresholds for all age groups in spectrally symmetrical markers were far smaller than in spectrally asymmetrical markers. In all conditions, gap thresholds were significantly smaller in young adults than in children or older adults. For all age groups, gaps between spectrally asymmetrical speech markers were detected better than gaps between analogous non-speech stimuli. It is argued that phonological knowledge compensates for auditory processing difficulties in both adult groups. For the young adults tested under conditions of simulated auditory aging, gap detection thresholds in the symmetrical conditions were significantly larger than for either age group when intact stimuli were used; however, performance in the asymmetrical conditions was not worse. Both markers used in the symmetrical conditions were periodic (tones or vowels) whereas the leading markers in the asymmetrical conditions were aperiodic (consonants or noise bands); therefore, it is not surprising that the simulation of auditory aging has a pronounced effect on gap detection in the former but not in the latter conditions. The findings are discussed in terms of the different aspects of temporal processing involved in the detection of gaps in different types of stimuli and the age-related changes in different aspects of temporal processing as they may relate to speech perception. Research funded by the International Dyslexia Association, the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research.

Liang Li

Auditory Memory of Fine Details in Humans

To recognize when a delayed noise is a copy of the leading noise, the auditory system needs to store detailed information of the leading noise over a period of time. Thus, the memory of fine details of sound waveforms is important for perceptually grouping correlated sounds and segregating uncorrelated sounds in reverberant environments, where older listeners experience certain difficulties that young listeners do not experience. To determine the temporal extent of the auditory memory and whether it is affected by aging, this study investigated the detection of a break in correlation (BIC) between two interaurally correlated broadband noises in young and older normal-hearing listeners. The results show that young listeners could detect a 100-ms BIC up to interaural delays ranging from 6.3 to 23.0 ms, while older listeners could detect the BIC only up to interaural delays ranging from 6.7 to 9.7 ms. Moreover, with the rise of inter-sound delay from 2 to 10 ms, the shortest BIC duration necessary for listeners to correctly detect the occurrence of the BIC increased rapidly, and the listener's reaction time in response to the BIC, but not to a comparable silent gap, elongated quickly. We propose that higher-order central mechanisms beyond the brainstem delay lines are likely to be involved in maintaining a memory trace of the fine details of the acoustic waveform, and the rapidly fading auditory memory is important for perceptually grouping correlated sounds and segregating uncorrelated sounds in noisy, reverberant environments.

Bob Shannon

Temporal Information in Speech: The Role of Envelope vs Fine Structure

Speech recognition is highly robust with regard to distortions in temporal information. Temporal information has been categorized into envelope (below 50 Hz), periodicity (50-500 Hz), and fine (>500 Hz). Although implant listeners and normal hearing listeners can perceive temporal pitch and modulation up to 300-500 Hz, results from experiments show that implant speech recognition in quiet is unaffected by reductions in temporal information above 20 Hz. Synchronization of timing across frequency regions is also not critical for speech, as recognition is unaffected by cross-frequency disruptions in timing of more than 200 ms. As listening conditions worsen by the addition of noise and/or distortion more temporal information is necessary to maintain speech recognition. This pattern suggests that slowly changing speech spectral patterns and temporal fine structure contribute independently to intelligibility.

Erin Hannon

Perceptual Learning of Musical Temporal Structure During Infancy and Adulthood

Culture-specific knowledge powerfully constrains how adults perceive and produce temporal structures in a musical context. North American adults have difficulty reproducing and remembering rhythmic patterns that fail to conform to Western conventions of temporal isochrony and simple duration ratios. For example, North American Western adults readily detect disruptions to an isochronous (Western) rhythm, but they fail to detect comparable disruptions of a non-isochronous (Balkan) rhythm. We recently demonstrated that such asymmetries were absent in adults from Bulgaria and in 6-month-old infants, indicating culture-general perception in infancy but culture-specific perception in adulthood. In this talk I will present evidence that 12-month-old infants continue to differentiate variations in a familiar, Western context, but fail in a foreign, Balkan context, suggesting that culture-specific biases in rhythm perception emerge by one year of age. By providing at-home exposure to Balkan music in the weeks prior to testing, we reversed this decline in infants but not in adults. Our findings suggest that perception of musical temporal structure undergoes a process of experience-dependent tuning during the first year of life, paralleling similar developmental changes in speech and face perception. Age-related changes in perceptual learning may constrain when and how individuals learn about temporal structures in music.

Ewen MacDonald

Simulation of Temporal Aspects of Auditory Aging

A jittering technique to disrupt the periodicity of the signal was used to simulate the effect of the loss of temporal synchrony coding believed to characterize auditory aging. In one experiment jittering was used to distort the frequency components below 1.2 kHz and in a second experiment the components above 1.2 kHz were distorted. To control for spectral distortion introduced by jittering, comparison conditions were created using a smearing technique (Baer & Moore, 1993). In both experiments, 16 normal hearing young adult subjects were presented with SPIN sentences in three conditions (intact, jittered, and smeared) at 0 and 8 dB SNR. When the low frequencies were distorted, speech intelligibility in the jittered conditions was significantly worse than in the intact and smeared conditions, but the smeared and intact conditions were equivalent. When the high frequencies were distorted, speech intelligibility was reduced similarly by jittering and smearing. On low-context jittered sentences, results for young adults mimicked results found previously for older listeners with good audiograms (Pichora-Fuller et al., 1995). It is argued that the jittering technique could be used to simulate the loss of neural synchrony associated with age-related changes in temporal auditory processing.

DAY 2

Pierre Divenyi

The Magic Duration of 100 ms in Auditory Segregation

The time constant for temporal integration in the auditory cortex has been estimated to be about 100 ms but not longer (Schreiner & Urbas, 1988), in the cat. MEG studies in humans (Yabe *et al.*, 1998) have estimated it to be in the neighborhood of more than 160 ms. One concept reconciling the two estimates derives from Hirsh's (1974) division of temporal processing into three regions: the short-range (up to 25 ms) of simultaneity, the middle range (25-100 ms) of Gestalt formation, and the long-range (100 ms and up) of separate auditory events. Interestingly, the long range is implicated in results of our experiments on auditory segregation of simultaneous speech-analog streams. These streams are simultaneously presented harmonic sinusoidal complexes with different f_0 's in the 100 to 200-Hz range and a single formant generated by a band-pass filter with dynamically changing center frequency; the trajectory of the formants start at around 1.5 kHz and have an upward (or downward) transition to a certain maximum frequency deflection followed by a downward (or upward) transition back to the starting frequency. Segregation of the two streams is easiest when the duration of a single transition is exactly 100 ms, i.e., at the boundary of the region at which the formant sweep can be perceived as a separate event and is likely to rely on processes taking place in the auditory cortex.

Christian Giguère

Working in Noise

In many occupational settings, speech communication, signal detection and sound localization tasks are performed in noise. These environments can sometimes be very challenging, particularly for individuals with hearing loss. Individuals who must perform these tasks and whose functional hearing abilities are impaired may constitute a risk to themselves and others. Diagnostic measures of hearing, such as the audiogram, are not adequate to make accurate predictions of speech intelligibility in real-world noise environments. Instead, a direct functional measure of hearing, the Hearing-In-Noise Test (HINT), has been identified and validated for use in predicting speech intelligibility in a wide range of speech communication situations in real-world noise environments. The prediction approach takes into account the voice level of the talker in noise due to the Lombard effect, the communication distance between the talker and the listener, a statistical model of speech perception in specific occupational noises, and the functional hearing abilities of the listener. The latter is taken as the elevation of the individual's speech reception threshold in noise above the normative value for the HINT test. This test and normative values are available in several languages, so that language-specific needs can be addressed. The detailed approach will be presented with an emphasis placed on application examples based on recent work to screen individuals working in hearing-critical tasks in the Department of Fisheries and Oceans Canada.

Claude Alain

Discussion of Real-World Listening: Neuroelectric Correlates of Age-Related Decline in Temporal Acuity

Age-related declines in the ability to code the temporal properties of the speech envelope are thought to contribute to the speech perception problems often experienced by older adults. Because the speech envelope is defined by energy fluctuations in the signal in which low-energy periods (gaps) are interspersed with high-energy periods of varying length, the two psychophysical measures most relevant to the processing of the speech envelope are the ability to detect a gap in a continuous sound and the ability to discriminate between two sounds on the basis of their duration. In this presentation, I will focus on the neural correlates of temporal acuity and present evidence that normal aging impairs listeners' ability to automatically process sound duration and detect rapid fluctuation in acoustic energy (i.e., gap). The implications of these findings for current models of temporal acuity and aging will be discussed.

Art Wingfield

Understanding Speeded, Noisy, and Complex Speech at the Word, Sentence and Discourse Levels: Implications for Adult Aging and Hearing Loss

Comprehension of rapid speech in complex environments is constrained by a number of factors. A common real-world source of this complexity is the task of following a single speaker being heard in a background of other voices, where both sensory and cognitive factors come into play. On the sensory side the listener must deal with rapid, often poorly articulated speech, a challenge that is exacerbated in older adults with high frequency hearing loss and reduced efficiency in temporal processing and frequency discrimination. These "bottom up" declines can be ameliorated by "top down" use of linguistic context for recognition of words as the speech unfolds in time, and also for retrospective recognition of an indistinct word based on the context that follows it. A second major factor in speech comprehension is the use of prosody, to include pitch contour, stress, and temporal patterning, such as the lengthening of clause-final words to signal that a clause boundary has been reached. Speech prosody can be a valuable aid to syntactic resolution as a step toward semantic comprehension. In all adults, and especially older adults, these operations are constrained by limitations in attentional or processing resources, a factor that shows bi-directional interaction with sensory challenge. This presentation will report research from our laboratory on speech comprehension and memory investigating each of these factors in young and older adults with good hearing and with mild-to-moderate hearing loss.

Antje Heinrich

The Effects of Temporal Distortion on Speech Perception and Memory Performance in Young and Old Listeners

Using peripheral auditory measures such as gap detection and duration discrimination, a number of studies have shown that the accuracy of temporal coding declines with age. Moreover, it is also well known that memory performance on episodic memory tasks declines with age. This study employed an auditory paired-associate memory paradigm and investigated the influence of compromised temporal accuracy on speech perception and memory performance. For that reason, word stimuli were temporally distorted (jittered) and speech perception accuracy as well as memory performance for these jittered words was determined for both age groups. In addition, memory performance in young and old adults for jittered words was compared to memory recall of masked words. The results show that the effect of temporal distortion on memory was quite different from the effect of masking. Moreover, the effect of temporal distortion was similar in pattern in young and old listeners but exacerbated in old adults presumably due to more advanced levels of auditory sensory decline.

Astrid vanWieringen

Temporal Processing in Cochlear Implants

In order to perceive speech sounds, the auditory system must be able to discriminate frequency, changes in amplitude and duration, and gaps between sounds. This presentation will highlight three different studies related to auditory temporal processing in cochlear implants. First, gap detection experiments for different complex patterns of electrical stimulation will be described, as these allowed us to examine the nature of within- and across channel auditory processes in detail. The study showed that stimulus complexity does not need to affect gap detectability and that cochlear impairment does not have to affect the temporal resolution of the auditory system. Second, several psychophysical experiments were carried out with normal-hearing and implanted subjects under conditions where place-of-excitation was held constant, and where pitch was therefore derived from “purely temporal” cues. Different theories state that pitch is determined from the intervals between each pulse and every other pulse (“autocorrelation”), from only those intervals between each pulse and the next (“1st-order intervals”), or simply from the total number of pulses per unit time (“mean rate”). First experiments showed that the pitch change should be determined not just by the physical timing of the pulses, but by the auditory nerve responses to them. Subsequently, it was examined how normal-hearing listeners and cochlear implantees derive temporal pitch information from amplitude modulated pulse trains. The acoustical and electrical pulse trains consisted of pulses whose amplitudes alternated between a high and a low value, and whose inter-pulse intervals alternated between 4 and 6 ms. The attenuated pulses occurred after the 4-ms intervals in condition A, and after the 6-ms intervals in condition B. For both normal-hearing subjects and cochlear implantees, the period of an isochronous pulse train equal in pitch to this ‘4-6’ stimulus increased from near 6 ms at the smallest modulation depth to nearly 10 ms at the largest depth. Additionally, the modulated pulse trains in condition A were perceived as being lower in pitch than those in condition B. Data are interpreted in terms of

Auditory Temporal Processing Symposium Abstracts

increased refractoriness in condition A, where the larger pulses are more closely followed by the smaller ones than in condition B. Third, the presentation will touch on a user-friendly research platform consisting of highly animated psychophysical procedures. These have been modified to examine possible temporal resolution deficits in young, normal-hearing pre-school children who have a risk of being reading-impaired.

Jeff Bondy

Cochlear Nonlinearities and Temporal Neural Clues

Sensorineural hearing impairment is devastating to a listener's ability to hear in temporally modulated noise and competing speech. Many different psychophysical deficits are seen with sensorineural impairment, but very few delineate the handicap as well as the difficulty of hearing in competing speech. A great step in understanding speech perception would be determining how a normal ear can parse a competing-speech acoustic landscape and how that is different for the hearing impaired. We looked at the optimal neural coding across amplitude, frequency and time of audio signals, and how the loss of adaptive nonlinearities in the cochlea such as compression, suppression and adaptation would effect the representation. In particular, adaptation is shown to be functionally important for coding the onset cue. This onset cue is used to improve robustness and reduce noise by normal hearing people, but is heavily distorted in the hearing impaired. Changes to fast adaptation brought on by hair cell damage are equivalent to misaligning these onset cues, that are important to parse the time-frequency plane.

Ian Bruce

How Lab Researchers can Interact with Industrial Partners

Basic scientific research in the academic environment often seems worlds apart from the development of commercial products. However, our research results in auditory neuroscience should surely make some impact in the design of assistive devices for the hearing impaired such as hearing aids and cochlear implants, in treatment strategies for tinnitus, and in training regimes for those with central auditory processing deficits. In this talk I will discuss what motivations we might have for interacting with an industrial partner, what manner of interactions are possible, and what benefits can be gained from different approaches.

DAY 3

Roy Patterson

What Limits the Notes that Composers Can Use to Make Tonal Melodies?

This is a talk for a general scientific audience with lots of interesting new sounds and a little hearing theory. The argument is that it is peripheral processing in the cochlea and brainstem that determines which notes can be used to make melodies. Then I show how we can make creative use of the rules once we know them to make new classes of notes with specific perceptual effects. In the course of the lecture, the examples will provide a short course on modeling temporal processing in the auditory system with examples of how brain imaging can be used to determine where the different stages of temporal processing are performed in the brain.

Bruce Schneider

Listening in Everyday Situations: From Hearing to Comprehension

Consider a real life situation in which a listener is trying to follow a conversation between two talkers in a complex auditory scene such as a restaurant. To comprehend what is being said by these two individuals, the listener first has to segregate two sound streams from one another and from the auditory background. Then the listener has to process the individual words and phrases spoken by each person in order to comprehend each person's message. Finally, the listener must integrate the two messages with past input and world knowledge in order to fully comprehend the exchange. When the acoustics are poor or the background noise is considerable, a listener's ability to do this will be adversely affected. Moreover, when poor acoustics are combined with virtually any kind of auditory problem (even those which would not normally merit clinical attention), all of these listening difficulties will be considerably exacerbated. For example, a number of studies have demonstrated that older adults with clinically normal hearing are considerably more disadvantaged than normal-hearing younger adults in adverse listening conditions. Indeed, hearing status in older adults is, arguably, the best predictor of their performance on a number of different cognitive tasks. In this presentation we will investigate the reasons for the tight linkage between hearing status, listening conditions, and cognitive performance. In doing so we will show that features of the acoustical environment can have a surprisingly large effect on the ability to segregate sound sources in an auditory scene, comprehend spoken language, and recall information from monologues and dialogues. Finally, we will argue that the nature of these complex interrelationships among acoustical factors, hearing status, and cognitive performance will force us to re-evaluate our notions of how information is processed to include a greater role for top-down, attentional control over auditory processing.

Bob Shannon

The Ear is for Music - The Brain is for Speech

Recent research on cochlear implants demonstrates an interesting dichotomy in auditory perception. Excellent speech recognition can be obtained with a coarse prosthetic signal in which most of the fine temporal information has been removed and much of the spectral pattern is heavily degraded. This result demonstrates that the pattern recognition in the brain for speech is so robust that it can obtain the intended linguistic message from a highly degraded signal from the ear. In contrast, harmonic pitch and melody information are eliminated by even mild degradations in cochlear processing. This dichotomy highlights the need for additional research to illuminate the relative roles of ear and brain in hearing. A better understanding of the relative contributions of the ear and brain is important for designing prosthetic devices and for designing rehabilitation strategies.

Jeff Bondy & Ian Bruce

From Neurons to Hearing Aids

Many individuals suffering from hearing loss find substantial benefit from a hearing aid in simple listening conditions, for example, talking with one other person in a quiet room. However, a hearing aid typically does not provide the same benefit in a more difficult listening situation, such as joining in a group discussion in a noisy environment. The normal hearing ear can adjust its operation to fit different environments, while hearing impairment makes the ear less adaptive. We hope to build new hearing aid algorithms that reestablish this environmental adaptation.

Compounding the complexity of different environments needing different hearing aid algorithms is that people with the same audiogram may need very different hearing aid fittings. We are trying to develop new hearing aids that take into account more than an audiogram. In order to deal with the intertwined environmental and pathology issues we utilize a computer model of the ear, analyzing the information that the brain receives about different sounds. In this talk I will discuss some of the effects hearing loss may have on the auditory brain, and hopefully give some insight into how special each hearing impaired person's type of loss is.