coustics'08
Paris
June 29-July 4, 2008
www.acoustics08-paris.org

Investigating the Influence of Pulse Rate and Duration on Pitch Perception in Cochlear Implants

Joshua Stohl, Chandra Throckmorton and Leslie Collins

Duke University, Department of Electrical and Computer Engineering, Box 90291, Durham, NC 27708, USA
lcollins@ee.duke.edu

Multi-rate sound processing strategies may potentially provide improved speech recognition and music perception in cochlear implants, although it may be necessary to tune these strategies using psychophysical data to provide maximum benefit to the user. To investigate the need for tuning, cochlear implant users participated in several rate-based psychophysical experiments. Subjects performed single- and two-rate pitch ranking tasks to investigate the influence of pulse rate on the overall pitch structure. The introduction of a second rate resulted in frequent pitch reversals between closely spaced electrodes. In addition to the multi-rate pitch structure, the duration necessary to perceive a rate change must also be considered. Experiments were performed using ABA stimuli, where rate B was higher than rate A. Subjects completed an embedded rate discrimination task in which the duration of B was fixed and its rate varied adaptively for durations ranging from 20-200 ms. The rate of segment B was then fixed, and its duration varied adaptively. Results imply that embedded rate difference limens may not be a function of duration, and that the minimum duration required for detecting a change in pulse rate is subject specific.

1 Introduction

While cochlear implant users are able to perform reasonably well on speech recognition tasks in quiet conditions, performance in noisy conditions is still significantly worse than that of normal hearing individuals. Speech recognition in tonal languages and musical tasks remain a challenge for implant users as well, and this is thought to be due to poor spectral representations of the subtle cues required for such tasks [1, 2]. A number of researchers have proposed the use of multiple stimulation rates in a speech processing strategy as a means of incorporating some of the fine structure information that is missing from current amplitude-modulation only algorithms [3, 4, 5, 6, 7, 8].

Multi-rate strategies attempt to transmit fundamental frequency (F0), within-channel dominant frequency, and fine timing information by using time-varying stimulation rates. The relative independence of place and rate in electric hearing [9] provides the opportunity to transmit spectral cues not only as a function of place of stimulation, as is typically done in commercial cochlear implant speech processors, but by varying the pulse rate at each location as well. Acoustic models of multi-rate strategies have demonstrated the potential benefit of this type of stimulation in noise [6, 7]; however, it is important to first understand the influence that using multiple stimulation rates will have on pitch perception.

Through the use of various psychophysical tasks, it may be possible to gain a better understanding of the rate pitch percept and what specific parameters must be considered in a multi-rate strategy to account for the effects of time-varying rate. Here, we will investigate the impact that using two stimulation rates has on the overall pitch structure, as well as examine the effects of duration on the rate-pitch percept. Subjects completed single and two-rate pitch ranking tasks, a short duration pitch ranking task, isolated and embedded rate discrimination tasks, and an embedded duration detection task.

2 Single-rate versus Two-rate Pitch Ranking

The tonotopic ordering of the cochlea is well established [10]; however, it is not clear what effect electrical stimulation with rectangular biphasic pulses at multiple rates may have on the well known place pitch structure. Previous studies have suggested that it may be possible to induce pitch reversals by varying the stimulation rate on multiple electrodes [11, 12]. Here, the pitch structure as a function of rate and place is examined via a series of pitch ranking tasks. Further details of this study can be found in [13].

2.1 Subjects

Five subjects participated in the single and two-rate pitch ranking tasks (S2, S4, S5, S6, S7). Demographic information for these subjects can be found in Table 1. All subjects were implanted with Cochlear Corporation's Nucleus CI24 cochlear implants at least two years prior to this study. Monopolar 1+2 (MP1+2) mode of stimulation was used for all subjects as this was also their clinical mode of stimulation. All subjects were compensated for their time except for Subject S7 who chose to volunteer his time. This study and the compensation exception made for Subject S7 were approved by the Institutional Review Board at Duke University.

Subject ID	Gender	Age	Duration of Deafness (yrs)	Mode of Stimulation
S2	F	71	20	MP1+2
S4	M	19	8	MP1+2
S5	F	58	28	MP1+2
S6	M	71	3	MP1+2
S7	M	53	<1	MP1+2
S8	M	55	17	MP2

Table 1: Demographic Information

2.2 Stimuli

All stimuli in this study were biphasic pulse trains with 25 μs pulse widths and an 8 μs interphase gap and were presented via direct stimulation with the SPEAR3 research sound processor [14]. In this pitch ranking task, subjects heard two 300 ms pulse trains separated by a 500 ms interstimulus interval and were asked to select the interval that contained the higher pitch. Two sets of stimuli were used. The first set of stimuli consisted of pulse trains presented at approximately 200 pulses per second (pps) presented to all active electrodes. This set of stimuli provided the pitch structure as a function of place of stimulation alone and was similar to the study done by [15]. The second set of stimuli included 300 ms pulse trains presented at approximately 200 and 400 pulses per second (pps) and were also presented to all active electrodes. Given the large number of stimuli in the second set (22 electrodes x 2 rates), this set was broken down into three subsets of stimuli that contained electrodes located in either the apical, middle, or basal region of the cochlea. Both sets (single-rate and two-rate) were presented seven to ten times for each subject, and all stimuli were loudness balanced prior to testing to avoid any loudness cues arising due to the difference in stimulation rate or location.

2.3 Results

Pitch ranking data were evaluated with the row sum analysis technique [16, 17], and results are plotted with stimulus on the abscissa and percent wins on the ordinate. Percent wins indicates the percentage of time that any stimulus was chosen as higher than all other stimuli in the set. In the cochlear convention, numbering begins at the basal end of the cochlea and increases toward the apical end. Representative results from Subject S2 are plotted Figure 1. The single rate results indicate that place pitch follows the tonotopic ordering of the cochlea, and these results are in agreement with previous studies [18, 19, 20]. Also in agreement with the rate pitch literature, 400 pps was consistently ranked higher than 200 pps on any given electrode for most subjects (e.g., [21]). However, the two rate results imply that frequent overlapping pitch percepts occur when comparing a basal electrode that was stimulated at approximately 200 pps to its apical neighboring electrode that was stimulated at approximately 400 pps. This phenomenon was seen regardless of location along the cochlea.

Although the introduction of a second rate seems to have little effect on the monotonic nature of place pitch (i.e., 400 pps across electrode is still monotonically decreasing from base to apex), the combined effects of place and rate pitch result in a pitch structure that is no longer a simple function of place on the cochlea. This has implications for multi-rate cochlear implant speech processing strategies that generally assume that no overlapping pitch percepts occur as a function of stimulation rate, but rather that the tonotopic ordering of the cochlea remains intact even under multi-rate stimulation. Subjects S4 and S7 displayed similar rate-place pitch structures to that of Subject S2; however, Subjects S5 and S6 demonstrated a much less patterned pitch structure as a function of both place, rate, and a combination of the two. This variation between subjects suggests that subject-specific tuning may provide some benefit to implant users with respect to a multi-rate strategy. Specifically, pitch ranking results similar to those obtained in this study could be used to map the output of a filterbank to the appropriate electrode-rate combination required to maintain a monotonic pitch percept.

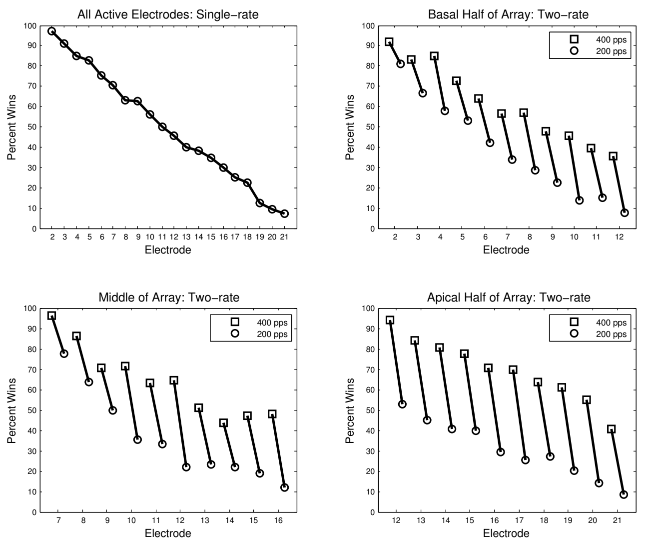

Figure 1: Single-rate and Two-rate pitch ranking results from Subject S2 demonstrate the frequent overlapping pitch percepts that occur when a second rate is introduced.

3 Place Pitch as Function of Duration

While the previous experiment demonstrates that place and rate may interact differently than is often assumed, that experiment, like other rate-based experiments, did not include temporal constraints that are present when using multiple rates in a sound processing strategy. Specifically, changes in rate will occur instantaneously (i.e., without an interstimulus interval), and in multi-rate strategies in general, rate is a time-varying parameter where the duration between changes may be very short. The following three experiments address the effects of duration on both the place and rate pitch percept.

The first duration experiment examines place pitch as a function of duration. A pitch ranking task was implemented in which subjects compared all active electrodes stimulated at 200 pps, and after comparing all stimuli, the duration was then reduced to determine if at some point the place pitch percept became unperceivable.

3.1 Subjects

Subjects S2, S5, S7, and S8 participated in a the following three psychophysical experiments, and demographic information for these subjects can be found in Table 1. All subjects were implanted with Cochlear Corporation's Nucleus CI24 cochlear implants at least two years prior to this study. Monopolar (MP1+2 or MP2) mode of stimulation was used for all subjects. All subjects were compensated for their time except for Subject S7 who chose to volunteer his time. This study and the compensation exception made for Subject S7 were approved by the Institutional Review Board at Duke University.

3.2 Stimuli

Subjects were presented two pulse trains on different electrodes, both at a rate of 200 pps, and asked to select the interval containing the higher pitch. The task was repeated for durations of 10, 20, 50, 100, and 200 ms in decreasing order. All stimuli within the set, or all active electrode pairs, were compared before reducing the duration. Sets were repeated 3-7 times depending on subject, and all durations were tested an equal number of times for any given subject. Stimuli were also loudness balanced to account for loudness differences between electrodes.

3.3 Results

Pitch ranking data was again analyzed using row sum analysis. Electrode number is listed along the horizontal axis, and the corresponding percent wins for each stimulus is given along the vertical axis. Figure 2 contains short duration pitch ranking results for all four subjects that participated in this experiment. Duration is indicated by line style, and these results indicate that the ability to rank electrodes according to place is not a function of duration for the given stimulation parameters.

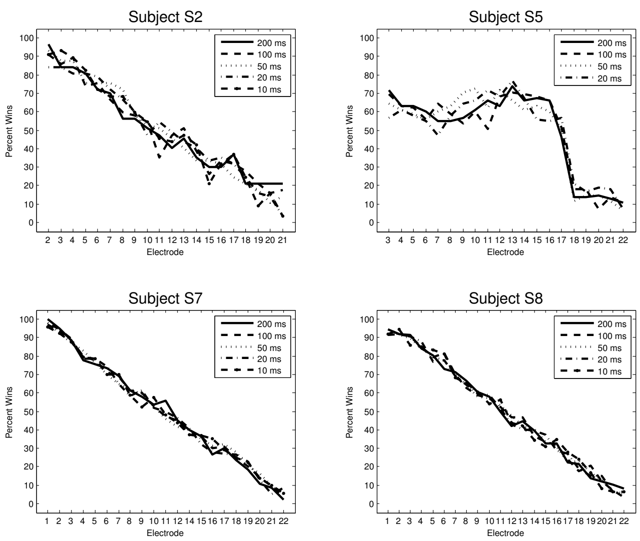

Figure 2: Short duration pitch ranking results indicate that place pitch is not generally affected by duration.

4 Isolated Versus Embedded Rate Discrimination

While the short duration pitch ranking results provide information about place pitch as a function of duration, the effects of duration on rate pitch are of even greater interest when implementing a multi-rate strategy. This experiment was designed to compare pulse rate difference limens (PRDL) in isolation (with an interstimulus interval) to PRDLs for embedded rate changes presented in an ABA pattern as discussed below.

4.1 Stimuli

This experiment is broken down into two separate tasks: Isolated rate discrimination and embedded rate discrimination. In both cases the base rate was fixed at 200 pps, and subjects performed a two interval, forced choice, adaptive procedure with flanking cues to determine the PRDL with respect to 200 pps. The target rate was always higher than the base rate and was typically initialized at 400 pps. The task was considered complete after twelve reversals or sixty trials. For the first four reversals a 1-up, 1-down rule was used, and a 1-up, 2-down rule was applied for the final eight reversals. The target rate was multiplied by a factor of 1.4 in the case of an incorrect response and divided by 1.4 in the case of a correct response or two correct responses after the fourth reversal. Three electrodes were selected for each subject to represent the basal, middle, and apical regions of the cochlea. All stimuli were loudness balanced at 200 and 400 pps, and loudness was also roved to prevent the subject from relying on consistent loudness cues that may arise as a function of pulse rate difference.

During the isolated rate discrimination task, subjects were presented constant rate, 200 ms pulse trains with a 500 ms interstimulus interval to a single electrode and asked to select the interval containing the different stimulus. During the embedded rate discrimination task, subjects were also asked to identify the interval that contained the different stimulus. In this part of the experiment, the target stimulus was of the ABA type, where A is the base rate (200 pps) and B is the adaptive, target rate ($>$ 200 pps). The three nontarget stimuli (the reference interval and the two flanking intervals) contained an AAA pattern of stimulation. The total stimulus duration was fixed at 600 ms, and stimuli were presented with a 500 ms interstimulus interval. The embedded rate discrimination task was repeated for four durations of B: 20, 50, 100, and 200 ms, and each DL measurement was repeated three to eight times.

4.2 Results

Isolated rate discrimination difference limens were first calculated for all four subjects at each of the three locations measured. Average isolated PRDLs were less than 200 pps (with respect to 200 pps) for all subjects except for Subject S5 who demonstrated elevated DLs, particularly at the apical end of her array. Embedded PRDLs were then calculated and normalized by the average isolated difference limen at each location of the cochlea. Combined isolated and embedded rate discrimination results are shown in Figure 3. Duration is plotted along the horizontal axis, and the mean of the normalized PRDLs are plotted in log scale on the vertical axis along with error bars indicating one standard error ($\sqrt{\sigma^2/n}$). Each location on the cochlea is plotted in a different subplot, and subjects are indicated by line style. These results seem to imply that embedded pulse rate difference limens are typically higher than isolated PRDLs. This suggests that while subjects may be able to discriminate closely spaced pulse rates when stimuli are separated by an interstimulus interval, these same stimulation rates may not be discriminable in a multi-rate strategy where rates are switching instantaneously. In general, embedded pulse rate difference limens do not appear to be a function of duration.

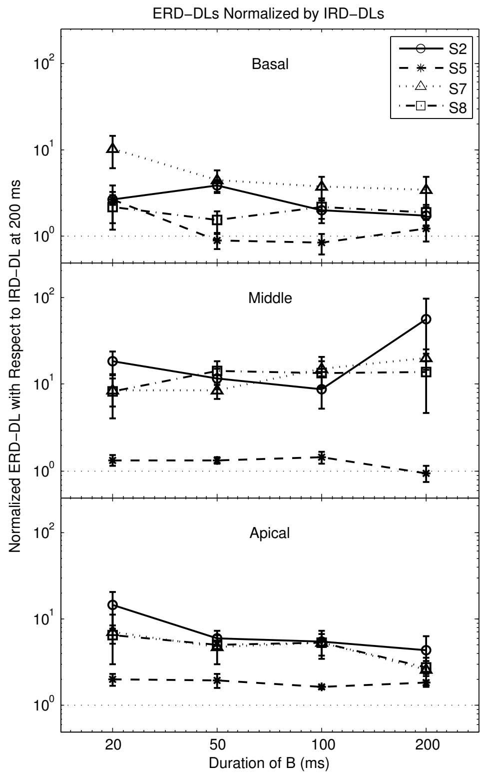

Figure 3: Normalized pulse rate difference limens show that embedded PRDLs are generally higher than isolated PRDLs.

5 Detection of an Embedded Rate Change as a Function of Duration

Although embedded rate discrimination results suggest that embedded PRDLs are not a function of duration when the duration is fixed, it may also be informative to investigate the response to a change in pulse rate while making duration the variable in question. This may be a more realistic model of the behavior of a multi-rate strategy, as spectral changes will not typically occur at predefined, fixed durations.

5.1 Stimuli

A two alternative, forced choice, adaptive procedure was again implemented for three locations of the cochlea. All stimuli had a total duration of 600 ms, and subjects were asked to identify the different interval. In this case, reference intervals contained AAA stimuli, and the target stimulus contained an ABA style stimulus. Here, the rate of B was fixed at a discriminable rate (determined from the previous experiment), and the duration of the middle section of each stimulus (A in the reference intervals and B in the target interval) was varied adaptively. Each experiment ended when either twelve reversals or sixty trials were reached. Adaptive durations followed a 1-down, 1-up rule for the first four reversals and a 2-down, 1-up rule for the final eight reversals. An incorrect reversal resulted in the duration of the middle section of the stimulus being multiplied by 1.4, and one or two correct responses, depending on the number of previous reversals, resulted in the duration of the middle section of the stimulus being divided by 1.4. The maximum duration of the middle section of the stimulus was limited to 200 ms. Loudness balancing was performed for all rates used in this task, and loudness was also roved to account for any additional loudness cues that may have arisen due to difference in pulse rate. Each minimum detectable duration was measured four to eight times.

5.2 Results

Figure 4 contains the results from the embedded duration detection task. Electrode position is indicated along the horizontal axis, and the minimum duration required to detect an embedded rate change is plotted along the vertical axis. Mean results are plotted for each subject as indicated by line style, and error bars indicate one standard error ($\sqrt{\sigma^2/n}$). A wide range of variability can be seen both within and across subject. In some cases these results do not agree with those seen in the previous embedded PRDL experiment, as subjects who were previously able to achieve reasonable pulse rate discrimination at 20 ms are unable to achieve those durations here.

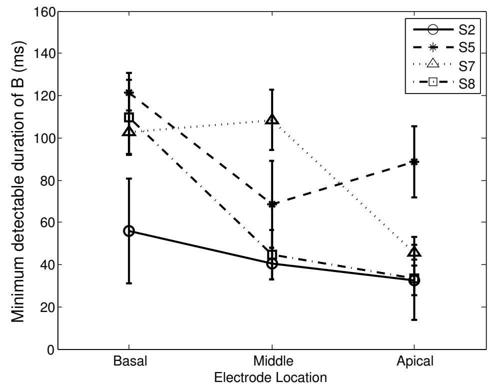

Figure 4: Minimum detectable duration results show large variability both within and across subjects.

6 Conclusion

Pitch ranking results indicate that frequent pitch overlaps occur when stimulating multiple electrodes at two rates [13]. Subject specific pitch maps that preserve a monotonic pitch structure may be beneficial to implant users. Embedded rate discriminate DLs suggest the relatively large changes in rate may be required for detection in a multi-rate strategy, and embedded duration detection results may imply that in some cases, it may not be possible to update rates often enough to transmit phonemic changes and other rapidly varying spectral cues to all subjects.

Together these results suggest that two main conclusions to be drawn from these psychophysical experiments: 1) Previous assumptions about the rate pitch percept as applied in multi-rate strategies may have been invalid. 2) The large amount of variability between subjects suggests that subject-specific tuning may be necessary for implant users to obtain maximum performance benefits from multi-rate strategies.

Acknowledgments

This work is supported by the National Institutes of Health, 1-R01-DC007994-01. The authors would also like to thank the subjects that participated in these studies for their time and dedication.

References

[1] Y.-Y. Kong and F.-G. Zeng, "Temporal and spectral cues in mandarin tone recognition", The Journal of the Acoustical Society of America **120**, 2830–2840 (2006).

[2] H. J. McDermott, "Music perception with cochlear implants: A review", Trends in Amplification **8**, 49–82 (2004).

[3] D. Grayden, A. Burkitt, O. Kenny, J. Clarey, A. Paolini, and G. Clark, "A cochlear implant speech processing strategy based on an auditory model", in *Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004. Proceedings of the 2004*, 491–496 (2004).

[4] R. J. M. van Hoesel and R. S. Tyler, "Speech perception, localization, and lateralization with bilateral cochlear implants", The Journal of the Acoustical Society of America **113**, 1617–1630 (2003).

[5] N. Lan, K. Nie, S. Gao, and F. Zeng, "A novel speech-processing strategy incorporating tonal information for cochlear implants", Ear and Hearing **51**, 752–760 (2004).

[6] K. Nie, G. Stickney, and F.-G. Zeng, "Encoding frequency modulation to improve cochlear implant performance in noise", IEEE Transactions on Biomedical Engineering **52**, 64–73 (2005).

[7] C. S. Throckmorton, M. S. Kucukoglu, J. J. Remus, and L. M. Collins, "Acoustic model investigation of a multiple carrier frequency algorithm for encoding fine frequency structure: Implications for cochlear implants", Hearing Research **218**, 30–42 (2006).

[8] C. Arnoldner, D. Riss, M. Brunner, M. Durisin, W.-D. Baumgartner, and J.-S. Hamzavi, "Speech and music perception with the new fine structure speech coding strategy: preliminary results", Acta Oto-Laryngologica **127**, 1298 (2007).

[9] C. M. McKay, H. J. McDermott, and R. P. Carlyon, "Place and temporal cues in pitch perception: are they truly independent?", Ear and Hearing **1**, 25–30 (2000).

[10] D. Greenwood, "Critical bandwidth and the frequency coordinates of the basilar membrane", The Journal of the Acoustical Society of America **33**, 1344–1356 (1961).

[11] D. Eddington, W. Dobelle, D. Brackmann, M. Mladejovsky, and J. Parkin, "Place and periodicity pitch by stimulation of multiple scala tympani electrodes in deaf volunteers", Transaction of the American Society for Artificial Internal Organs **24**, 1–5 (1978).

[12] H. J. McDermott and C. M. McKay, "Musical pitch perception with electrical stimulation of the cochlea", The Journal of the Acoustical Society of America **101**, 1622–1630 (1997).

[13] J. S. Stohl, C. S. Throckmorton, and L. M. Collins, "Assessing the pitch structure associated with multiple rates and places for cochlear implant users", The Journal of the Acoustical Society of America **123**, 1043–1053 (2008).

[14] J. S. Stohl, C. S. Throckmorton, and L. M. Collins, "Developing a flexible spear3-based psychophysical research platform for testing cochlear implant users", Technical Report, Duke University, Durham, NC, USA (2008).

[15] L. M. Collins, T. A. Zwolan, and G. H. Wakefiled, "Comparison of electrode discrimination, pitch ranking, and pitch scaling data in postlingually deafened adult cochlear implant subjects", The Journal of the Acoustical Society of America **101**, 440–455 (1997).

[16] H. David, *The Method of Paired Comparisons* (Oxford University Press, 200 Madison Avenue, New York, NY 10016) (1988).

[17] M. Kendall and B. B. Smith, "On the method of paired comparisons", Ear and Hearing **31**, 324–345 (1940).

[18] L. M. Collins and C. S. Throckmorton, "Investigating perceptual features of electrode stimulation via a multidimensional scaling pradigm", The Journal of the Acoustical Society of America **108**, 2353–2365 (2000).

[19] D. A. Nelson, D. J. V. Tasell, A. C. Schroder, S. Soli, and S. Levine, "Electrode ranking of "place pitch" and speech recognition in electrical hearing", The Journal of the Acoustical Society of America **98**, 1987–1999 (1995).

[20] B. Townshend, N. Cotter, D. V. Compernolle, and R. White, "Pitch perception by cochlear implant subjects", The Journal of the Acoustical Society of America **82**, 106–114 (1987).

[21] F.-G. Zeng, "Temporal pitch in electric hearing", Hearing Research **174**, 101–106 (2002).