TEMPORALLY JITTERED SPEECH PRODUCES PI-PB ROLLOVER IN YOUNG NORMAL-HEARING LISTENERS

Terence T. Miranda and M. Kathy Pichora-Fuller
The University of British Columbia, 5804 Fairview Avenue, Vancouver, BC, V6T 1Z3

1. INTRODUCTION

In general, elderly listeners, even those with audiograms in the normal range, have great difficulty when trying to understand language spoken in background noise (e.g., CHABA, 1988). One possible explanation for their poor performance is the existence of an auditory temporal processing deficit such as neural asynchrony (e.g., Pichora-Fuller & Schneider, 1992). Recent studies in our lab have used temporally jittered SPIN-R sentences to test young, normal hearing adults in a range of signal-to-noise (S/N) ratios (e.g., Pichora-Fuller, Schneider, Pass & Brown, submitted). One conclusion from these studies was that the external jitter introduced into the stimuli seemed to simulate, in young normals, the increased internal neural asynchrony hypothesized to exist in the elderly. It is also possible that the results of Pichora-Fuller et al (submitted) may be due to something other than the simulation of neural asynchrony. For example, the jitter may be simulating a type of auditory processing deficit other than, or in addition to, neural asynchrony and it is this other deficit that is behind the pattern of results.

This study uses a different approach to help answer the question of whether or not temporal jitter simulates neural asynchrony in young, normal hearing listeners. The question is addressed by temporally jittering word discrimination lists and presenting them to young adults with normal hearing to see if they will show performance-intensity phonetically balanced (PI-PB) rollover. PI-PB rollover occurs when word discrimination scores decrease with increases in presentation level. There exists substantial clinical and theoretical evidence to suggest that neural PI-PB rollover (e.g., the type found in acoustic neuroma cases) is due to increased neural asynchrony in the auditory system. Results from a number of studies (e.g., Jerger and Jerger, 1971; Meyer and Mishler 1985) provide clinical evidence suggesting a connection between the measure of PI-PB rollover and the existence of neural asynchrony in both the acoustic neuroma and elderly populations. A theoretical link between PI-PB rollover and neural asynchrony can be established using the Average Localized Synchronized Rate (ALSR) computational model of Young and Sachs (1979).

2. METHODS

2.1 Participants

Sixteen young listeners (mean = 27.3 years, SD = 3.5 years), eleven females and five males, were tested. All were native English speakers, had normal middle ear function, and had bilateral pure-tone air-conduction thresholds at 0.25, 0.5, 1, 2, 4, and 8 kHz less than or equal to 20 dB HL. Each participant gave informed consent and received remuneration of $10 following completion of each experimental session.

2.2 Design

Each participant attended two sessions of 1 hour each. The sessions were separated by at least one week to reduce the effects of practice. Uncomfortable listening level (UCL) for speech was determined for each participant and PI-PB functions were created by measuring speech discrimination scores at 40, 55, 65, and (UCL-5) dB HL in each of three conditions: one intact and two different jittered conditions. Participants were randomly assigned to one of four groups, with four participants per group. Experimental conditions (e.g., word list presentation, intensity presentation) were counterbalanced between and within groups. Speech discrimination was tested in the intact and first jitter condition using 50-word lists of Northwestern University Auditory Test No. 6 (NU6). The second jitter condition was tested using 50-word Central Institute for the Deaf (CID) W22 lists.

2.3 Procedures

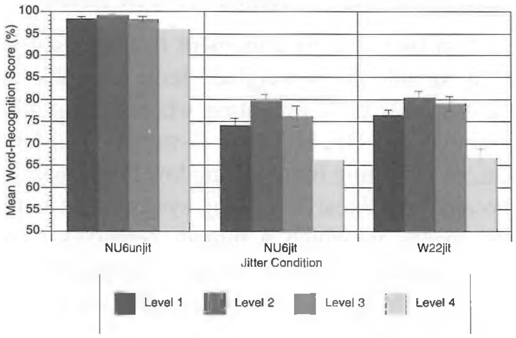
Digitized CD recordings of all stimuli were fed from a JVC XL-Z232 compact disc player, into a Grason-Stadler GSI-16 audiometer and then into TDH-50P headphones (left ear only). In order to prevent crossover, speech noise was delivered to the right ear at an effective masking level. All equipment was calibrated to ANSI 3.6 1969/ISO 389 1975 standards.

2.4 Stimuli

The intact NU6 and W22 word lists were purchased on a commercial CD recording (Auditec of St. Louis). Each target word and the accompanying carrier phrase was extracted from the original CD, redigitized at a sampling rate of 20 kHz and saved as a *.SND file on a PC hard drive. These soundfiles were used to produce a jittered version of the word discrimination lists. A Fast Fourier Transform (FFT) was used to separate the incoming signal into its component frequencies. For the first jitter condition, the speech (NU6 lists) was then divided into two bands, one above and the other below 1.2 kHz. For the second jitter condition, the speech (W22 lists) was divided into four bands, one above 1.2 kHz, and three below 1.2 kHz (0-.4, .4-.8, and .8-1.2 kHz). For both jitter conditions, only the components below 1.2 kHz were jittered. Jittering was accomplished using in-house software (Jaeger, 2000). Using our method, the sequence of amplitude values in the soundfile is altered by shifting them by delay values. The delay value applied to each sample is determined using a low-pass (LP), band-limited white noise model. Such a noise has amplitude values that are normally distributed with a mean of 0 and a specified standard deviation (SD). Bandwidth (BW) represents the upper cut-off frequency of the LP band-limited noise. The higher the BW value, the more rapid are the changes in amplitude of the noise. The larger the SD, the greater the range of the delay values that can be used in jittering the signal. For each data point in the digitized sound file, the program selects a delay value by referring to a noise generated with an experimenter-specified SD and BW, determining the amplitude value of the noise at the corresponding point in time, and then converting this amplitude into a delay value. The delay value determines the position (in time) of the sample in the original file whose amplitude value is to be substituted in for the data point under consideration. For both jitter conditions, the specified values for jittering were SD = 0.50 msec and BW = 0.5 kHz. In the first jitter condition, all signal components below 1.2 kHz were jittered using one noise exemplar. In the second jitter condition, each of the three bands below 1.2 kHz was jittered using a unique noise exemplar; therefore, at any given sample point, the delay value applied to one band was independent of the delay value applied to the other two bands. After jittering, the jittered low-frequency band(s) and the intact high-frequency band were recombined, re-sampled at 44.1 kHz, saved as *.WAV files, and written back to CD, with a different CD for each of the three conditions. The 1.0 kHz calibration tone was saved to each CD and used for calibration.

3. RESULTS

Speech discrimination scores were measured for each of the 16 participants using intact NU6, jittered NU6 and jittered W22 word lists presented at four levels, 40, 55, 65, and (UCL-5) dB HL, denoted respectively as Level 1, Level 2, Level 3 and Level 4. These results are summarized in Figure 1. While the scores for the intact lists remain high as presentation level increases, scores for the jittered lists decline as presentation level increases. For each PI-PB function (one per condition, 3 per participant, 48 in total), rollover was calculated (Rollover = PBmax – PBmin, where PBmin is the lowest score obtained at an intensity greater than the intensity at which PBmax is observed). Figure #1 clearly shows that the average amount of rollover obtained in the jittered conditions is substantially greater than the rollover in the intact condition. This is confirmed by an ANOVA with group as a between-subjects factor and jitter condition as a within-subjects factor. There was no significant main effect of group on rollover [F(3,12)=0.42, p=0.75], but there was a significant main effect of jitter condition [F(2,24)=27.91, p<0.001]. A Student Newman-Keuls test confirmed that rollover in the intact condition was significantly (p<0.001) less than in the jitter conditions which did not differ significantly from each other.



4. CONCLUSIONS

Because the participants in the present study had ipsilateral acoustic reflexes within the normal range and did not exhibit rollover when intact speech was presented, natural mechanical and neural bases for the rollover are ruled out. Thus, the rollover that was observed must be attributed to the simulation of neural asynchrony that involved externally jittering the signal. Importantly, the simulated asynchrony disproportionately disrupted speech discrimination at high presentation levels, thereby ruling out the possibility that the jittering simply degraded the signal in a level-independent fashion. In contrast, other kinds of signal degradation, such as low-pass filtering, would be expected to yield better scores at high presentation levels than at lower levels. Consistent with the clinical and theoretical considerations presented in the introduction, the results of the present study support the hypothesis that PI-PB rollover is due to disruptions of synchrony coding. Furthermore, the present findings support the theoretical notion that synchrony coding plays an important role in the perception of high-level speech. The lack of difference between the two jitter conditions suggests that various types of asynchrony could produce PI-PB rollover and further research will be required to determine how to characterize the exact nature of the asynchrony found in particular pathologies or individual cases.

5. REFERENCES

CHABA: Committee on Hearing, Bioacoustics, and Biomechanics, Working Group on Speech Understanding and Aging. (1988). Speech understanding and aging. JASA. 83: 859-895.

Jaeger CP. (2000). Jitter14.exe [computer program]. Vancouver: CJA Consulting.

Jerger J, Jerger S. (1971). Diagnostic significance of PB word functions. Arch Otolaryngol 93:573-580.

Meyer DH, Mishler ET. (1985). Rollover measurements with Auditec NU-6 word lists. J Speech Hear Disord 50:351-356.

Pichora-Fuller MK, Schneider BA. (1992). The effect of interaural delay of the masker on masking-level differences in young and old adults. JASA. 91(4): 2129-2135.

Pichora-Fuller MK, Schneider BA, Pass HE, Brown S. (Submitted). Temporal jitter disrupts speech intelligibility and memory.

Young ED, Sachs MB. (1979). Representation of steady-state vowels in the temporal aspects of the discharge patterns of auditory-nerve fibers. JASA. 66:1381-1403.