The relationship between pure-tone sequential stream segregation and perceptual separation of male and female talkers by listeners with hearing loss

Carol L. Mackersie

Department of Communicative Disorders
San Diego State University
San Diego, CA 92101
cmackers@mail.sdsu.edu

ABSTRACT

The purpose of this study was to describe the relationship between sequential stream segregation abilities of listeners with hearing loss and the ability to recognize pairs of sentences spoken simultaneously by a man and a woman. In the streaming task, the fusion threshold was measured as the frequency separation at which listeners could no longer perceptually separate a series of fixed and varying-frequency pure-tones. The varying frequency tones started at frequencies either below (ascending) or above (descending) the frequency of the fixed tone. Seven of 11 subjects showed significant differences between the ascending and descending fusion thresholds. Ascending fusion thresholds were found to predict the intelligibility of the male talker, but not the female talker. Conversely, descending fusion thresholds were found to predict the intelligibility of the female talker, but not the male talker. In both cases, higher speech perception scores were associated with lower (better) fusion thresholds.

1. INTRODUCTION

Individuals with hearing loss often report difficulties understanding speech in a competing background of other talkers. Moreover, the substantial individual differences in speech understanding in competing backgrounds is not well predicted by the pure-tone audiogram [1,2]. These individual differences have been attributed to differences in supra-threshold auditory capacities [1,2,3] as well as individual differences in the ways in which listeners use cues present in the incoming signal [4,5,6].

The hearing aid industry has targeted the pervasive problem of reduced speech intelligibility in noise by developing instruments that incorporate advanced signal processing algorithms and directional microphones. Despite these advances, individuals with hearing loss continue to experience difficulties understanding speech in challenging listening environments. Further progress in rehabilitative strategies may arise from the consideration of factors related to the listeners themselves. The identification of factors underlying individual differences in the ability to understanding speech in the presence of competing speech is an important component of this process.

Recently, we examined the relationships between simultaneous sentences perception by listeners with hearing loss and two psychoacoustic measures: frequency selectivity and pure-tone sequential stream segregation [7]. The recognition of two competing sentences was strongly related to the ability to segregate two pure-tone pitch patterns as measured by the sequential stream segregation task. Simultaneous sentence perception was not related to frequency selectivity, however, suggesting that peripheral masking was not a key factor in recognition of the two competing sentences.

A common element of the streaming and simultaneous sentence task is the presence of pitch cues, which are present in the changing frequencies in the streaming task and in the intonation patterns of the talkers in the simultaneous sentence recognition task. To the extent that listeners rely on these pitch cues, both tasks may involve the processing of pitch patterns over time. The role of pitch information in the separation of simultaneous talkers has been demonstrated by studies showing that increasing the difference between fundamental frequencies of competing sentences improves the recognition of the sentences by both listeners with normal hearing and listeners with hearing loss [8,9,10].

This paper presents further observations regarding relationships between simultaneous talker separation and stream segregation by listeners with hearing loss based on a subset of data from our previous study [7]. The paper addresses differences between simultaneous sentence recognition of a male versus female talker, differences between ascending and descending frequency patterns in a stream segregation task, and the relationships between talker gender and streaming patterns.

2. METHODS

2.1 Participants

The 11 participants of the experiment had mild-moderate sensorineural hearing loss with pure-tone averages (500, 1000, 2000 Hz) ranging form 33-47 dB HL and monosyllabic word recognition scores in quiet of 76% or greater. Participants ranged in age from 51-87 years with a mean of 73 years. Mean thresholds ± 1 standard deviation are shown in Figure 1.

2.2 Simultaneous sentence recognition

The simultaneous sentences materials consisted of sentence pairs matched in length and equalized to the same average rms level (Leq). One sentence of each pair was spoken by a female talker and the complementary sentence was spoken by a male talker. Stimuli were played through a Creative-Labs SB-16 Sound Card and routed to the tape input of a Grason-Stadler GSI-16 audiometer. The output of the audiometer was delivered monaurally to TDH-50 earphones. Listeners were tested at an Leq level of 25 dB SL relative to their thresholds at 2000 Hz.

Four lists were administered to each participant in random order. Listeners were instructed to repeat both sentences of each pair. The percentages of words repeated correctly were

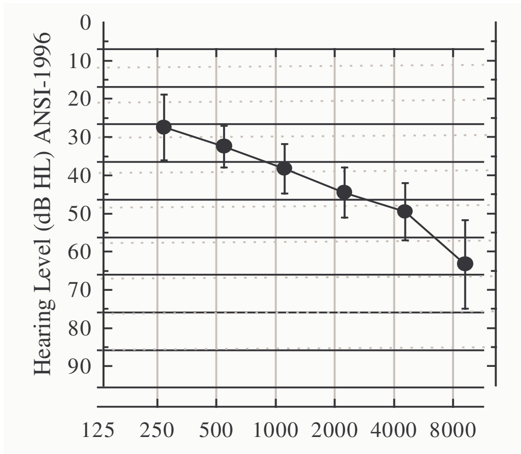

Figure 1. Mean pure-tone audiometric thresholds for the 11 participants. Error bars show + / - 1 standard deviation.

calculated separately for the male and female talkers'. Data were arcsine transformed before statistical analysis to increase the stability of the error variance [11].

2.3 Sequential stream segregation

Sequential stream segregation was measured using the method described by Rose and Moore (1997). A 1000 Hz tone (" A") was alternated with a tone whose frequency varied ("B"). Each tone had a total duration of 100 ms including 10 ms rise/fall times. The tone sequences were presented in an ABA_ABA_ABA_… pattern, with 20 ms separating each tone and a 120 ms silence separating each triplet.

Tone B changed in either an ascending or descending pattern starting at a frequency either lower than or higher than the fixed tone. The starting frequencies of tone B were selected randomly from the range between 100-200 Hz for the ascending series and from a range between 2000-3000 Hz for the descending series. The frequency difference between the fixed 1000 Hz tone and varying tone was gradually decreased using a successively decreasing step size. The step size was determined using an equation described by Rose and Moore [12]. This equation is

$$F_{vary'} = F_{fix} (F_{vary} / F_{fix})^{(0.91)} \quad (1)$$

where $F_{vary'}$ is the frequency of the varying tone in the (i + 1) ABA sequence, F_{fix} is the frequency of the fixed tone, and F_{vary} is the frequency of the varying tone in the ith ABA sequence. The ABA sequence was played four times at each frequency separation. Stimuli were presented at 25 dB SL relative to the threshold for the 100 ms 1000 Hz tone.

Following a practice session, listeners heard 12 ascending and 12 descending series, alternating between ascending and descending series in blocks of four. Listeners pushed a "Start" button on a computer keyboard to begin each series. They were instructed to listen to the changing tone and follow it as long as it could be heard as separate from the fixed tone. Listeners pressed a "Stop" button on the keyboard when they could no longer perceptually separate the two tonal patterns. At this point, the subjects heard a single tonal series with a "galloping" pattern. The frequency separation at which subjects could not longer separate the tones was taken as the listener's fusion threshold. Mean fusion thresholds were calculated separately for the ascending (rising pitch) and descending (falling pitch) patterns.

3. RESULTS AND DISCUSSION

3.1 Sequential stream segregation

The group means ascending and descending fusion thresholds are shown in the left panel of Figure 2. The 0.2 semitone difference between ascending and descending thresholds is not statistically significant [t (10) = 0.34, $p > .05$].

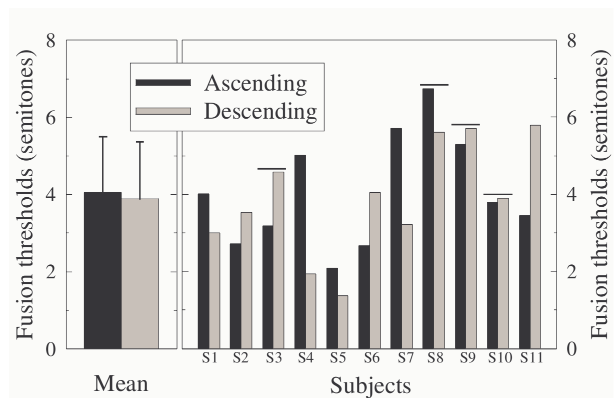

Figure 2. Group mean ascending and descending fusion thresholds are shown in the left panel. Error bars indicate +1 standard deviation. Individual fusion thresholds are shown in the right panel. Horizontal lines indicate no significant difference between thresholds.

Mean ascending and descending fusion thresholds for individual listeners are shown in the right panel of Figure 2. Analyses of individual listener data show that 7 of the 11 listeners have significantly lower (better) fusion thresholds for either the ascending or descending patterns ($p < .05$). As shown in Figure 2, four of the listeners had higher fusion thresholds for the ascending series (S1, S4, S5, S7) and three of the listeners had higher fusion threshold for the descending series (S2, S6, S11) indicating that some listeners are better at maintaining perception separation of pitch patterns when the changing pitch pattern is rising whereas others are better at maintaining separation when the pitch pattern is falling. Rose and Moore [12] also showed differences between ascending and descending fusion thresholds for some of their listeners with bilateral hearing loss, although an analysis of these differences was not reported. It is difficult to account for individual differences between ascending and descending fusion thresholds in the current study on the basis of differences in auditory sensitivity, as the audiograms of the 11 participants were fairly similar. It is possible that the individual differences reflect differences in the weighting of pitch information among listeners.

.

3.2 Relationship between sentence recognition scores and fusion thresholds

Figure 3 shows the relationship between scores for the male talker and the two streaming thresholds. There was a significant correlation between recognition scores and the ascending fusion thresholds (right panel: $r = -.81$, $p < .01$), but not the descending thresholds (left panel: $r = -.36$, $p > .05$).

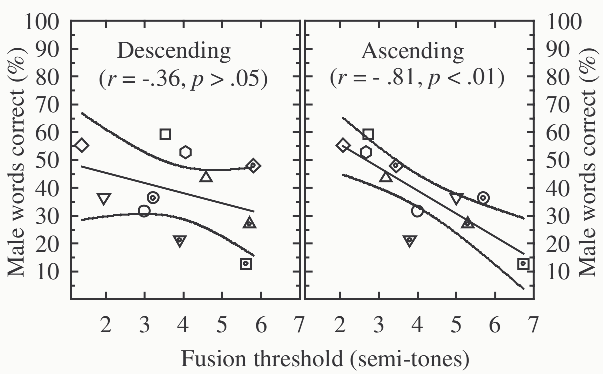

Figure 3. Sentence recognition scores for the male talker as a function of fusion thresholds for the ascending and descending patterns are shown in the right and left panels, respectively. Linear regression functions and 95% confidence intervals are shown in each panel.

The opposite pattern was observed for the female talker, shown in Figure 4. Recognition scores for the female talker were strongly related to the descending fusion thresholds ($r = -.82$, $p < .01$), but not to the ascending thresholds ($r = -.35$, $p > .05$). For both talkers, higher speech recognition scores are associated with lower (better) fusion thresholds. That is, there was a strong tendency for listeners who were better able to perceptually separate and identify the two sentences to also be better at separating interleaving pitch patterns in the pure-tone streaming task.

These patterns might be explained by considering the role of attention in these tasks. In the sentence recognition task, listeners must extract and attend to the stream of acoustic and linguistic information produced by the talkers. To the extent that a listener uses a talker's intonation patterns as one of the cues to facilitate this process, the listener may adopt a strategy in which he/she attends to the lower frequency intonation patterns of the male talker and ignores the higher frequency intonation patterns of the female talker when extracting the male talker's speech. When extracting the female talker's speech, the listener may use a strategy that involves attending to the higher frequency intonation patterns of the female talker while ignoring the lower frequency intonation patterns of the male talker.

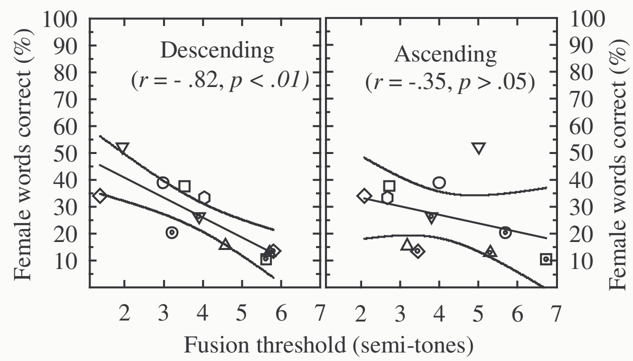

Figure 4. Sentence recognition scores for the female talker and fusion thresholds for the ascending and descending patterns are shown in the right and left panels, respectively. Linear regression functions and 95% confidence intervals are shown in each panel.

The role of attention in sequential stream segregation is somewhat controversial. Some maintain that attention is not required to perform a streaming task [13,14[. There is some evidence, however, that attention does plays a role in sequential stream segregation. Carlyon and his colleagues, for example, showed that the build-up of streaming was reduced when subjects attended to a competing task in the opposite ear [15].

To the extent that attentional processes were engaged during the streaming task in the current study, listeners may have focused their attention on the rising or falling pitch pattern while ignoring the fixed-frequency tones. A strategy of attending to the low frequency ascending pattern while ignoring the higher frequency fixed tone, for example, corresponds to a strategy of attending to the lower frequency intonation patterns of the male talker in the presence of the higher frequency intonation patterns of the female talker. This explanation is consistent with the finding that the ascending fusion thresholds are highly correlated with the recognition scores for the male talker, but not the female talker.

4. CONCLUSIONS

These findings support the conclusion from our earlier report that the dynamic process involved in streaming might also underlie a listener's ability to use intonation contours to facilitate the perceptual separation of two simultaneous talkers. In addition, the present findings suggest that the importance of streaming in the perceptual separation of talkers depends on the nature of the information provided by the changing pitch stream. Differences in the way listeners use this information may be related to individual differences in the perceptual weighting of different pitch regions. Researchers interested in improving performance of individuals with hearing loss in difficult listening situations must take account of individual differences in the perceptual strategies used in auditory processing.

5. REFERENCES

[1] Festen, J.M. & Plomp, R. (1983). "Relations between auditory functions in impaired hearing". *J. Acoust. Soc. Am., 73,* 652-662, 1983

[2] Glasberg, B., & Moore, B. "Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech". *Scand. Audiol., 32,* 1-25, 1989.

[3] Moore, B. "Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids". *Ear & Hear., 17*(2), 133-61.

[4] Doherty, K., & Lutfi, R. "Spectral weights for overall level discrimination in listeners with sensorineural hearing loss". *J. Acoust. Soc. Am., 99*, 1053-1058, 1996.

[5] Hedrick, M., & Jesteadt, W. "Effect of relative amplitude, presentation level, and vowel duration on perception of voiceless stop consonants by normal and hearing- impaired listeners". *J. Acoust. Soc. Am., 100,* 3398-3407, 1996.

[6] Summers, V., & Leek, M. (1992). "The role of spectral and temporal cues in vowel identification by listeners with impaired hearing". *J. Speech Hear. Res., 35,* 1189-1199, 1992.

[7] Mackersie, C., Prida, T., & Stiles, D. (2001). "The role of sequential stream segregation and frequency selectivity in the perception of simultaneous sentences by listeners with sensorineural hearing loss". *J. Speech Hear. Res., 44,* 19-28, 2001.

[8] Bird, J. & Darwin, C.J. "Effects of a difference in fundamental frequency in separating two speech messages". In A.R. Palmer, A. Rees, Q. Summerfield, & R. Meddis (Eds) *Psychophysical and Physiological Advances in Hearing.* (pp. 263-269). Whurr: London, 1998.

[9] Brokx, J., & Nooteboom, S. "Intonation and the perceptual separation of simultaneous voices". *J. Phonetics, 10,* 23-36, 1982.

[10] Summers, V., & Leek, M. "Fo processing and the separation of competing speech signals by listeners with normal hearing and with hearing loss". *J. Speech Hear. Res., 41,* 1294-1306, 1998.

[11] Studebaker, G. "A 'rationalized' arcsine transform". *J. Speech Hear. Res., 28,* 455-462, 1985.

[12] Rose, M., & Moore, B.C.J. "Perceptual grouping of tone sequences by normally hearing and hearing-impaired listeners". *J. Acoust. Soc. Am., 102,* 1768-1778, 1997.

[13] Bregman, A. (1990). *Auditory Scene Analysis.* Cambridge, MA: MIT Press.

[14] Moore, C., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception & Performance, 23,* 339-352.

[15] Carlyon, R. P., Cusack, R. F. J. M., & Robertson, I. H. "Effects of attention and unilateral neglect on auditory stream segregation". *J Exp. Psych.: Hum. Percept. & Perf.., 27*, 115-127, 2001