Investigating the effects of stimulus duration and context on pitch perception by cochlear implant users

Joshua S. Stohl, Chandra S. Throckmorton, and Leslie M. Collins^{a)}

Department of Electrical and Computer Engineering, Duke University, 129 Hudson Hall, Box 90291, Durham, North Carolina 27708-0291

(Received 28 June 2008; revised 21 April 2009; accepted 22 April 2009)

Cochlear implant sound processing strategies that use time-varying pulse rates to transmit fine structure information are one proposed method for improving the spectral representation of a sound with the eventual goal of improving speech recognition in noisy conditions, speech recognition in tonal languages, and music identification and appreciation. However, many of the perceptual phenomena associated with time-varying rates are not well understood. In this study, the effects of stimulus duration on both the place and rate-pitch percepts were investigated via psychophysical experiments. Four Nucleus CI24 cochlear implant users participated in these experiments, which included a short-duration pitch ranking task and three adaptive pulse rate discrimination tasks. When duration was fixed from trial-to-trial and rate was varied adaptively, results suggested that both the place-pitch and rate-pitch percepts may be independent of duration for durations above 10 and 20 ms, respectively. When duration was varied and pulse rates were fixed, performance was highly variable within and across subjects. Implications for multi-rate sound processing strategies are discussed. © 2009 Acoustical Society of America. [DOI: 10.1121/1.3133246]

PACS number(s): 43.66.Ts, 43.66.Hg, 43.66.Fe [BCM]

Pages: 318–326

I. INTRODUCTION

Individuals with cochlear implants (CIs) often exhibit speech recognition performance in quiet conditions that is comparable to that of normal-hearing individuals; however, there continue to be situations in which discrepancies in performance between implant users and normal-hearing individuals exist. For example, speech recognition in situations where there is environmental noise, or there are multiple talkers, remain a challenge for implant users. In addition, the subtle pitch cues that are used to communicate in tonal languages are not accurately conveyed via current implants (Wei *et al.*, 2004; Kong and Zeng, 2006), and the lack of fine spectral cues also renders music identification and appreciation a challenge (Gfeller *et al.*, 2007; McDermott, 2004).

The majority of clinical speech processing algorithms divide the acoustic input obtained from the microphone of the CI into M subbands that correspond to M electrodes, extract the slowly-varying temporal envelope within each band, and use those envelopes to modulate fixed-rate biphasic pulse trains that are presented on N of the M implanted electrodes, $N \leq M$, (i.e., Wilson *et al.*, 1991; Kiefer *et al.*, 2001; Koch *et al.*, 2004). While many subjects are able to use envelope cues to achieve high speech recognition scores in quiet conditions, it has been hypothesized that representing the fine structure (FS) of the input signal may be necessary to obtain further improvements in CI user performance on more challenging tasks such as those mentioned above (Wilson *et al.*, 2004). The envelope and FS may be defined in the following way: A signal may be represented analytically as $s(t) = A(t)\cos\phi(t)$, where $A(t)$ may be referred to as

the envelope, and $\cos\phi(t)$ may be referred to as the FS, with $\phi(t) = \arctan(x_i(t)/x_r(t))$, and $x_i(t)$ being the Hilbert transform of the original signal, $x_r(t)$. It is possible that some of the FS information is contained in the envelope of the signal if this information is below the cutoff frequency used during envelope extraction, and the stimulation rate is high enough to adequately represent that information. However, the majority of clinically available speech processing algorithms do not explicitly encode FS (Wilson *et al.*, 2004; Zeng, 2004).

One proposed method of increasing the amount of FS information that is available to CI users via a CI speech processing strategy is to encode the FS by using time-varying stimulation rates (Grayden *et al.*, 2004; Büchner *et al.*, 2005; van Hoesel and Tyler, 2003; Lan *et al.*, 2004; Nie *et al.*, 2005; Throckmorton *et al.*, 2006; Arnoldner *et al.*, 2007). Unlike in acoustic hearing, the place- and rate-pitch mechanisms may function independently in electric hearing (Tong *et al.*, 1983; McKay *et al.*, 2000), and these proposed multi-rate strategies use the independence of place and rate in an attempt to improve upon existing algorithms that transmit spectral cues only as a function of electrode placement in the cochlea. By conveying fundamental frequency (F0) (Lan *et al.*, 2004; Büchner *et al.*, 2005), the dominant frequencies in a bank of subbands (Nie *et al.*, 2005; Throckmorton *et al.*, 2006), or fine timing cues (van Hoesel and Tyler, 2003; Grayden *et al.*, 2004; Arnoldner *et al.*, 2007) with time-varying stimulation rates, researchers hope to provide additional information about the spectrum of the acoustic signal within the constraints of existing implant hardware.

The implementation of a CI strategy that incorporates time-varying stimulation rate requires practical knowledge regarding the perception of such stimuli. Specifically, the pulse rate difference and the minimum duration of any single rate stimulus required for distinct percepts across rates may

^{a)}Author to whom correspondence should be addressed. Electronic mail: lcollins@ee.duke.edu.

TABLE I. Demographic information for implanted subjects.

Subject ID	Gender	Age (year)	Age at onset of deafness (years)	Age at implantation (years)	Mode of stimulation	Speech recognition (percent correct)
S2	F	72	46	66	MP1 + 2	92
S5	F	59	26	54	MP1 + 2	90
S7	M	53	50	50	MP1 + 2	97
S8	M	55	35	52	MP2	100

be incorporated into a multi-rate strategy. Historically, researchers have investigated the first issue, the ability of CI users to discriminate one pulse rate from another, when each pulse rate is presented in isolation, i.e., with stimuli separated by an inter-stimulus interval (Fearn *et al.*, 1999; McDermott and McKay, 1997; Pijl, 1995; Shannon, 1983; Tong and Clark, 1985; Townshend *et al.*, 1987; Zeng, 2002). These studies demonstrated that this percept saturates somewhere between 300–500 pps (pulses per second) for most implant users. However, proposed multi-rate sound processing strategies typically incorporate time-varying signals in which pulse rate may be changing rapidly and without the relatively long interstimulus gaps that are used in most rate discrimination studies (approximately 200–500 ms).

Rate-based psychophysical experiments in CI subjects have traditionally focused on frequency modulation (FM) detection rather than the investigation of perceptual responses to instantaneous changes in stimulation rate (Chen and Zeng, 2004; Tong *et al.*, 1982; Luo and Fu, 2007). FM difference limens (FMDL) increased with increasing base rate (Chen and Zeng, 2004), and when coupled with amplitude modulation (AM), FMDL increased with an increase in AM depth. This finding is relevant to multi-rate CI strategies in which AM is used to convey the slowly varying envelope information within each band, and simultaneous changes in stimulation rate are used to transmit FS information. Expanding upon previous studies, the experiments presented below were designed to measure pulse rate difference limen (PRDL) for stimuli containing relatively large (i.e., 200 pps), instantaneous changes in rate. These data are applicable to multi-rate strategies that propose quantizing the variations in the FS and mapping those changes on to predefined stimulation rates (e.g., Throckmorton *et al.*, 2006).

In addition to investigating the rate difference required for listeners to detect an instantaneous change in rate, the influence of duration on the rate-pitch percept is of interest for implementation of a multi-rate strategy. Stimulation rate may be updated as often as on a pulse-by-pulse basis, yet there remains some question about whether or not rapid changes in stimulation rate are detectable by CI listeners. Chen and Zeng (2004) observed a significant difference between FMDLs at low sinusoidal FM rates (<80 Hz) and higher FM rates (160 and 320 Hz), suggesting that some threshold may exist at which the ability to identify rapid changes in rate increases in difficulty for CI listeners. Tong *et al.* (1982) investigated frequency discrimination as a function of duration using linear FM sweeps. Decreasing the duration of the FM sweep resulted in a decreased ability to

detect the difference between a stimulus with a linear sweep and a stimulus with a constant pulse rate for the single subject tested. These FM studies suggest that both the range of pulse rates and the duration of each rate may affect the ability of strategies that rely on time-varying stimulation rate to convey FS cues. In this study, direct measurement of these two variables was conducted using quantized changes in rate.

The study consisted of three experiments. First, the consistency of the place-pitch percept was measured as a function of duration using a pitch ranking task. The ability to consistently pitch rank electrodes independent of duration would imply that salient percepts remained when presenting single-rate stimuli at short durations. The subsequent experiments were designed to investigate the pulse rate separation and single-rate duration required for distinct percepts on one electrode. PRDLs were measured with an adaptive procedure for stimuli both in isolation and when embedded (i.e., with no interstimulus interval between stimuli with different pulse rates). Repeating the experiment in both conditions allowed for the comparison of PRDLs measured with rapid transitions in pulse rate to traditional PRDLs. This also provided a direct measurement of the ability to detect rate changes as they would be presented in a multi-rate sound processing strategy. In the third experiment, embedded PRDLs were used to set pulse rates, and an adaptive procedure was implemented to determine the minimum duration required for detection of a change between two fixed pulse rates with no interstimulus interval between them. The goal of this experiment was to determine how often rate may be changed while still evoking a perceptual change. These experiments may provide insight regarding the ways in which time-varying pulse rates may best be implemented in a CI speech processing algorithm to provide maximum benefit to the individual user.

II. METHODS

A. Subjects and stimuli

Four postlingually deafened subjects participated in the experiments described below. Demographic information for the subjects is presented in Table I. All subjects were implanted with a version of Cochlear Corporation's CI24 implant and had a minimum of 3 years experience with their device prior to testing. Testing occurred over six to nine sessions, and each session lasted between 2 and 4 h. All subjects were paid for their time except for subject S7 who elected to volunteer his time. These experiments were ap-

proved by the Institutional Review Board at Duke University, as was subject S7's voluntary participation.

Pulse trains consisting of biphasic rectangular pulses with 25 μ s pulse widths and an 8 μ s interphase gap were presented via the SPEAR3 research sound processor (Stohl *et al.*, 2008). Only electrodes available in the subjects' clinical MAP were used for testing, and threshold and comfort levels were measured at the beginning of each test session. While all four subjects use a monopolar 1+2 (MP1+2) mode of stimulation in their clinical devices, only subjects S2, S5, and S7 were stimulated using an MP1+2 mode in this study. Subject S8 was stimulated in Monopolar 2 mode due to safety concerns regarding the use of the Freedom Implant (CI24RE) and MP1+2 stimuli with the SPEAR3.¹

B. Experiment 1

In general, electrical stimulation via a multichannel CI elicits percepts that follow the tonotopic arrangement of the cochlea (Eddington *et al.*, 1978; Townshend *et al.*, 1987). The first experiment was designed to determine what effects, if any, duration had on pitch ranking due to place of stimulation alone.

1. Stimuli and experimental task

A two-interval, forced-choice, pitch ranking task was used. In each trial, subjects were presented with two intervals and instructed to pick the interval containing the sound with the higher pitch (Townshend *et al.*, 1987). All pulse trains were presented at a rate of 200 pps throughout this experiment. Pulse train durations were 10, 20, 50, 100, and 200 ms, and the interstimulus interval was 300 ms. An experimental block consisted of one comparison of each active electrode to all other active electrodes at a single duration, and a set included the presentation of one block at all of the durations listed above. Electrode pairs were chosen randomly within each block and were presented in a random order during each trial. Blocks were presented from the longest duration to the shortest, and stimuli were presented at all durations before the set of durations was repeated. Each set was repeated three to seven times.

All stimuli were loudness balanced across electrode prior to testing, and subjects S5 and S7 required increased stimulation levels for stimuli with shorter durations. Estimates of equal loudness were obtained via the method of adjustment. Two pulse trains were presented repeatedly in an alternating fashion. The first, or reference, stimulus was a 200 pps pulse train presented on the most apical active electrode at a comfortable level determined by listener before the task, and the second, or target, stimulus was a pulse train presented on the adjacent active electrode in the basal direction. The subject was instructed to use a knob that was connected to the test personal computer (PC) to adjust the volume (current level) of the target stimulus. The initial amplitude of the target stimulus was between 5 and 10 current steps below the amplitude of the first stimulus depending on the dynamic range for that electrode. A repeat measurement was made with the initial amplitude of the target stimulus set above the reference stimulus amplitude but be-

low the maximum comfortable loudness (MCL) for the electrode being tested. If the two estimates of equal loudness were separated by more than four current steps, this procedure was repeated. The resulting amplitudes were arithmetically averaged to determine the level at which the two stimuli were considered to be equally loud. The target electrode became the reference electrode, and the next active electrode in the basal direction became the target electrode for the next trial. All electrodes were loudness balanced to the adjacent apical electrode using this procedure (Throckmorton and Collins, 2001).

2. Results

Pitch ranking data were analyzed using row sum analysis (Collins *et al.*, 1997; David, 1988), and the results are presented in Fig. 1. The horizontal axis shows electrode number, where electrode 1 is the most basal electrode in the array and numbering increases toward the apex. The vertical axis shows each electrode's percent wins. This value is the percentage of time that a given electrode was ranked as higher in pitch than all other active electrodes. Each duration is indicated by a unique line style and is also accompanied by a coefficient of consistence in the legend labeled ζ . The coefficient of consistence, $\zeta \in [0, 1]$, is unity when no inconsistencies occur in the ranking of the stimuli throughout the experiment (Kendall and Smith, 1940).

As can be seen in Fig. 1, pitch ranking as a function of place generally remained consistent when pulse rate was fixed, regardless of the stimulus duration. Furthermore, subjects were able to rank electrodes with the same consistency across durations, as is indicated by stable values of ζ for all durations. These data suggest that a pitch percept is available at durations as short as 10 ms, and that this percept remains constant with respect to all other electrodes in the array.

C. Experiment 2

One goal of this experiment was to determine PRDL as a function of duration. While PRDLs are typically measured with an interval between two stimuli with different pulse rates (e.g., Zeng, 2002), this does not accurately reflect the time-varying stimulation rate that would be observed on any single electrode for a multi-rate strategy. Therefore, this experiment was designed to test CI subjects' ability to discriminate pulse rates that change instantaneously, i.e., without an interstimulus interval. Another goal of this experiment was to determine what impact changing rate instantaneously may have on PRDLs, relative to those measured in the traditional fashion. Therefore, PRDLs were measured for both the traditional (isolated) and embedded patterns of stimulation.

1. Stimuli and experimental task

Two rate discrimination tasks were implemented. Both tasks used four-interval, two-alternative, forced-choice, adaptive procedures (Levitt, 1971) in which the subject was instructed to identify the interval that sounded "different." Intervals 2 and 3 contained possible targets, and intervals 1 and 4 were always fixed as reference intervals. Subjects selected the interval that they perceived to be different. In the

FIG. 1. Row sum analysis results of pitch ranking data are plotted for four CI subjects. The horizontal axis indicates electrode, and the vertical axis indicates the corresponding percent wins. Pulse train duration is indicated by line style, and the mean coefficient of consistency (ζ) for each duration is listed in the legend.

first task, subjects heard 200 ms fixed-rate pulse trains. The base or reference rate was always 200 pps, and the target rate was higher and changed from trial-to-trial. This task will be henceforth referred to as isolated rate discrimination (IRD), given that the stimuli associated with different pulse rates were always isolated from one another by a 500 ms inter-stimulus interval. During the second task, the target interval was of an ABA format in which the rate of A was always the reference rate of 200 pps, and the rate of B varied from trial-to-trial. Reference intervals were always presented at a constant 200 pps (AAA). Stimuli were approximately 600 ms in duration, and this task was repeated using four durations for rate B (20, 50, 100, and 200 ms). For each trial, the number of rate B pulses was calculated such that the error between the actual duration and the desired duration was minimized, where duration refers to the time between the onset of the first pulse and end of the period of the final pulse (i.e., 4 pulses at 200 pps has a duration of 20 ms). The number of rate A pulses was then calculated so that the total duration of the target interval was as close to 600 ms as possible, and rate B was always embedded in the middle of the stimulus. The second task will be henceforth referred to as embedded rate discrimination (ERD), as the rate B pulse train may be thought of as embedded in a pulse train of rate A. Figure 2(a) illustrates an example target stimulus with an embedded rate change.

Stimuli were loudness balanced via the method of adjustment prior to testing, and amplitudes were roved an average of ± 4 current steps² to minimize loudness cues due to differences in stimulation rate (McKay *et al.*, 2001). Each interval was roved in the IRD task, whereas in the ERD task, each third of each 600 ms interval was roved independently. The durations of each of the three A sections of the reference intervals were equal to the desired target durations (e.g., 275-50-275 ms) to minimize any cues arising from recogni-

tion of any residual loudness differences at the rate transitions in the target stimulus that remained after loudness balancing the stimuli across pulse rate. The stimulation levels of each third of the reference and target stimuli were roved to prevent changes in loudness in the target interval from providing an additional cue to the subjects.

Each task was considered complete after 12 reversals or 60 trials, whichever occurred first. For the first four reversals, a one-down, one-up rule was used, and a two-down, one-up rule was applied for the final eight reversals. The step size was a factor of 1.4, and the geometric mean of the DL at the final eight reversals was taken as the DL that corresponds to 70.7% probability of a correct decision (Levitt, 1971). For the cases in which 60 trials preceded 12 reversals, the geometric mean of the DL at the last N reversals was computed and taken as the DL, where N was even and $4 < N < 12$. Sixty trials were reached in less than 3% of all repetitions and were never reached for subject S5.

Difference limens (DLs) were measured for three electrode locations for each subject, one basal, one middle, and

FIG. 2. Examples of the stimuli used in the ERD task and the MDD task, respectively. During the ERD task, the rate of section B is varied while the duration of section B is held constant. During the MDD task, the rate of section B is fixed, and its duration is varied adaptively.

FIG. 3. Isolated and embedded rate DLs for all four subjects as a function of electrode and the duration of B in the ABA target stimulus. ERDDL is marked by a gray background. The median DL, upper and lower quartiles, and 1.5 times the interquartile difference are plotted for each condition. Outliers are indicated by a plus (+) symbol.

one apical. IRDDLs were measured between four and eight times with an average of approximately seven measurements per subject, and ERDDLs were measured between four and ten times with an average of approximately seven measurements per subject. DLs were measured for all three electrodes in a random order for a given task before repeat measurements were taken. The tasks were interleaved randomly as well.

2. Results

IRDDL and ERDDL are plotted as a function of electrode in Fig. 3, with ERDDL further subdivided by the duration of B and indicated by a gray background. The median DL, upper and lower quartiles, 1.5 times the interquartile range, and outliers, as indicated by a plus symbol, are plotted for each DL (McGill *et al.*, 1978). Data for all four subjects are shown, and subjects are grouped by panel. Nonparametric statistics were used to analyze these data. The use of nonparametric statistical analysis eliminates the need to assume that observations of the DL are drawn from some specific underlying distribution and are appropriate for analyzing these data given the relatively small sample size (four to eight observations per DL) (Moses, 1952).

A Kruskal–Wallis test, an extension of the two-class Wilcoxon rank sum test and a nonparametric one-way analysis of variance that tests for equal medians across k classes, $k > 2$, was performed for each subject and electrode with duration as factor (Gibbons, 1985). Due to multiple comparisons being made with the same data, Bonferroni adjustment was applied, and in general ERDDLs were not significantly affected by duration ($p < 0.05$). Two exceptions were found. The ERDDL for S5's electrode 4, when the duration of B was 20 ms, was significantly higher than the ERDDL measured when B was 50 and 100 ms ($p < 0.02$). A significant

difference was also found between the ERDDL for S8's electrode 20 when measured at a duration of B equal to 200 ms and compared to measurements taken when the duration of B was equal to 20, 50, and 100 ms ($p < 0.02$).

Given that there was generally no significant difference in ERDDL as a function of duration, ERDDLs were collapsed and the same analysis was applied to ERDDL within subject with electrode as factor. ERDDL and IRDDL are plotted for each subject and the group in Fig. 4. The group ERDDL includes all measurements, whereas those ERDDLs that were determined to be significantly different as a function of electrode were removed in the individual data shown.

FIG. 4. Isolated and embedded rate DLs for all four subjects and the group. ERDDLs shown in Fig. 3 have been combined for each subject and electrode due to the lack of significant effect of duration. DL medians, upper and lower quartiles, and whiskers indicating 1.5 times the interquartile difference are shown. Outliers are indicated by a plus (+) symbol.

TABLE II. Stimulation rate of B in ABA stimulus, amplitude used for reference rate A (200 pps) and amplitude used for B that was adjusted for equal loudness with respect to A for used in Experiment 3. All amplitudes are listed in current steps (0–255) by subject and electrode.

Subject		S2				S5			S7			S8	
Electrode	3	11	20	4	13	21	3	11	20	3	12	20	
Rate of B	400	400	400	400	500	700	500	500	500	600	400	400	
Amplitude of A	177	196	193	163	175	176	199	202	198	151	149	146	
Amplitude of B	176	190	193	161	172	172	198	200	197	150	145	145	

A significant effect of electrode was observed in two cases. ERDDLs on S5's electrode 21 and S8's electrode 3 were significantly higher than the other two electrodes tested for each subject ($p < 0.001$), and both ERDDLs were also significantly higher than the IRDDL for the same electrode ($p < 0.001$). When removing these two electrodes and comparing IRDDL and ERDDL within subject via the Wilcoxon rank sum test that checks for equal medians between two groups, ERDDL was significantly higher than IRDDL for S4, S7 and S8 ($p < 0.002$). S5 suffers from tinnitus and thus had a competing percept that made it difficult to perform tasks involving pitch. This is reflected in elevated IRDDLs, and the resulting lack of significant difference between IRDDLs and ERDDLs. Nonetheless, group ERDDLs were observed to be higher than group IRDDLs ($p < 0.001$) when pooling data from all four subjects. The median IRDDL for the group for all electrodes tested was found to be approximately 17.5 pps, and the median ERDDL for the group across electrode and duration was found to be approximately 63.8 pps. These DLs expressed as Weber fractions with respect to 200 pps are 0.09 and 0.31 for IRDDL and ERDDL, respectively. The significant increase in DL suggests that DLs should be measured in the embedded condition when determining discriminable rates for use in a multi-rate sound processing strategy.

D. Experiment 3

The lack of effect of duration on rate discrimination in the embedded case has implications for implementation of a multi-rate CI strategy. However, changing rate over a fixed duration does not necessarily reflect what may happen in a class of multi-rate strategies in which the stimulation rates are predefined (Throckmorton *et al.*, 2006) and the selected rate of stimulation is a time-varying function of the location of spectral peaks in the input signal. Furthermore, there is evidence in the normal-hearing literature that suggests that uncertain time of arrival and uncertain duration may both have a negative impact on signal detection ability (Wright and Fitzgerald, 2004; Wright, 2005). Therefore, in the third experiment the duration of the embedded stimulus was varied adaptively and rate of stimulation was fixed. Results from this experiment may provide insight into the minimum duration required for listeners to detect a change in stimulation rate when pulse rates are fixed. An example of the target stimulus used in the minimum detectable duration (MDD) task can be seen in Fig. 2(b). This is in contrast to the second experiment in which duration was fixed and rate varied adaptively [see Fig. 2(a) for comparison].

1. Stimuli and experimental task

The stimuli used in Experiment 3 were identical to those used in Experiment 2, except that the duration of the middle third of each stimulus was varied adaptively on a trial-by-trial basis instead of the stimulation rate. The stimulation rates of A and B were fixed, and the duration of the first and third sections of the stimuli were adjusted to maintain a total interval duration of approximately 600 ms. Each interval was separated by a 500 ms interstimulus interval. The reference rate was fixed at 200 pps for all subjects; however, the rate of B was selected using the ERDDL data obtained in Experiment 2 and was always at least 400 pps. Stimuli of different pulse rates were loudness balanced prior to testing using the same method as Experiment 2, and amplitudes were roved by an average of 4 current steps above and below the loudness-balanced levels. Subject specific stimulation rates that were used for the middle section of the target stimulus (B of ABA) are listed in Table II along with the corresponding current levels required for equal loudness.

The same procedure that was used in Experiment 2 was also used in Experiment 3. Twelve reversals were reached in all cases but one, in which S5 reached 60 trials on her first attempt at the task. In that case, eleven reversals occurred, and the geometric mean of the MDD at the last six reversals was taken as the MDD for that experimental run. Four to eight MDD measurements were taken using each electrode for each subject. Measurements were taken using all three electrodes of interest in a random order before repeating the task.

2. Results

The MDD for a fixed difference in rate is plotted in Fig. 5 as a function of electrode location for each subject as well as across electrode for the group, with each subject and the group in different panels. As before, each MDD is represented by a horizontal line at the median, lines at the upper and lower quartiles, and whiskers that indicate values within 1.5 times the interquartile range. MDDs outside of the whiskers are marked with a plus symbol and are considered outliers.

As is shown in Fig. 5, there is substantial variability of MDD within and across electrode for each subject and across subject as well. When applying a Kruskal–Wallis test with Bonferonni adjustment, no significant difference was observed between MDD when the factor was electrode for subjects S2 and S5; however, the MDD for S7's electrode 20 with the rate of A fixed at 200 pps and the rate of B fixed at 500 pps was significantly lower than the other two electrodes tested ($p < 0.05$). Similarly, the MDD for S8's electrode 20

FIG. 5. The minimum duration required to detect an embedded rate change is plotted as a function of electrode location for each subject and for the group irrespective of electrode. Data for each subject are contained in a separate panel. Associated with each electrode for each subject as well as for the group is the MDD median, upper and lower quartiles, whiskers extending to those points inside of 1.5 times the interquartile range, and outliers indicated with a plus (+) symbol.

was significantly lower than the MDD for electrode 3 with the rates of B equal to 400 and 600 pps, respectively ($p < 0.05$).

This variability suggests that the task was extremely challenging for some subject and electrode combinations. Similar inter-electrode trends and intra-electrode variability can be seen when comparing the ERDDL and MDD data. The presence of MDD values of 100 ms and higher suggests that in some cases, sound processing algorithms may need to limit how often stimulation rate changes in order to be effective.

III. DISCUSSION

The results from Experiment 1 suggest that place-pitch is not a function of duration in electric hearing for durations greater than 10 ms. Tong *et al.* (1982) found that the ability to discriminate stimuli containing shifts in electrode position from stimuli with static electrode position was independent of duration for one subject. Tong *et al.* (1982) measured the ability to discriminate place as a function of duration using a same-different task, and the pitch ranking results from Experiment 1 expanded upon that finding by using a larger number of electrodes (22 versus 4) and more subjects. Tong *et al.* (1982) concluded that the ability to detect rapid transitions in electrode position suggested that place of stimulation may be effectively used to transmit rapidly varying (segmental) speech information, and this hypothesis is supported by the pitch ranking data collected in the first experiment.

Comparing the IRDDL and ERDDL data collected in Experiment 2 indicated a significant increase in PRDL in the embedded case. This increase in PRDL may be due in part to the fact that the stimuli used in the ERD task were more complicated as they had longer durations and a higher number of level changes (12 total changes in level in a single trial as opposed to only 4 changes in level in a single trial of the

IRD task). The ERD task may have therefore required a greater amount of cognitive effort on the part of the listener. Nonetheless, the stimuli used to collect ERD data better mimic the stimulation patterns that may be used in a multi-rate strategy, and thus when determining the appropriate rates for use in such a strategy, it may be more appropriate to measure the DLs without an interstimulus interval between rate changes. Results from Experiment 2 also imply that changes in stimulation rate are detectable for durations as brief as approximately 20 ms, and that the ability to detect an embedded change in rate is not significantly impacted by duration when the duration of the embedded stimulus is fixed across observations and greater than 20 ms. These results, in combination with those data obtained in Experiment 1, suggest that in electric hearing the ability to detect changes in both the place and rate of stimulation are independent of duration above some minimum duration. Chen and Zeng (2004) observed a similar lack of significant effect of duration, or FM rate, when measuring sinusoidal FM detection at low FM rates (≤ 80 Hz); however, there was a significant increase in FMDLs at high sinusoidal FM rates (160 and 320 Hz) when compared to low FM rates. The lack of effect of duration on FM detection at longer durations, or lower FM rates, is similar to the trend observed in the data collected in Experiment 2 in which changes in rate were quantized.

In Experiment 3, MDD was measured with one pair of pulse rates for each electrode per subject. Results showed considerably more variability both within and across subject than the ERDDL results from Experiment 2, and no clear lower bound on the ability to detect instantaneous changes in rate was observed in general. In addition to measuring the detectability of shifts in electrode position, Tong *et al.* (1982) measured the ability of one subject to detect linear changes in pulse rate on single electrodes, and found an inability to perform FM sweep detection when the duration of the sweep was less than 25 ms. Note that in that study, the onset of the sweep was consistent across trial, as was the case in Experiment 2 of this study. Although inter-electrode trends were similar when comparing the ERDDL and MDD data from this study, data from Experiment 3 suggest that a time-varying duration that changes from trial-to-trial or observation-to-observation may hinder performance of embedded pulse rate change detection. It has been suggested in the normal-hearing literature that when performing a signal detection task listeners form a template for the target signal that includes both frequency and duration information (Wright, 2005), and it has also been shown that signal detection performance worsens when the onset time of the target stimulus is unknown (Wright and Fitzgerald, 2004). In Experiment 2, only the frequency, or pulse rate, of B changed from trial-to-trial, while the duration and onset time of B remained relatively constant. In Experiment 3, both the duration and the onset of time of B changed from trial-to-trial while the pulse rates for A and B remained constant, and it is possible that changing these two features (duration and onset time) between each trial resulted in a much more challenging task for the subjects that participated in this study.

Tong *et al.* (1982) concluded that while pulse rate cues may be appropriate for transmitting longer duration (supra-

segmental) information contained in the fundamental frequency of speech, the inability to perform FM sweep detection when the duration of the sweep was less than 25 ms may prevent these cues from being used to transmit segmental information. This supposition is supported by the variability in the data collected in Experiment 3 and that many of the MDDs obtained in Experiment 3 were greater than the average duration of an English phoneme (Umeda, 1975, 1977). Thus, using rate to transmit information about changes in a speech spectrum may not be effective for all subjects. In the event that rate update durations are restricted to the values determined by Experiment 3, spectral smearing across phonemes is likely to occur. However, the ability of some subjects to identify changes in rate at durations well below the median MDD (i.e., Subject S8, electrode 20) suggests that further investigation may be necessary to identify a true lower bound on rate change detection. As can be observed by comparing ERDDL from Experiment 2 summarized as the rate of B for each electrode in Table II, the probability of accurately detecting rate B with respect to rate A (200 pps) may have varied according to subject and electrode. Repeating Experiment 3 with multiple pulse rates assigned to B in an ABA stimulus and varying the duration of B (and thus A) adaptively would potentially provide a better representation of the interaction between duration and CI listeners' ability to detect changes in pulse rate.

Given that the ultimate goal of using multiple stimulation rates is to transmit FS cues in a CI speech processing algorithm, it may be most beneficial to determine sets of discriminable rates that allow pulse rate to be updated as often as possible, thus potentially transmitting the most information per unit time to the listener. Furthermore, a multi-rate CI speech processing algorithm will necessarily include AM in addition to time-varying rates, and extending the work of Luo and Fu (2007) to include changes in pulse rate that are quantized instead of slowly varying may provide further insight into the interaction of pulse rate, duration, and amplitude change.

IV. CONCLUSION

Improvement in speech recognition with multi-rate strategies may require subject specific tuning as supported by the variability in the data collected in this study. Specifically, ERDDL and MDD both varied considerably within and across subject, and parameters such as rate of stimulation and minimum duration for a single pulse rate may need to be adjusted according to subject-specific psychophysical data for users to take advantage of the additional information encoded via pulse rate. Based on the results from Experiment 2, PRDLs measured in the embedded case are more likely to reflect detectable pulse rates in a multi-rate strategy, but the duration at which they are measured may not be a significant factor above approximately 20 ms. Results from Experiment 3 suggest that update durations that provide discriminable changes in rate are both subject and electrode position specific. These differences have the potential to limit a multi-rate strategy's ability to transmit rapidly varying spectral information. Even if only for suprasegmental aspects of

speech, the use of multi-rate strategies may still provide additional, usable information that when deliberately included may aid in speech recognition in noisy conditions. Inconsistencies observed in the data from Experiment 2 and Experiment 3 indicate the need for further investigation on how to use pulse rate to encode FS information and how to best implement a multi-rate sound processing strategy such that the listener is able to extract the FS information that is available.

ACKNOWLEDGMENTS

The authors would like to thank the subjects who participated in this study for their time and patience, and the anonymous reviewers and Brian Moore for their comments and suggestions on this manuscript. This work was supported by the National Institutes of Health R01-DC007994.

¹The steps required to safely present stimuli to a Freedom Implant in Monopolar 1+2 mode via the SPEAR3 have been implemented since the completion of this experiment (Personal Communication: Colin Irwin of Cochlear Corporation). See Stohl *et al.*, 2008 for details.

²A current step is a unit defined by Cochlear Corporation, and has a logarithmic relationship to current: $I = a175C/255$, where I is current in amperes, $a = 10 \mu A$, and $C \in [0, 255]$ is current step.

- Arnoldner, C., Riss, D., Brunner, M., Durisin, M., Baumgartner, W.-D., and Hamzavi, J.-S. (2007). "Speech and music perception with the new fine structure speech coding strategy: Preliminary results," *Acta Oto-Laryngol.* **127**, 1298–1303.
- Büchner, A., Edler, B., and Nogueira, W. (2005). "Fundamental frequency coding in NofM strategies for cochlear implants," in *Audio Engineering Society 118th Convention Barcelona, Spain*, p. 6515.
- Chen, H., and Zeng, F.-G. (2004). "Frequency modulation detection in cochlear implant subjects," *J. Acoust. Soc. Am.* **116**, 2269–2277.
- Collins, L. M., Zwolan, T. A., and Wakefield, G. H. (1997). "Comparison of electrode discrimination, pitch ranking, and pitch scaling data in postlingually deafened adult cochlear implant subjects," *J. Acoust. Soc. Am.* **101**, 440–455.
- David, H. (1988). *The Method of Paired Comparisons* (Oxford University Press, New York).
- Eddington, D., Dobelle, W., Brackmann, D., Mladejovsky, M., and Parkin, J. (1978). "Place and periodicity pitch by stimulation of multiple scala tympani electrodes in deaf volunteers," *Trans. Am. Soc. Artif. Intern. Organs* **24**, 1–5.
- Fearn, R., Carter, P., and Wolfe, J. (1999). "The dependence of pitch perception on the rate and place of stimulation of the cochlea: A study using cochlear implants," *Ear Hear.* **27**, 41–43.
- Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R., and Olszewski, C. (2007). "Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise," *Ear Hear.* **28**, 412–423.
- Gibbons, J. D. (1985). *Nonparametric Statistical Inference*, Statistics: Textbooks and Monographs, Vol. **65**, 2nd ed. (McGraw-Hill, Dallas, TX).
- Grayden, D., Burkitt, A., Kenny, O., Clarey, J., Paolini, A., and Clark, G. (2004). "A cochlear implant speech processing strategy based on an auditory model," in *Proceedings of the Intelligent Sensors, Sensor Networks and Information Processing Conference*, 2004, pp. 491–496.
- Kendall, M., and Smith, B. B. (1940). "On the method of paired comparisons," *Ear Hear.* **31**, 324–345.
- Kiefer, J., Hohl, S., Sturzebecher, E., Pfennigdorff, T., and Gstöettner, W. (2001). "Comparison of speech recognition with different speech coding strategies (SPEAK, CIS, and ACE) and their relationship to telemetry measures of compound action potentials in the nucleus CI24M cochlear implant system," *Audiology* **40**, 32–42.
- Koch, D. B., Osberger, M. J., Segel, P., and Kessler, D. (2004). "Hiresolution and conventional sound processing in the hiresolutiontm bionic ear: Using appropriate outcome measures to assess speech recognition ability," *Audiol. Neuro-Otol.* **9**, 214–223.
- Kong, Y.-Y., and Zeng, F.-G. (2006). "Temporal and spectral cues in man-

- darin tone recognition," *J. Acoust. Soc. Am.* **120**, 2830–2840.
- Lan, N., Nie, K., Gao, S., and Zeng, F. (2004). "A novel speech-processing strategy incorporating tonal information for cochlear implants," *Ear Hear.* **51**, 752–760.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Luo, X., and Fu, Q.-J. (2007). "Frequency modulation detection with simultaneous amplitude modulation by cochlear implant users," *J. Acoust. Soc. Am.* **122**, 1046–1054.
- McDermott, H. J. (2004). "Music perception with cochlear implants: A review," *Trends Amplif.* **8**, 49–82.
- McDermott, H. J., and McKay, C. M. (1997). "Musical pitch perception with electrical stimulation of the cochlea," *J. Acoust. Soc. Am.* **101**, 1622–1630.
- McGill, R., Tukey, J. W., and Larsen, W. A. (1978). "Variations of box plots," *Am. Stat.* **32**, 12–16.
- McKay, C. M., McDermott, H. J., and Carlyon, R. P. (2000). "Place and temporal cues in pitch perception: Are they truly independent?," *Ear Hear.* **1**, 25–30.
- McKay, C. M., Remine, M. D., and McDermott, H. J. (2001). "Loudness summation for pulsatile electrical stimulation of the cochlea: Effects of rate, electrode separation, level and mode of stimulation," *J. Acoust. Soc. Am.* **110**, 1514–1524.
- Moses, L. E. (1952). "Non-parametric statistics for psychological research," *Psychol. Bull.* **49**, 122–143.
- Nie, K., Stickney, G., and Zeng, F.-G. (2005). "Encoding frequency modulation to improve cochlear implant performance in noise," *IEEE Trans. Biomed. Eng.* **52**, 64–73.
- Pijl, S. (1995). "Musical pitch perception with pulsatile stimulation of single electrodes in patients implanted with the nucleus cochlear implant," *Ann. Otol. Rhinol. Laryngol. Suppl.* **166**, 224–227.
- Shannon, R. V. (1983). "Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics," *Hear. Res.* **11**, 157–189.
- Stohl, J. S., Throckmorton, C. S., and Collins, L. M. (2008). "Developing a flexible spear3-based psychophysical research platform for testing cochlear implant users," Technical Report, Duke University, Durham, NC.
- Throckmorton, C. S., and Collins, L. M. (2001). "A comparison of two loudness balancing tasks in cochlear implant subjects using bipolar stimulation," *Ear Hear.* **22**, 439–448.
- Throckmorton, C. S., Kucukoglu, M. S., Remus, J. J., and Collins, L. M. (2006). "Acoustic model investigation of a multiple carrier frequency algorithm for encoding fine frequency structure: Implications for cochlear implants," *Hear. Res.* **218**, 30–42.
- Tong, Y., and Clark, G. (1985). "Absolute identification of electric pulse rates and electrode positions by cochlear implant patients," *J. Acoust. Soc. Am.* **77**, 1881–1888.
- Tong, Y. C., Clark, G. M., Blamey, P. J., Busby, P. A., and Dowell, R. C. (1982). "Psychophysical studies for two multiple-channel cochlear implant patients," *J. Acoust. Soc. Am.* **71**, 153–160.
- Tong, Y., Dowell, R., Blamey, P., and Clark, G. (1983). "Two-component hearing sensations produced by two-electrode stimulation in the cochlea of a deaf patient," *Science* **219**, 993–994.
- Townshend, B., Cotter, N., Compemolle, D. V., and White, R. (1987). "Pitch perception by cochlear implant subjects," *J. Acoust. Soc. Am.* **82**, 106–114.
- Umeda, N. (1975). "Vowel duration in American English," *J. Acoust. Soc. Am.* **58**, 434–445.
- Umeda, N. (1977). "Consonant duration in American English," *J. Acoust. Soc. Am.* **61**, 846–858.
- van Hoesel, R. J. M., and Tyler, R. S. (2003). "Speech perception, localization, and lateralization with bilateral cochlear implants," *J. Acoust. Soc. Am.* **113**, 1617–1630.
- Wei, C.-G., Cao, K., and Zeng, F.-G. (2004). "Mandarin tone recognition in cochlear-implant subjects," *Hear. Res.* **197**, 87–95.
- Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (1991). "Better speech recognition with cochlear implants," *Nature (London)* **352**, 236–238.
- Wilson, B., Sun, X., Schatzer, R., and Wolford, R. (2004). "Representation of fine structure or fine frequency information with cochlear implants," in *Eighth International Cochlear Implant Conference*, edited by R. T. Miyamoto (Elsevier, New York), Vol. **1273**, pp. 3–6.
- Wright, B. A. (2005). "Combined representations for frequency and duration in detection templates for expected signals," *J. Acoust. Soc. Am.* **117**, 1299–1304.
- Wright, B. A., and Fitzgerald, M. B. (2004). "The time course of attention in a simple auditory detection task," *Percept. Psychophys.* **66**, 508–516.
- Zeng, F.-G. (2002). "Temporal pitch in electric hearing," *Hear. Res.* **174**, 101–106.
- Zeng, F.-G. (2004). "Trends in cochlear implants," *Trends Amplif.* **8**, T1–T34.