6
Psychophysics

Understanding how simple and complex electrical stimuli are perceived and their relationship with speech perception has been important in the development of cochlear implants, and was fundamental to the University of Melbourne's approach. Although the relationship between psychophysics and speech perception is not well defined, the perception of temporal and place pitch perception and loudness are key elements. Furthermore, the relationship between both pitch and loudness and speech features has become better understood, in part due to cochlear implant research. The research has also helped explain the effectiveness of the formant-based as well as fixed filter speech-processing strategies (Tong and Clark 1982; Tong et al 1982, 1983a,b), and has guided research leading to improvements in speech processing (Tong et al 1983b). It has also aided in explaining plasticity in children, and differences in their performance. This research will be needed in the development of bimodal and bilateral speech processing. The perceptual studies discussed below are important also in understanding the underlying electrophysiology.

Acoustic Stimulation

Frequency coding correlates predominantly with pitch perception. The time/period (rate) code for frequency results in temporal pitch, and the place code in place (spectral) pitch. As the perception of pitch is highly relevant to the perception of speech, early research at the University of Melbourne focused on the relative importance of the rate and place coding of frequency, and how well they could be reproduced by electrical stimulation.

With sound it is difficult to determine the relative importance of the time/period and place codes in the perception of pitch as the two codes operate together. The temporal responses of the neurons vary according to the site of excitation. On the other hand, with electrical stimulation of auditory nerves the two codes can be reproduced separately to study their relative importance.

Psychoacoustic data were also the basis for studying the percepts obtained with simple and complex electrical stimulation (Tong et al 1979; Tong et al 1982; Tong et al 1983). For more comprehensive descriptions of psychoacoustics the reader is referred to the texts of Stevens (1975), Plomp (1976), and Moore (1997).

Pitch and Timbre

Pitch is a basic percept underlying speech recognition, and is related to its intelligibility. Timbre is also important in speech perception and in musical appreciation.

Definition

Pitch is the subjective attribute of tones that correlates most closely with the physical dimension of frequency. It is also related to duration and intensity. In 1960 the American Standards Association defined pitch as "that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale."

Timbre relates to the quality of the sound. It depends primarily on the spectra of the stimuli (i.e., their harmonic content) and to a lesser extent on their relative phases. Tones with strong lower harmonics sound more mellow. The American Standards Association defined timbre as "that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar."

Measurement

Pitch is measured in units referred to as mels (Stevens and Volkmann 1940). The pitch in mels is determined by scaling the pitch of a frequency relative to the pitch of the reference frequency of 1000 Hz. For example, the pitch of the reference frequency of 1000 Hz is 1000 mels (Fig. 6.1); thus a tone with a pitch of 500 mels sounds half as high in pitch as a tone of 1000 mels.

Scaling

Pitch scaling can be done by asking subjects to adjust a variable tone until its pitch appears to be half that of a fixed tone (fractionation) or to adjust a variable tone to a pitch halfway between the pitches of two fixed tones (bisection). As seen in Figure 6.1 the pitch of pure tones does not vary linearly as a function of frequency.

A single numerical estimation method can also be used to measure pitch. The subjects are instructed to assign a number in the range 1 to 100 for the pitch of a single presentation. The scale could be expanded in either direction if required. This procedure is based on the research of Stevens and Greenbaum (1966) on magnitude estimation in matching numbers to loudness.

Intensity

Pitch can affect loudness and vice versa. For example, increasing the intensity of a sound not only increases the loudness, but also there may be a small change in

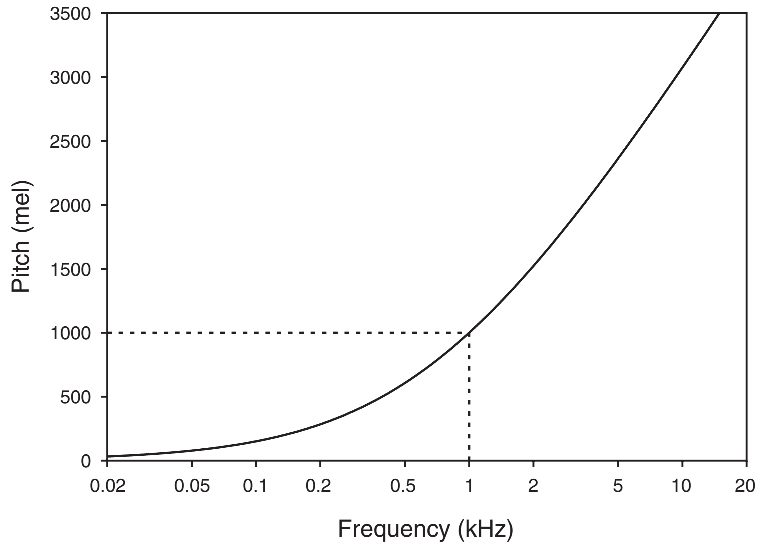

FIGURE 6.1. The pitch in mels versus frequency (Stevens and Volkmann 1940).

pitch. Stevens (1935) noted that increasing the intensity decreased pitch for tones less than 800 Hz and increased pitch for tones above 3000 Hz. Later studies showed the effects highly dependent on the subject, but overall supported the findings of Stevens even though the effects were small (Morgan et al 1951; Cohen 1961; Verschuure and van Meeteren 1975). For frequencies below 2000 Hz there can be a maximum 5% decrease in pitch with an increase in intensity, and a 5% increase in pitch for frequencies above 4000 Hz (Moore 1997). Fortunately, with music, intensity does not influence the pitch of the complex tones of the instruments.

Duration

Recognizing pitch requires a tone of minimal duration. Regardless of frequency, durations of only a few milliseconds are heard as clicks. As the duration is lengthened the click develops a tonal quality that allows some listeners to discriminate among clicks on the basis of click pitch. The minimum period of time for a frequency to be perceived as pitch varies with frequency below 1000 Hz when two to three cycles are needed. Above 1000 Hz a minimum duration of 10 ms is required.

Selectivity

Frequency selectivity is the ability to select separate pitch for one frequency from that at another when the two are sounded together. The ability to select the different harmonics in a complex sound is related to the critical band. The critical band is the frequency region in which there is masking of a pure tone, the loudness of a band of noise of fixed intensity remains constant, and phase differences in amplitude and frequency modulated sounds are detected. It is described in some detail below (see Critical Band and Ratio).

Plomp (1976) found that a harmonic could be heard from a multitone complex only if that harmonic were separated from neighboring harmonics by about one critical band. For a two-tone complex the harmonics could be identified for a separation less than this. Studies by Plomp (1964) and Plomp and Mimpen (1968) have shown that from five to eight harmonics are the maximum number that can be detected from a complex tone. The above studies were important for the development of a multiple-channel implant. They suggested that electrodes need not be spaced closer than a critical band for frequency selectivity, but would need to be closer if phase differences were important in the perception of frequencies within the critical band.

Discrimination

Frequency discrimination is the ability to distinguish a change or difference in frequency when the two tones are played one after the other. It must not be confused with frequency selectivity. Smaller differences in frequency can be detected if the frequencies are played one after the other (discrimination), than if they are to be selected from the sounds presented simultaneously (selectivity). Discrimination is measured as a difference limen (DL). The absolute DL is the smallest detectable change in frequency (Δf). The relative DL is the change relative to the initial frequency ($\Delta f/f$). The DL was measured initially by the modulation method (Shower and Biddulph 1931). This involved modulating frequencies up and down about four times per second. The degree of modulation required to detect a frequency variation is the DL. It is not the more reliable method as the frequency changes are accompanied by loudness fluctuations. The better method is to present two sounds, and to ask the subject to determine which is the higher pitch (a two-alternative forced choice task). The DL is the frequency separation where the subject achieves 75% correct. When measured this way, the results show that the relative DL ($\Delta f/f$) is fairly constant from 50 Hz to 5000 Hz (i.e., it obeys the Weber-Fechner law), and is approximately 0.2% (Moore 1974; Wier et al 1977), but it varies with the duration of the tone. The ability to discriminate frequency is remarkably good with DLs as small as 2 Hz for a 1000-Hz tone.

It was important for understanding the coding of frequency to know whether the detection of a change in frequency could be explained in terms of place or temporal frequency codes. It was also essential that some of the first psychophysical studies on the first University of Melbourne's patient were to determine whether changes in rate or place could be perceived for different stimulus durations required for the perception of vowels and consonants (Tong et al 1982).

With regard to place coding, the tuning curves of cells in the auditory brainstem are too broad to account for the small size of the frequency DL. However, a model was developed by Zwicker (1970) that predicted a change in frequency could be detected whenever the site of neuronal excitation varied by more than a threshold value. The greatest difference between the two patterns would occur on the steeply sloping high-frequency side. The steepness of the tuning curve was shown by Maiwald (1967) to be constant when expressed in units of the critical bandwidth. This predicted that frequency DLs at any frequency should be a constant fraction of the critical bandwidth at that frequency. Nevertheless, this has not been found to be the case. Moore and Glasberg (1989) discovered with tones randomized in intensity that the DLs were lower than the model predicted for frequencies from 500 to 4000 Hz, but were consistent at higher frequencies.

Temporal frequency coding is an alternative explanation for DLs, but the DL at 1000 Hz requires an accuracy of about 2 μs in the neural processing of information. This is unlikely, as there is a jitter in the initiation of nerve impulses with a standard deviation of 100 μs for a 1000-Hz tone (de Boer 1969). It is more likely that both place and temporal coding together are the mechanisms involved, and they may contribute to different degrees over the whole frequency range.

Complex Tones

Sounds composed of more than one sinusoid are complex tones. Ohm's acoustic law states that the ear can separate a complex tone into the individual sinusoids so that the listener can perceive the separate components. The law does not always hold, as the listener may not hear all the components or may hear different ones.

When two tones are presented together, a number of phenomena occur. If they are widely separated in frequency they are heard as two distinct tones on the basis of Ohm's law. If they are very close (e.g., 1000 Hz and 1003 Hz), a tone will be heard that beats at a frequency that is the difference between the two frequencies (3 Hz). Beating is used to tune pianos.

If two tones have frequencies that are not too close, other sounds of low intensity not in the signal can be detected. These are summation or difference tones and are referred to as combination tones. The most common one is the low pitch called the missing fundamental or the residue pitch (Schouten 1940a,b). The pitch persists even when the low frequencies are masked, and this is presumed to arise from the high-frequency end of the cochlea. This can be explained through the physiology of temporal coding, as Rose et al (1967) have recorded the intervals between nerve action potentials for two tones, and found a dominant interval equaling the missing fundamental from the high-frequency neurons.

Other tones not present can also be heard if a probe tone is used to beat with them. These combination tones occur at higher intensities and can be masked at their frequency regions on the basilar membrane. Ones most commonly heard are $f_2 - f_1$, $2f_1 - f_2$ and to a lesser extent others of the class $f_1 - n\,(f_2 - f_1)$, where f_1 and f_2 are the first and second harmonics respectively. They are produced by a nonlinearity of the middle ear and basilar membrane. The $2f_1 - f_2$, however, can be heard at low intensities and cannot be masked on the basilar membrane location.

Springer

http://www.springer.com/978-0-387-95583-4

Cochlear Implants
Fundamentals and Applications
Clark, G.
2003, XXXVIII, 831 p., Hardcover
ISBN: 978-0-387-95583-4