The Effect of Fundamental Frequency on Intelligibility in Simulated Cochlear-Implant Listening

Nicole M. Scherrer, Christopher A. Brown, and Sid P. Bacon
Department of Speech and Hearing Science, Arizona State University, Tempe AZ

ABSTRACT

• Individuals with residual hearing only at low frequencies are candidates for electric-acoustic stimulation (EAS). When low-frequency acoustic information is added to high-frequency electric stimulation, speech recognition often improves dramatically. This may reflect the availability of fundamental frequency (F0) information in the acoustic region. The purpose of the present study was to determine (1) whether similar benefit exists when replacing low-frequency speech with a tone that tracks F0 and (2) whether that tone can be shifted to lower frequencies and provide similar benefit. A 4-channel vocoder simulated electric stimulation. A female talker (mean F0 = 184 Hz) was presented in various backgrounds. Intelligibility improved when a tone tracking F0 was added to vocoder stimulation, even at shifts of 100 Hz (mean F0 = 84 Hz), as long as the sensation level of that tone was comparable to that of the unshifted tone. This confirms the importance of F0 information (at least under simulated EAS), and also indicates that significant information can be provided by a tone that tracks F0. That the mean frequency of this tone can be shifted to extremely low frequencies suggests that individuals may only need residual hearing up to about 100 Hz to benefit from EAS.

INTRODUCTION

• Individuals with residual hearing restricted to the low-frequency region (below 500 – 750 Hz) now benefit from the combination of electrical stimulation in the mid- to high-frequency range with acoustic stimulation in the low frequencies (electric-acoustic stimulation; EAS).

• Speech recognition by both real and simulated EAS listeners is significantly improved with the addition of low-frequency acoustic information, particularly in a competing background (Qin & Oxenham, 2006; Turner *et al*., 2004).

• It has been suggested that this benefit is due to the availability of fundamental frequency (F0) information that occurs across an utterance.

• Pilot data from our lab using a 4-channel vocoder simulation of cochlear implant processing in normal-hearing listeners are consistent with this.

• Figure 1 shows significant improvement in intelligibility when a tone modulated with the dynamic changes in F0 was added to vocoder stimulation (V/TF0).

• Further benefit was seen when the F0 tone was modulated by the amplitude envelope of the low-pass speech (V/TF0-env).

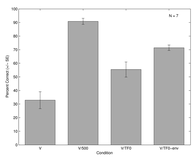

Fig. 1. Preliminary results showing the effects of adding a low-frequency modulated tone to higher frequency vocoder stimulation. The target was a female talker. The background was a male talker.

The aims of this investigation were to determine:

(1) if the benefit in intelligibility that is observed when low-pass speech is added to vocoder stimulation can be achieved by replacing the low-pass speech with a tone that is modulated in both frequency and amplitude (F0-env)

(2) if this benefit is observed when the mean F0 of the tone is shifted down in frequency.

• If improvement is observed at shifted frequencies, this may suggest that individuals who have residual hearing restricted to very low frequencies (perhaps up to only 100 Hz) may stand to benefit from EAS.

METHOD 1

• A 4-channel vocoder was created by band-pass filtering the female target's speech into 4 bands, extracting the amplitude envelope from each band, and then using it to modulate four pure tones – each centered within one of the four bands. The frequencies were 992, 1633, 2687, and 4421 Hz (see top row of Fig. 2).

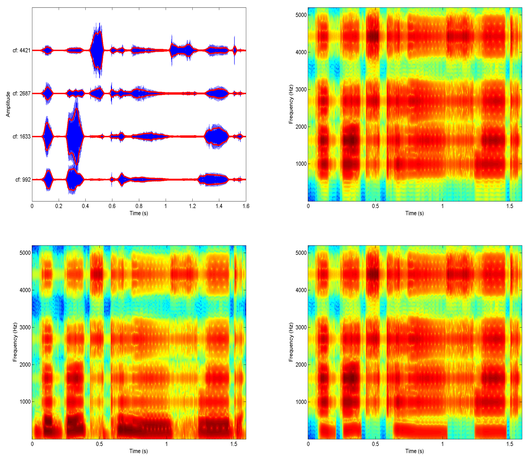

Fig. 2. Top row: output of a 4-channel vocoder as a time waveform (left) and spectrogram (right). Bottom row: spectrograms for the same 4-channel vocoder, with either low-pass speech (left) or a frequency-modulated tone (right) added.

• Prior to testing, the time-varying fundamental (F0) was extracted for each target sentence.

• 13 normal-hearing listeners were tested. Each listener was first presented the female target (mean F0 of 184 Hz) combined with a male distractor at various signal-to-noise ratios (SNRs) to determine the SNR that would produce 30% correct sentence recognition.

• The distractor was either a different female talker (mean F0 of 235 Hz), a male talker (mean F0 of 127 Hz), multi-talker babble, or speech-shaped noise.

• Processing conditions consisted of a vocoder-alone condition and vocoder plus a tone that was modulated in frequency with the extracted F0 information.

• The tone had a mean F0 that was either 184 Hz (the correct mean F0) or shifted downward in frequency to either 159, 134, 109, or 84 Hz.

• In some conditions, the frequency-modulated tone was also amplitude modulated by the envelope of the target speech below 500 Hz (F0-env).

• Participants responded verbally and percent correct scores were calculated.

RESULTS 1

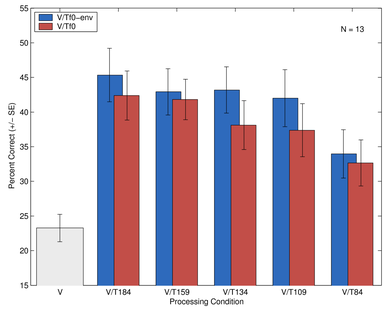

Fig. 3. Group means for each processing condition in experiment 1.

• The pattern of results was similar across background, so results were collapsed across this variable.

• Performance in the V/TF0-env condition was significantly better than that in the V/TF0 condition.

• The presence of the tone (F0 or F0-env) in the low-frequency region resulted in significant improvement in speech intelligibility (re V alone) at each mean F0.

• This improvement was not significantly affected by shifts in mean F0, except when the shift was 100 Hz.

DISCUSSION 1

• One possible reason that F0 improves performance is that the dynamic changes in frequency that naturally occur across an utterance may help distinguish the target speech from the background by providing cues for segregation (Qin & Oxenham, 2003; 2006).

• Given the high-pass nature of the low-frequency region of the normal audibility curve, can the decrease in performance at 84 Hz be explained by reduced audibility?

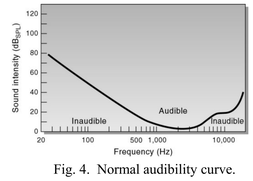

Fig. 4. Normal audibility curve.

METHOD 2

• 21 normal-hearing listeners were tested following the experimental design previously outlined.

• Prior to testing, participants' quiet thresholds were measured for 200-ms pure tones with frequencies of 184 and 84 Hz.

• The target was a female talker. The distractor was multi-talker babble.

• The processing conditions were vocoder-only and vocoder plus TF0-env.

• The mean F0 was either 184, 159, 134, 109, or 84 Hz.

• In addition, the SL of the 84-Hz modulated tone was adjusted to be equal to the SL of the 184-Hz tone.

• Participants responded verbally and percent correct scores were calculated.

RESULTS 2

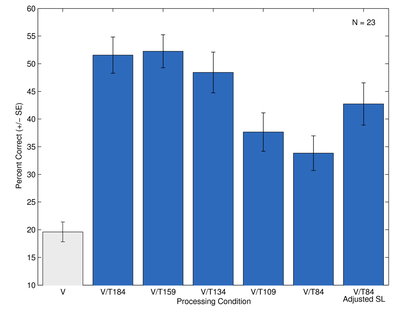

Fig. 5. Group means for each processing condition in experiment 2.

• Performance in all V/T conditions was significantly improved over vocoder-alone (V).

• Performance in the V/T84 condition was significantly worse than in V/T184, V/T159, and V/T134.

• However, when the 84-Hz tone was adjusted in SL to be equal to the 184-Hz tone (V/T84-adjusted), performance improved to equivalent levels (compare V/T184 and V/T84-adjusted).

DISCUSSION 2

• Shifting the mean F0 of the carrier tone down in frequency did not degrade performance, even at shifts of 75 Hz (V/T109 condition)

• Shifting the tone by 100 Hz (V/T84) resulted in significantly worse performance

• However, when adjusting for audibility, performance was equivalent, even at this very low frequency.

CONCLUSION

• There was no adverse effect of shifting F0 by as much as 100 Hz, when equating SL, suggesting that patients with extremely frequency-limited hearing may benefit from acoustic stimulation by a tone that tracks F0, even when that tone is shifted well below the "correct" F0.

• These results may have important implications for signal-processing strategies in individuals with restricted low-frequency hearing.

REFERENCES

Turner, C. W., Gantz, B. J., Vidal, C., Behrens, A., and Henry, B. A. (**2004**). "Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing," J. Acoust. Soc. Am. **115**(4), 1729-1735.

Qin, M. K. and Oxenham, A. J. (**2006**). "Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech," J. Acoust. Soc. Am. **119**(4), 2417-2426.

Qin, M. K. and Oxenham, A. J. (**2003**). "Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers," J. Acoust. Soc. Am. **114**(1), 446-454.

Work supported by NIDCD (R01-DC01376)