

The Role of Temporal Cues in Voluntary Stream Segregation for Cochlear Implant Users

Trends in Hearing
Volume 22: 1–13
© The Author(s) 2018
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/2331216518773226
journals.sagepub.com/home/tia



Andreu Paredes-Gallardo¹ , Sara M. K. Madsen¹, Torsten Dau¹, and Jeremy Marozeau¹

Abstract

The role of temporal cues in sequential stream segregation was investigated in cochlear implant (CI) listeners using a delay detection task composed of a sequence of bursts of pulses (B) on a single electrode interleaved with a second sequence (A) presented on the same electrode with a different pulse rate. In half of the trials, a delay was added to the last burst of the otherwise regular B sequence and the listeners were asked to detect this delay. As a jitter was added to the period between consecutive A bursts, time judgments between the A and B sequences provided an unreliable cue to perform the task. Thus, the segregation of the A and B sequences should improve performance. The pulse rate difference and the duration of the sequences were varied between trials. The performance in the detection task improved by increasing both the pulse rate differences and the sequence duration. This suggests that CI listeners can use pulse rate differences to segregate sequential sounds and that a segregated percept builds up over time. In addition, the contribution of place versus temporal cues for voluntary stream segregation was assessed by combining the results from this study with those from our previous study, where the same paradigm was used to determine the role of place cues on stream segregation. Pitch height differences between the A and the B sounds accounted for the results from both studies, suggesting that stream segregation is related to the salience of the perceptual difference between the sounds.

Keywords

auditory streaming, cochlear implant, auditory perception

Date received: 15 January 2018; revised: 27 March 2018; accepted: 30 March 2018

Introduction

The cochlear implant (CI) is among the most successful neural prostheses (Zeng, Rebscher, Harrison, Sun, & Feng, 2008), making it possible for severely hearing-impaired listeners to achieve relatively high levels of speech intelligibility in quiet environments. However, CI listeners typically experience difficulties when listening in complex listening situations, where the auditory system is required to segregate the target signal from other competing sounds (e.g., Nelson, Jin, Carney, & Nelson, 2003; Stickney, Zeng, Litovsky, & Assmann, 2004). The auditory processes underlying the segregation of sounds have been a topic of research for many years (Bregman, 1990; Bregman & Campbell, 1971; Miller & Heise, 1950; Van Noorden, 1975). Bregman (1990) proposed two different mechanisms for the segregation process: *primitive* and *schema-driven* stream segregation.

Primitive stream segregation is assumed to be an involuntary and preattentive process, driven by the incoming acoustic data. Schema-driven stream segregation is, instead, assumed to require the attention of the listener and to involve the activation of stored knowledge of familiar patterns.

Auditory stream segregation has often been investigated using an *auditory streaming* paradigm (Bregman, 1990; Carlyon, 2004; Moore & Gockel, 2002, 2012; Van Noorden, 1975). In this paradigm, two repeating sounds

¹Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark, Kongens Lyngby, Denmark

Corresponding author:

Andreu Paredes-Gallardo, Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark, Ørstedes Plads Building 352, 2800 Kongens Lyngby, Denmark.
Email: apaga@elektro.dtu.dk



(A and B), typically two pure tones with different frequencies, are presented sequentially to the listener who might integrate them into a single stream or segregate them into two separate streams. Whereas large frequency differences between the sounds facilitate segregation, small frequency differences promote integration (e.g., Bregman & Campbell, 1971; Van Noorden, 1975). In normal-hearing (NH) listeners, stream segregation is also influenced by other stimulus properties than frequency differences, such as differences in the temporal envelope (e.g., Cusack & Roberts, 2000; Grimault, Bacon, & Micheyl, 2002; Iverson, 1995; Singh & Bregman, 1997; Vliegen and Oxenham, 1999), the phase spectrum (e.g., Roberts, Glassberg, & Moore, 2002), or the spatial characteristics (e.g., David, Grimault, & Lavandier, 2015; Sach & Bailey, 2004; Stainsby, Fu, Flanagan, Waldman, & Moore, 2011). Therefore, it has been hypothesized that sequential stream segregation may be directly related to the degree of the perceptual difference between the sounds (Moore & Gockel, 2002, 2012).

Van Noorden (1975) showed that the frequency difference needed to perceptually segregate the A sound from the B sounds depends on the intention of the listener. A smaller difference is needed when the listener is encouraged to segregate the sounds (i.e., *voluntary* stream segregation) than when the listener is encouraged to integrate them (i.e., *obligatory* stream segregation). Based on this, Van Noorden defined two boundaries: the *fission boundary*, which represents the smallest difference at which the sounds can be segregated, and the *temporal coherence boundary*, representing the largest difference at which the sounds can still be integrated. Thus, the temporal coherence boundary and the fission boundary represent the thresholds of obligatory and voluntary stream segregation, respectively. The temporal coherence boundary depends on the presentation rate of the sounds, with faster presentation rates promoting the segregation of the sounds. Conversely, the fission boundary is independent of the presentation rate of the sounds. The duration of the sequence of the A and B sounds also affects its perception, as the probability of achieving a segregated percept has been reported to increase over time (for a review, see Bregman, 1990; Moore & Gockel, 2002, 2012). This phenomenon, often referred to as the *build-up* of stream segregation, has been reported both in studies using integration-promoting listening instructions (e.g., Roberts, Glassberg, & Moore, 2008; Thompson, Carylton, & Cusack, 2011), neutral listening instructions (e.g., Anstis & Saida, 1985; Bregman, 1978; Van Noorden, 1975) as well as segregation-promoting listening instructions (e.g., Micheyl, Carylton, Cusack, & Moore, 2005; Nie & Nelson, 2015).

In electric hearing, perceptual differences can be elicited by varying the electrode (place cues) or the pulse

rate (temporal cues) of the stimulation (e.g., Eddington, Dobbelle, Brackmann, Mladejovsky, & Parkin, 1978; Lamping, Santurette, & Marozeau, 2018; Landsberger, Vermeire, Claes, Van Rompaey, & Van De Heyning, 2016; Shannon, 1983). Both electrode and pulse rate of stimulation can contribute to the perception of pitch height. The stimulation of apical electrodes and the use of low pulse rates are generally associated with a lower pitch percept than the stimulation of basal electrodes and the use of a high pulse rate (e.g., Lamping et al., 2018; Landsberger et al., 2016). It has been suggested that CI listeners might be able to combine place and rate information (e.g., Luo, Padilla, & Landsberger, 2012; McKay, McDermott, & Carylton, 2000; Rader, Döge, Adel, Weissberger, & Baumann, 2016). However, McKay et al. (2000) reported no advantage of consistent combinations of place and rate information (e.g., a slow pulse rate stimulating an apical electrode) over inconsistent combinations (e.g., a slow pulse rate stimulating a basal electrode) for the discrimination of sounds. Thus, place and temporal cues are considered to be perceptually orthogonal and independent cues in electric hearing (Marimuthu, Swanson, & Mannell, 2016; McKay et al., 2000; Tong, Blamey, Dowell, & Clark, 1983).

Previous studies investigating auditory stream segregation in CI listeners have focused on the role of place cues. In contrast to the results from studies in NH listeners, some of these studies did not observe any effect of the sequence duration (i.e., build-up) or the tone presentation rate on the ability to segregate the sounds (Chatterjee, Sarampalis, & Oba, 2006; Cooper & Roberts, 2007, 2009). Nevertheless, other studies found results consistent with those from studies in NH listeners, suggesting that there are circumstances in which CI listeners can use place cues to segregate sounds (Böckmann-Barthel, Deike, Brechmann, Ziese, & Verhey, 2014; Chatterjee et al., 2006; Hong & Turner, 2006; Paredes-Gallardo, Madsen, Dau, & Marozeau, 2018; Tejani, Schwartz-Leyzac, & Chatterjee, 2017). Moreover, the results from several studies suggest that CI listeners need time to build up a segregated percept (Böckmann-Barthel et al., 2014; Hong & Turner, 2006; Paredes-Gallardo et al., 2018), even though the build-up might be slower for CI listeners than for NH listeners (Paredes-Gallardo et al., 2018).

The effect of temporal cues on stream segregation in CI listeners has been investigated by Chatterjee et al. (2006), Duran, Collins, & Throckmorton (2012), and Hong & Turner (2009). The results from these studies suggest that large differences in the amplitude modulation or the pulse rate between the A and the B sounds facilitate both voluntary stream segregation (Chatterjee et al., 2006; Hong & Turner, 2009) and obligatory stream segregation (Duran et al., 2012). Chatterjee et al. observed a larger probability of a two-

stream percept with increasing sequence duration (i.e., build-up) in one listener. To our knowledge, no other study has investigated whether CI listeners experience a build-up as a function of pulse rate or amplitude modulation differences.

The present study examined the role of temporal cues on voluntary stream segregation in CI listeners. Delay detection performance was measured in a paradigm where the listeners were required to make time judgments between consecutive sounds of a target stream while ignoring a temporally irregular distractor stream. The task became easier if the listeners could segregate the target from the distractor and, thus, the performance in the detection task was affected by the stream segregation ability of the listeners. This paradigm has previously been used to investigate the role of spectral and temporal cues on stream segregation in NH listeners (e.g., Nie & Nelson, 2015; Nie, Zhang, & Nelson, 2014) and the role of place cues in CI listeners (Paredes-Gallardo et al., 2018). Here, temporal cues were induced by varying the pulse rate at a fixed cochlear location. The aim of the present study was to clarify whether CI listeners can use pulse rate differences (Δ rate) to segregate the streams and whether a two-stream percept builds up over time. The fission boundary was estimated as a function of Δ rate. Furthermore, the contribution of place versus temporal cues for voluntary stream segregation was assessed by combining the results from this study with those from Paredes-Gallardo et al. (2018). Electrode and Δ rate were converted to pitch height differences (Δ pitch) using data from a verbal attribute magnitude estimate experiment (Lamping et al., 2018). If stream segregation is related to the salience of the perceptual difference between the sounds, the Δ pitch between the target and the distractor streams should account for the results from both studies.

Methods

Listeners

Seven CI listeners (six women and one man) participated in this experiment. The listeners were aged between 19 and 74 years (mean: 50.8 years, standard deviation [*SD*]: 21.5 years; see Table 1), had no residual hearing and were bilateral CI users. All listeners were users of the Cochlear Ltd. (Sydney, Australia) implant. Six of the listeners had previously participated in a study that used the same paradigm to assess the effect of place cues on stream segregation (Paredes-Gallardo et al., 2018). The same listener IDs were used in both studies. All listeners provided informed consent prior to the study and all experiments were approved by the Science-Ethics Committee for the Capital Region of Denmark (reference H-16036391).

Table 1. Relevant Information About CI Listeners.

Listener	Age	Gender	Onset of deafness	Implant model (ear)	Years of experience
1	19	F	Prelingual	CI24RE (right)	16
4	74	F	Postlingual	CI24R (left)	13
5	73	M	Postlingual	CI24RE (right)	3
6	64	F	Perilingual	CI24R (right)	15
8	61	F	Perilingual	CI24RE (right)	3
9	21	F	Perilingual	CI24RE (left)	16
10	44	F	Prelingual	CI24RE (left)	5

Note. CI24RE and CI24R are two implant models, also known as Nucleus-24 and Freedom, respectively. CI = cochlear implant.

Stimuli and Conditions

The stimulation paradigm is illustrated in Figure 1, where the different panels represent the different conditions. A sequence of 50 ms bursts of pulses (B) presented on a single electrode was interleaved with a sequence (A) presented on the same electrode with a different pulse rate. In half of the trials, a small delay (Δt) was added to the last burst of the otherwise regular B sequence (the target stream). The listeners were asked to indicate after each trial whether or not the last sound of the sequence was delayed. The nominal onset-to-onset interval between consecutive B sounds was 340 ms, and a random jitter was added to the onset-to-onset interval between consecutive A sounds. The duration of the jitter applied to each A sound was drawn from a rectangular distribution with a range of ± 110 ms. Thus, the onset-to-onset interval between the A and B sounds was $170 \text{ ms} \pm \text{jitter}$, as illustrated in Figure 2. Consecutive A and B sounds were always separated by a minimum interval of 10 ms. The temporal irregularity of the distractor stream made across-streams time judgments an unreliable cue to perform the task. Therefore, to optimize performance, the listeners needed to compare the time interval between the last two B sounds with those between previous B sounds. Thus, the task became easier when the A and B sequences fell into different streams (Micheyl and Oxenham, 2010; Nie & Nelson, 2015; Nie et al., 2014), encouraging the listener to segregate the streams.

Each A and B sound consisted of a 50-ms burst of biphasic pulses presented at Electrode 11,¹ located at the midpoint of the array, in monopolar mode. Each biphasic pulse had a phase width of $25 \mu\text{s}$ and a phase gap of $8 \mu\text{s}$. The stimuli were presented through the Nucleus Implant Communicator research interface (NIC v2, Cochlear Limited, Sydney).

The ability of CI listeners to perceive pitch changes as a function of pulse rate (temporal pitch) has been

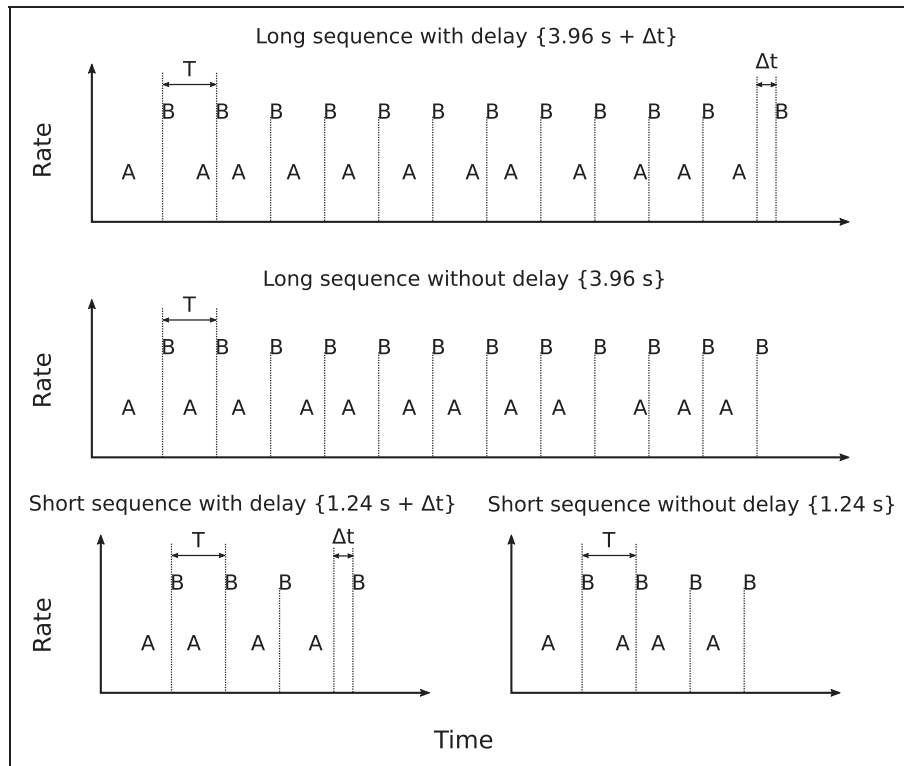


Figure 1. Graphical representation of the experimental paradigm. T represents the onset-to-onset interval and Δt is the delay of the last B sound. The long sequence, with and without Δt , is shown in the upper and middle panels. The short sequence, with and without Δt , is illustrated in the lower left and lower right panels, respectively. The rate difference between A and B sounds varied across conditions.

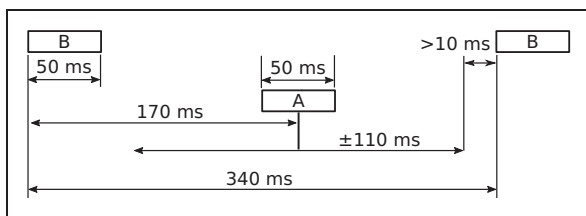


Figure 2. Graphical representation of the timing between the B and the A sounds.

reported to be limited to rates below 300/400 pulse per sound (pps; e.g., Shannon, 1983; Tong & Clark, 1985; Townshend, Cotter, Van Compennolle, & White, 1987). In this study, the target stream was played with a constant rate of 300 pps, while the A sequence was played with a lower pulse rate of either 80, 140, 200, or 260 pps, leading to a Δ rate between the streams of 220, 160, 100, or 40 pps depending on the condition.

All sequences started with the distractor stream (A) and ended with the target stream (B), as illustrated in Figure 1. Two sequence durations were tested. The long sequence consisted of 12 AB pairs (3.96 s without Δt) and the short sequence consisted of 4 AB pairs (1.24 s without Δt). Performance in the detection task for

the long and the short sequences was also measured without the distractor stream (control conditions). These conditions were easier than the test conditions and, thus, a shorter Δt was used to avoid ceiling effects. A *no difference* condition (Δ rate = 0) was also tested for the long sequence. In this condition, both the target and the distractor were presented from the same electrode and with the same pulse rate. Both the control and the *no difference* condition were identical to those described in the study by Paredes-Gallardo et al. (2018). Thus, only Listener 10, who did not participate in the study by Paredes-Gallardo et al., performed those conditions in this study. For the remaining listeners, the results from the control and the *no difference* condition were obtained from Paredes-Gallardo et al.

For each combination of Δ rate and sequence duration, 60 presentations of the delayed sequence and 60 presentations of the non-delayed sequence were used to calculate the listener's sensitivity (d') to the delayed target.

Loudness Balancing

Loudness has been found to be an effective cue for sound segregation in CI listeners (e.g., Cooper & Roberts, 2009; Marozeau, Innes-Brown, & Blamey, 2013). The stimuli

Table 2. Individual Δt As Obtained From the Delay Adjustment Procedure.

Listener	Δt (ms)	Δt (ms) for control condition
1	40	30
4	45	35
5	35	32
6	80	55
8	60	28
9	35	30
10	60	40

were therefore loudness-balanced in this study. Categorical loudness scaling was used to find the most comfortable levels for each listener and stimulus. Each pair of target and distractor sounds was then loudness matched by the listeners using the procedure described in Paredes-Gallardo et al. (2018). The loudness matching was performed in the beginning of each session. The level of the loudness balanced stimuli did not markedly change between sessions.

Delay (Δt) Adjustment Procedure

Individual Δt values were chosen such that listeners would be equally sensitive to the delayed target in a given condition, minimizing the effect of individual differences on the detection performance in the auditory streaming task. To facilitate the comparison of the results from this study and those from Paredes-Gallardo et al. (2018), the same individual Δt values were used in the two studies. For Listener 10, who did not participate in the study by Paredes-Gallardo et al., Δt was derived using the same criterion as in the study by Paredes-Gallardo et al.: Δt was defined as the delay leading to $d' = 2$ for the long sequence whereby the 50 ms bursts of pulses were presented at 900 pps to Electrodes 11 (A) and 19 (B) (see table 2).

The individual Δt to be used in the control condition was also derived as in Paredes-Gallardo et al. (2018), that is, the Δt corresponding to $d' = 3$ for the long sequence without the distractor stream. This d' value was chosen to keep the control conditions relatively easy while avoiding ceiling effects.

Procedure

The experiments took place in a double-walled sound-attenuating booth and were organized into two sessions, each lasting 2 h including short breaks. For Listener 10, the first session included a brief description of the task and the delay adjustment procedure, as well as a 10- to 15-min training on the detection task. The other listeners

had participated in the study by Paredes-Gallardo et al. (2018) and were therefore familiar with the paradigm.

A one-interval two-alternative forced-choice procedure was used, where the listeners were asked to report if the last target sound of the sequence was delayed or not. A one interval task was chosen to minimize the attentional effort required to perform the task (Nie & Nelson, 2015). The sequences were organized in 12 blocks, 6 with long sequences and 6 with short sequences, presented in random order. On a given block, the four Δ rate conditions were presented in pseudorandom order, ensuring that the same Δ rate condition would not be presented in consecutive sequences. Thus, the first sound of each sequence always had a different rate, contributing to resetting the build-up of a two-stream percept after each presentation (Roberts et al., 2008). Each Δ rate condition was presented 20 times in each block (10 delayed and 10 non-delayed presentations). The *no difference* condition was tested in a separate block.

The control conditions were tested in four blocks (two with long sequences and two with short sequences), with each block containing 30 repetitions of the delayed and 30 repetitions of the non-delayed sequences. The control blocks were randomly presented at the beginning or at the end of either session.

Ideal Observer Model

The distribution of possible onset-to-onset gaps between the last A and B sounds was different in the delayed and the non-delayed sequences. The gap between the last A and B sounds in the delayed sequence was, on average, Δt ms longer than the one in the non-delayed sequence. Therefore, the listeners had an extra cue, proportional to Δt , to perform the task. As in Paredes-Gallardo et al. (2018), an ideal observer (IO) model was used to simulate the best possible performance that each listener could achieve if the gap between the last A and B sounds would be the only available cue. The model categorized individual trials as delayed or non-delayed by evaluating the gap between the last A and B sounds of a given sequence and comparing it with the nominal gap between consecutive A and B sounds (i.e., the gap of 170 ms, when no jitter has been applied). A given trial was categorized as delayed if the gap between the last A and B sounds was larger than the nominal gap. Otherwise, the trial was categorized as non-delayed. Because Δt was adjusted individually, the probability of a correct response when fusing the A and B streams (chance level) was different for each listener. Thus, the gap between the last A and B sounds of each presentation, listener, and condition was used as input to the IO model. The IO model generated a d' estimate for each listener and condition. Segregation was considered to occur when the CI listeners' performance was significantly better than the one predicted by the IO model.

Statistical Analysis

Unless otherwise specified, statistical inference was performed by fitting a mixed-effects linear model to the computed d' scores. The experimental variables and their interactions were treated as fixed effects whereas listener-related effects were treated as random effects with random intercepts and slopes. The Δ rate values were calculated from the log-transformed pulse rate values and were back-transformed after the post hoc analysis for an easier interpretation of the results. The model was implemented in R using the *lme4* library (Bates, Mächler, Bolker, & Walker, 2014) and the model selection was carried out with the *lmerTest* library (Kuznetsova, Brockhoff, & Christensen, 2017) following the backward selection approach based on stepwise deletion of model terms with high p values (Kuznetsova, Christensen, Bavay, & Brockhoff, 2015). The p values for the fixed effects were calculated from F tests based on Sattethwaite's approximation of denominator degrees of freedom and the p values for the random effects were calculated based on likelihood ratio tests (Kuznetsova et al., 2015). The post hoc analysis was performed through contrasts of least-square means using the *lsmeans* library (Lenth, 2016) and the *lme4* model object. The p values were corrected for multiple comparisons using the Tukey method.

Statistical contrasts between the individual listeners' data and their respective IO model predictions were performed using t tests with the mean and standard error from each d' estimate. The resulting p values were adjusted for multiple comparisons controlling for the false discovery rate (Benjamini & Hochberg, 1995).

Results

The individual results are shown in Figure 3, where each row represents the results for an individual listener. Sensitivity scores for the short and long sequences are shown in the left and right columns, respectively. The experimental data are indicated by the blue circles. Estimates from the IO model are indicated by the green triangles. Statistically significant differences between the achieved d' scores and the IO model predictions are indicated by the asterisks.

For the long sequence (right column), d' scores generally increased with increasing Δ rate. For the largest Δ rate condition, all listeners achieved larger d' scores than the IO model. In contrast, for the *no difference* condition (Δ rate = 0), none of the listeners achieved significantly larger d' scores than the IO model. A large across-listener variability was observed, with some listeners exhibiting little or no improvement in detection performance toward larger Δ rate values. For the short sequence (left column), no general trend was observed in

the d' scores with increasing Δ rate. For some listeners, d' scores did not significantly increase with increasing Δ rate (i.e., Listeners 1, 5, 8, and 9). For three of the seven listeners (i.e., Listeners 4, 6, and 10), d' scores increased with increasing Δ rate and were larger than the IO model predictions for the largest Δ rate condition (i.e., 220 pps).

Figure 4 shows the d' scores for all listeners and conditions. The results from the short and long sequences are shown in separate panels. The results for the control (no distractor) condition for the short and the long sequences are shown in the rightmost panel. The *lsmeans*

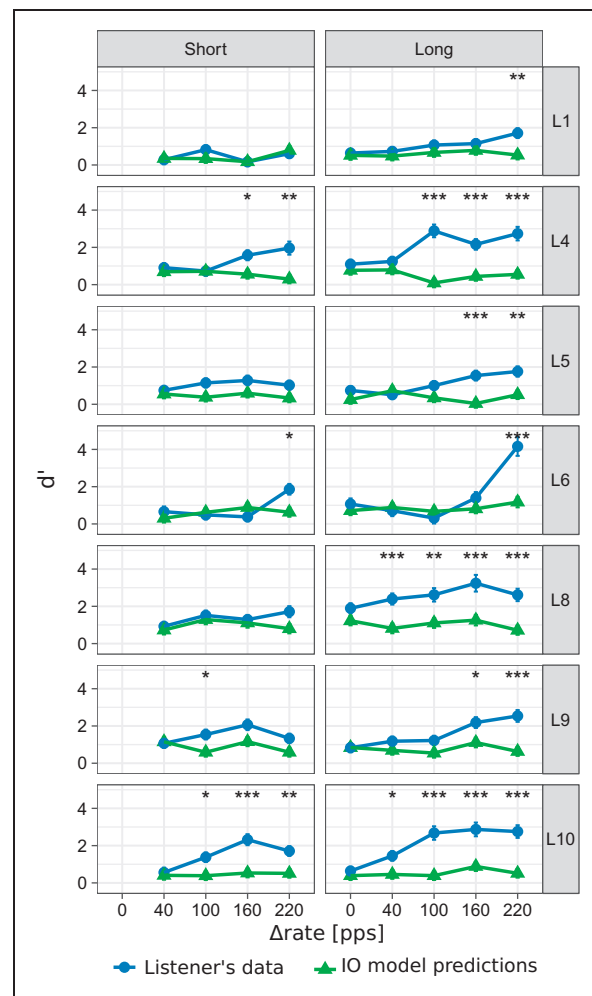


Figure 3. Individual sensitivity (d') scores to the delayed B sound for each Δ rate and sequence duration (blue circles) as well as the corresponding ideal observer model prediction (green triangles). Error bars represent the standard error of the d' estimates. The error bars often fall within the symbols and are therefore not always visible in the graph. A statistically significant difference between the IO model predictions and the listener's data is indicated by one asterisk if $.05 > p > .01$, two asterisks if $.01 > p > .001$ and three asterisks if $p < .001$.

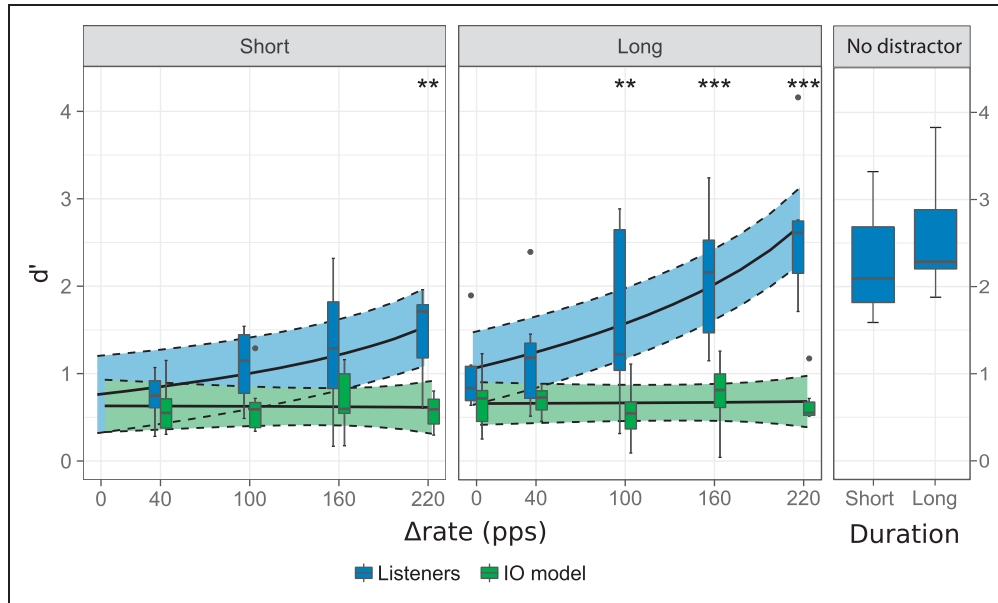


Figure 4. Sensitivity (d') scores to the delayed B sound for each Δ rate and sequence duration. The control condition (no distractor) for both long and short sequences is shown in the rightmost panel. The boxes illustrate data from the listeners (blue) and the corresponding IO model predictions (green). The solid lines represent the lsmeans estimate from the statistical model. Its 95% confidence interval is indicated with dashed lines and the corresponding shaded area. A statistically significant difference between the IO model predictions and the listener's data is indicated by one asterisk if $.05 > p > .01$, two asterisks if $.01 > p > .001$ and three asterisks if $p < .001$.

estimates and the 95% confidence interval from the statistical model fitted to the data are represented by solid and dashed lines, respectively. The data from the listeners and the predictions from the IO model are represented by boxes. Different colors represent the measured data and the IO model predictions.

The sensitivity scores (d') increased with Δ rate [$F(1, 29.95) = 23.051, p < .001$]. The main effect of sequence duration [$F(1, 98.70) = 1.990, p = .162$], and of the data type (listener's vs. IO model) [$F(1, 11.53) = 2.526, p = .139$], were found to be nonsignificant. However, a significant interaction was found between Δ rate and sequence duration [$F(1, 93.11) = 5.277, p = .024$]; Δ rate and data type [$F(1, 80.75) = 38.958, p < .001$]; and Δ rate, sequence duration, and data type [$F(1, 80.75) = 4.609, p = .035$].

The increase in the d' scores obtained by the listeners with increasing Δ rate was significantly steeper for the long sequence than for the short sequence [$t(86.47) = 3.134, p = .012$], indicating a greater effect of Δ rate for the long than for the short sequences. For the long sequences, the increase of the d' scores with increasing Δ rate was significantly steeper for the listeners than for the IO model predictions [$t(79.65) = 6.557, p < .001$]. The listeners performed significantly better than the IO model for the Δ rate values of 100 pps [$t(8.60) = 5.784, p = .006$]; 160 pps [$t(9.63) = 8.354, p < .001$]; and 220 pps [$t(27.59) = 9.454, p < .001$]. Thus, for the long sequence, the smallest Δ rate at which the listeners could segregate the streams (i.e., the fission boundary) was 100 pps (50%

relative to the distractor pulse rate). For the short sequences, the increase of the d' scores with increasing Δ rate was only marginally steeper for the listeners than for the IO model predictions [$t(79.65) = 2.664, p = .045$]. The listeners achieved significantly larger d' scores than those from the IO model only for the largest Δ rate condition (Δ rate = 220 pps) [$t(29.62) = 4.214, p = .007$], difference estimate = 0.915. Thus, for the short sequences, the fission boundary was 220 pps (275% relative to the distractor pulse rate).

A paired t test revealed no significant difference between the d' scores achieved for the long and short sequences in the control condition (no distractor) [$t(6) = 1.515, p = .180$].

In summary, the performance in the delay detection task improved with increasing Δ rate. A larger effect of Δ rate was observed for the long than for the short sequence, indicating the build-up of stream segregation.

Discussion

In this study, a delay detection task was used to assess the stream segregation abilities of CI listeners. The task became easier when the listeners could segregate the sounds—hence, larger d' scores were achieved in conditions facilitating a segregated percept. Segregation was considered to occur when the d' scores achieved by the CI listeners were significantly larger than those predicted by the IO model.

The Role of Δ rate in Stream Segregation

The d' scores obtained by the listeners increased with increasing Δ rate, both for the long and for the short sequences, suggesting that the listeners were able to use Δ rate to segregate the streams. These findings are consistent with earlier work suggesting that larger differences between the temporal envelopes of the A and the B sounds facilitate a segregated percept both in NH listeners (e.g., Grimault et al., 2002; Roberts et al., 2002; Vliegen, Moore, & Oxenham, 1999; Vliegen and Oxenham, 1999) and in CI listeners (Chatterjee et al., 2006; Duran et al., 2012; Hong & Turner, 2009). Hong & Turner (2009) investigated the role of amplitude modulation differences in stream segregation both in NH and in CI listeners. They used a rhythm detection task that became easier when the A and B sounds were perceptually segregated. In their study, both groups of listeners were found to be able to use differences in the temporal envelope of sequential sounds to voluntarily segregate them. Hong & Turner presented the stimuli through a loudspeaker and the CI listeners used their own speech processor. Therefore, they had only limited control over the exact stimuli delivered to the listeners, as noted by Cooper & Roberts (2009). The results from this study support the findings from Hong and Turner (2009). However, in this study, temporal cues were elicited by directly manipulating the stimulation rate at a fixed cochlear location, bypassing the listener's speech processor such that there was a better control of the signal delivered to the listeners.

It has been suggested that NH listeners might need larger differences to perceptually segregate two stimuli than to discriminate them (Rose & Moore, 2005). Hong and Turner (2009) measured amplitude modulation frequency discrimination thresholds in NH and CI listeners and compared them with the fission boundary obtained as a function of the amplitude modulation frequency difference between two noise bursts. In both groups, larger amplitude modulation frequency differences were needed to segregate the two sounds than to discriminate them. Previous studies assessing the pulse rate difference limen in CI listeners reported a large variability across listeners and a strong dependency of the difference limen on the pulse rate of the reference sound, that is, the *base rate* (e.g., Baumann & Nobbe, 2004; Townshend et al., 1987; Van Hoesel and Clark, 1997; Zeng, 2002). The difference limen was found to increase with increasing base rate, with values of about 10% at a base rate of 100 pps and about 20% for a base rate of 200 pps. Consistent with the findings from Hong and Turner, the results from this study suggest that CI listeners need larger differences to segregate the sounds than to discriminate them. This was particularly evident for the short sequence, where a pulse rate difference of

275% of the base rate (80 pps) was needed to segregate the sounds.

Duran et al. (2012) also assessed the role of temporal cues in stream segregation by changing the pulse rate at a fixed cochlear location. Their results suggested that CI listeners can use Δ rate to segregate sounds. While in this study the task became easier when the sounds were perceptually segregated (voluntary stream segregation paradigm), in the study by Duran et al., the task became easier if the sounds were integrated into a single stream (obligatory stream segregation paradigm). Together, these results suggest that CI listeners can use Δ rate for both voluntary and obligatory stream segregation of sequential sounds.

The Build-up of a Two-Stream Percept

The d' scores obtained by the listeners increased with increasing Δ rate. The effect of Δ rate was found to be dependent on the sequence duration, with a steeper growth of d' with increasing Δ rate for the long than for the short sequence. Given that in the absence of the distractor stream the performance was not affected by the duration of the sequence, as demonstrated by the results from the no distractor condition, these findings suggest that longer sequences facilitated the segregation of the A and B sounds (i.e., there was evidence of build-up). Chatterjee et al. (2006) observed evidence of build-up in one CI listener, who was instructed to qualitatively report whether a given sequence of sounds was integrated or segregated. In the present study, a detection task was used to assess stream segregation "objectively" (Cooper & Roberts, 2009; Hong & Turner, 2009; Micheyl and Oxenham, 2010; Roberts et al., 2002) and seven CI listeners performed the task. The results from this study support the observations reported in Chatterjee et al.

The results presented here are also consistent with the findings from Nie and Nelson (2015), who investigated the effects of amplitude modulation rate and sequence duration on voluntary stream segregation in NH listeners. Nie and Nelson used modulated bandpass noise bursts to simulate the degraded spectral cues present in electric hearing. With a similar task and using similar sequence durations, both studies found an interaction between the temporal cue (amplitude modulation or pulse rate difference) and the sequence duration, suggesting that a similar build-up process might be experienced by CI listeners and NH listeners. Nie and Nelson found that spectral cues (i.e., a difference between the center frequencies of the noise bands) elicited a build-up both in the presence and in the absence of temporal cues. However, temporal cues elicited a build-up when combined with moderate spectral differences, but not in the absence of spectral cues, suggesting that temporal cues could be a weaker or secondary cue for the segregation

of sounds. In this study, temporal cues elicited a build-up even in the absence of place cues.

Shorter Δt values were needed in the control condition (no distractor) to avoid ceiling effects. This reflects the difficulty experienced by the CI listeners in performing the task in the presence of a distractor stream, even when a large $\Delta rate$ and a long sequence duration were used. Thus, even though CI listeners seem to be able to achieve a segregated percept and exhibit a similar build-up process as NH listeners, they are not able to completely ignore a competing stream, which may reflect a slower build-up in CI listeners than in NH listeners.

Contribution of Temporal Regularity Differences to Stream Segregation

In this study, a temporally irregular distractor stream was used to ensure that temporal judgments between the A and the B sounds would be an unreliable cue to perform the task. Temporally irregular patterns are more likely to be segregated than predictive and temporally regular patterns (for a review, see Bendixen, 2014). Even though the temporal irregularity of the distractor stream cannot account for the increase of the d' scores associated with larger $\Delta rate$ values, it is possible that the CI listeners made use of both $\Delta rate$ and regularity differences to segregate the streams. Nie et al. (2014) investigated the role of spectral separation for stream segregation in NH listeners with a paradigm similar to the one used in this study. Their results suggested that NH listeners could segregate the sounds when the only available cue was the temporal regularity of one stream versus the temporal irregularity of the other. This condition is similar to the *no difference* condition from this study. Nevertheless, the results from the *no difference* condition suggest that CI listeners were not able to segregate the streams when the A and B streams were presented through the same electrode and at the same pulse rate. Thus, even though temporal regularity differences between the streams could contribute to their segregation, this cue was not sufficiently salient for it to elicit a segregated percept in the absence of $\Delta rate$.

Place Versus Temporal Cues in Stream Segregation: The Role of $\Delta pitch$

Paredes-Gallardo et al. (2018) and this study employed the same paradigm to assess the role of electrode separation versus the role of $\Delta rate$ in voluntary stream segregation. In electric hearing, both electrode and stimulation pulse rate contribute to the perception of pitch height (e.g., Lamping et al., 2018). If stream segregation is correlated with the overall perceptual difference between the sounds, the perceptual $\Delta pitch$ between

the target and the distractor stream may account for the results obtained in the two studies. To test this, the $\Delta rate$ and the electrode separation values from this study and from Paredes-Gallardo et al. were converted to $\Delta pitch$ between the target and the distractor streams. Data from a verbal attribute magnitude estimate experiment (Lamping et al., 2018) were used to map specific single electrode stimuli to a perceptual pitch height scale (see the Appendix for more details).

Figure 5 shows the d' scores for the long sequence as a function of $\Delta pitch$ between the target and the distractor streams. On the basis of the magnitude estimation experiment for pitch height, the $\Delta pitch$ values were normalized such that a $\Delta pitch$ value of 100% corresponded to the perceived $\Delta pitch$ between Electrodes 11 and 22 at a pulse rate of 900 pps. The data from the CI listeners and predictions from the IO model are indicated by the blue and green boxes, respectively. The pitch differences elicited by varying the pulse rate are shown with a lighter color than pitch differences elicited by changing the stimulation electrode. The solid and dashed lines represent the estimates from the statistical model and its 95% confidence intervals, respectively. The cue used to elicit the pitch differences (electrode vs. $\Delta rate$) was found to be a nonsignificant factor [$F(1, 107.31) = 1.216, p = .273$]. No significant interaction was found between the cue and $\Delta pitch$ [$F(1, 105.90) = 0.295, p = .588$]; the cue and data type [$F(1, 105.40) = 1.101, p = .296$]; or the cue, $\Delta pitch$, and data type [$F(1, 98.26) = 0.004, p = .950$]. Only the $\Delta pitch$ [$F(1, 6.51) = 18.166, p = .004$], data type (listeners' data vs IO model predictions) [$F(1, 11.08) = 5.236, p = .043$], and their interaction [$F(1, 101.85) = 38.612, p < .001$], were found to be significant effects in the model.

The d' scores from the listeners increased for larger $\Delta pitch$ values. Moreover, the cue used to elicit the pitch difference was revealed to be a nonsignificant factor in the statistical model. This suggests that CI listeners can use both place and temporal cues to segregate the streams as long as the perceptual pitch difference between the streams is larger than the fission boundary (i.e., about 20% of the pitch difference between Electrodes 11 and 22), supporting the hypothesis proposed by Moore and Gockel (2002, 2012). These findings suggest that the combination of cues may improve stream segregation for CI listeners, provided that a larger overall perceptual difference is elicited between the sounds.

Six of the listeners from this study had previously participated in the study from Paredes-Gallardo et al. (2018). Thus, a learning effect might have affected the d' scores obtained by the listeners in this study. Nevertheless, the lack of a significant effect of the cue used to elicit the pitch difference in the combined data from both studies implies that there was not a systematic change in the d' scores from the two studies.

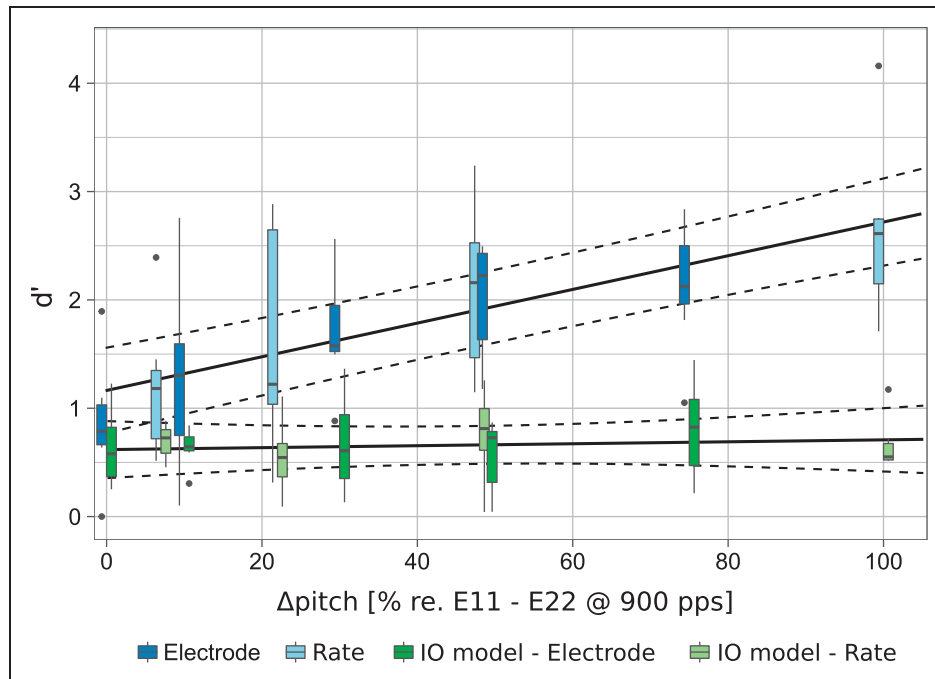


Figure 5. Sensitivity (d') scores to the delayed B sound for each Δ pitch between the target and the distractor streams. A Δ pitch of 100% corresponds to the pitch difference experienced between Electrodes 11 and 22 when stimulated with a pulse rate of 900 pps. The boxes illustrate data from the listeners (in blue) and the corresponding IO model predictions (in green). Dark colors represent the data from Paredes-Gallardo et al. (2018; i.e., electrode separation) and light colors represent the data from this study (i.e., Δ rate). The solid lines represent the lsmeans estimate from the statistical model. Its 95% confidence interval is indicated with dashed lines.

Summary and Conclusion

This study assessed the effect of temporal cues on voluntary stream segregation in CI listeners. The results suggested that CI listeners can make use of temporal cues to segregate sounds when attention is directed toward segregation. Moreover, a build-up process similar to that reported in NH listeners was observed. The similarity between the trends observed in this study for CI listeners and those reported for NH listeners suggest a common underlying mechanism for stream segregation in both groups. Furthermore, differences in the perceived pitch height accounted for the results from this study (temporal cues) as well as from Paredes-Gallardo et al. (2018) (place cues). This suggests that stream segregation is directly related to the salience of the perceptual difference between the sounds. Thus, the combination of cues may improve stream segregation in CI listeners.

Appendix A: Perceptual Mapping of Place and Temporal Cues

The Δ rate and the electrode separation values from this study and from Paredes-Gallardo et al. (2018) were converted to pitch height differences between the target and

the distractor streams (Δ pitch). Data from a verbal attribute magnitude estimation experiment (Lamping et al., 2018) were used to map specific single electrode stimuli to a perceptual pitch height scale. Lamping et al. (2018) collected responses from five CI listeners who were instructed to rate the pitch height of single electrode stimuli on a scale from 0 to 100. A rating of 100 reflected full agreement with the verbal attribute *high* and a rating of 0, no agreement. A combination of four electrodes (i.e., 10, 14, 18, 22) and five pulse rates (i.e., 80, 150, 300, 600, and 1,200 pulse per sound [pps]) were tested. A mixed-effects quadratic model was fitted to the median of the individual ratings over eight repetitions using the statistical software R (*lme4* and *lmerTest* libraries; Bates et al., 2014; Kuznetsova et al., 2017). Both stimulation electrode, pulse rate (log transformed) and their interaction were treated as fixed effects terms. Listener-related effects were treated as random effects with random intercepts and slopes. Pitch height ratings were defined as the lsmean estimates of the model (*lsmeans* library; Lenth, 2016) for each target and distractor stimuli. The Δ pitch between each target and distractor sound was then calculated. Values of Δ pitch were normalized such that a Δ pitch of 100 would correspond to the pitch difference between Electrodes 11 and 22 when stimulated at a pulse rate of 900 pps.

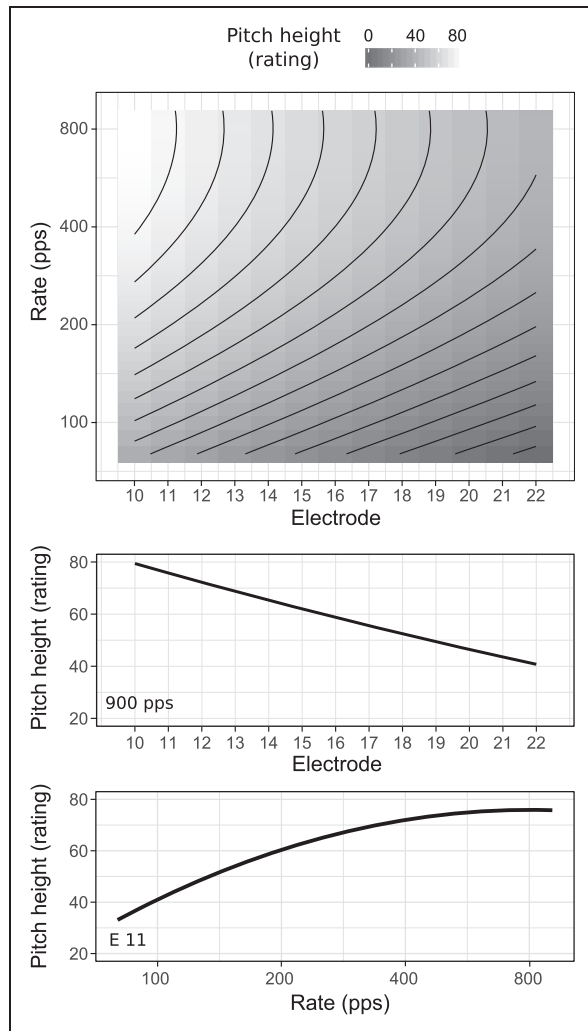


Figure A1. Upper panel: predictions of the pitch height ratings for different electrode and pulse rate combinations as obtained from the mixed-effects model fitted to the data from Lamping et al. (2018). Darker colors represent low ratings while brighter colors represent high ratings. The solid lines represent equal-rating contours. Middle panel: predictions of the pitch height ratings for different electrodes at a pulse rate of 900 pps. Bottom panel: predictions of the pitch height ratings for different pulse rates at Electrode 11.

Figure A1 shows the model predictions of pitch height ratings as a function of electrode and pulse rate (upper panel). High ratings are shown in white while low ratings are shown in dark gray. Equal-rating contours are indicated by the solid black lines. The middle and bottom panels show pitch height ratings for a fixed pulse rate and for a fixed electrode, respectively. The ratings of the pitch height decrease linearly as a function of stimulation electrode (middle panel). Conversely, pulse rate and pitch height exhibit a nonlinear relation (bottom panel) consistent with other studies (e.g., Landsberger

et al., 2016). Pitch height ratings increase up to a pulse rate of 300/400 pps and saturate for higher pulse rates.

Acknowledgments

The authors would like to thank all volunteers who participated in this study. The authors thank the two anonymous reviewers for the helpful and constructive comments on an earlier version of the article; Wiebke Lamping for sharing the data from the verbal attribute magnitude estimation experiment (and for many fruitful discussions); Per B. Brockhoff, Alexandra Kuznetsova, and Anders Stockmarr for their help and guidance with the statistical analysis of the data; and their colleagues from the Hearing Systems Group for valuable comments and stimulating discussions. Part of the research equipment was provided by Cochlear Ltd.

Declaration of Conflicting Interests

The authors declared no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Oticon Centre of Excellence for Hearing and Speech Sciences (CHeSS) and the Carlsberg Foundation.

Supplementary Material

The data presented here is publicly available at <http://doi.org/10.5281/zenodo.1126665>

Note

1. In the cochlear electrode array, Electrode 1 is the most basal electrode and Electrode 22 the most apical one.

ORCID iD

Andreu Paredes-Gallardo  <http://orcid.org/0000-0003-1371-3547>

References

- Anstis, S. M., & Saida, S. (1985). Adaptation to auditory streaming of frequency-modulated tones. *Journal of Experimental Psychology Human Perception & Performance*, *11*, 257–271. doi:10.1037/0096-1523.11.3.257
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 51. doi:10.18637/jss.v067.i01
- Baumann, U., & Nobbe, A. (2004). Pulse rate discrimination with deeply inserted electrode arrays. *Hearing Research*, *196*, 49–57. doi:10.1016/j.heares.2004.06.008
- Bendixen, A. (2014). Predictability effects in auditory scene analysis: A review. *Frontiers in Neuroscience*, *8*, 1–16. doi:10.3389/fnins.2014.00060
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*

- (*Methodological*), 57(1), 289–300. Retrieved from <http://www.jstor.org/stable/2346101>
- Bregman, A. S. (1978). Auditory streaming is cumulative. *Journal of Experimental Psychology Human Perception & Performance*, 4, 380–387. doi:10.1037/0096-1523.4.3.380
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: The MIT Press.
- Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 89, 244–249. doi:10.1037/h0031163
- Böckmann-Barthel, M., Deike, S., Brechmann, A., Ziese, M., & Verhey, J. L. (2014). Time course of auditory streaming: Do CI users differ from normal-hearing listeners? *Frontiers in Psychology*, 5, 775. doi:10.3389/fpsyg.2014.00775
- Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences*, 8, 465–471. doi:10.1016/j.tics.2004.08.008
- Chatterjee, M., Sarampalis, A., & Oba, S. I. (2006). Auditory stream segregation with cochlear implants: A preliminary report. *Hearing Research*, 222, 100–107. doi:10.1016/j.heares.2006.09.001
- Cooper, H. R., & Roberts, B. (2007). Auditory stream segregation of tone sequences in cochlear implant listeners. *Hearing Research*, 225, 11–24. doi:10.1016/j.heares.2006.11.010
- Cooper, H. R., & Roberts, B. (2009). Auditory stream segregation in cochlear implant listeners: Measures based on temporal discrimination and interleaved melody recognition. *The Journal of the Acoustic Society of America*, 126, 1975–1987. doi:10.1121/1.3203210
- Cusack, R., & Roberts, B. (2000). Effects of differences in timbre on sequential grouping. *Perception & Psychophysics*, 62, 1112–1120. doi:10.3758/BF03212092
- David, M., Grimault, N., & Lavandier, M. (2015). Sequential streaming, binaural cues and lateralization. *The Journal of the Acoustic Society of America*, 138, 3500–3512. doi:10.1121/1.4936902
- Duran, S. I., Collins, L. M., & Throckmorton, C. S. (2012). Stream segregation on a single electrode as a function of pulse rate in cochlear implant listeners. *The Journal of the Acoustic Society of America*, 132, 3849–3855. doi:10.1121/1.4764875
- Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, M. G., & Parkin, J. (1978). Place and periodicity pitch by stimulation of multiple scala tympani electrodes in deaf volunteers. *Transactions—American Society for Artificial Internal Organs*, 24, 1–5.
- Grimault, N., Bacon, S. P., & Micheyl, C. (2002). Auditory stream segregation on the basis of amplitude-modulation rate. *The Journal of the Acoustic Society of America*, 111, 1340–1348. doi:10.1121/1.1452740
- Hong, R. S., & Turner, C. W. (2006). Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients. *The Journal of the Acoustic Society of America*, 120, 360–374. doi:10.1121/1.2204450
- Hong, R. S., & Turner, C. W. (2009). Sequential stream segregation using temporal periodicity cues in cochlear implant recipients. *The Journal of the Acoustic Society of America*, 126, 291–299. doi:10.1121/1.3140592
- Iverson, P. (1995). Auditory stream segregation by musical timbre: Effects of static and dynamic acoustic attributes. *Journal of Experimental Psychology Human Perception & Performance*, 21, 751–763. doi:10.1121/1.403521
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82, 1–26. doi:10.18637/jss.v082.i13
- Kuznetsova, A., Christensen, R. H. B., Bavay, C., & Brockhoff, P. B. (2015). Automated mixed ANOVA modeling of sensory and consumer data. *Food Quality and Preference*, 40, 31–38. doi:10.1016/j.foodqual.2014.08.004
- Lamping, W., Santurette, S., & Marozeau, J. (2018, January). Verbal attribute magnitude estimates of pulse trains across electrode places and stimulation rates in cochlear implant listeners. *Proceedings of the International Symposium on Auditory and Audiological Research* (Vol. 6, pp. 215–222). Retrieved from <https://proceedings.isaar.eu/index.php/isaarproc/article/view/2017-26>
- Landsberger, D. M., Vermeire, K., Claes, A., Van Rompaey, V., & Van de Heyning, P. (2016). Qualities of single electrode stimulation as a function of rate and place of stimulation with a cochlear implant. *Ear and Hearing*, 37, e149–e159. doi:10.1097/AUD.0000000000000250
- Lenth, R. V. (2016). Least-squares means: The {R} package {lsmeans}. *Journal of Statistical Software*, 69, 1–33. doi:10.18637/jss.v069.i01
- Luo, X., Padilla, M., & Landsberger, D. M. (2012). Pitch contour identification with combined place and temporal cues using cochlear implants. *The Journal of the Acoustic Society of America*, 131, 1325–1336. doi:10.1121/1.3672708
- Marimuthu, V., Swanson, B. A., & Mannell, R. (2016). Cochlear implant rate pitch and melody perception as a function of place and number of electrodes. *Trends in Hearing*, 20, 1–20. doi:10.1177/2331216516643085
- Marozeau, J., Innes-Brown, H., & Blamey, P. (2013). The acoustic and perceptual cues in melody segregation for listeners with a cochlear implant. *Frontiers in Psychology*, 4, 1–11. doi:10.3389/fpsyg.2013.00790
- McKay, C. M., McDermott, H. J., & Carlyon, R. P. (2000). Place and temporal cues in pitch perception: Are they truly independent? *Acoustic Research Letters Online*, 1, 25. doi:10.1121/1.1318742
- Micheyl, C., Carlyon, R. P., Cusack, R., & Moore, B. C. J. (2005). Performance measures of auditory organization. In D. Pressnitzer, A. de Cheveigné, S. McAdams, & L. Collet (Eds.), *Auditory Signal Processing—Physiology, Psychoacoustics, and Models* (pp. 203–211). New York, NY: Springer.
- Micheyl, C., & Oxenham, A. J. (2010). Objective and subjective psychophysical measures of auditory stream integration and segregation. *Journal of the Association for Research in Otolaryngology*, 11, 709–724. doi:10.1007/s10162-010-0227-2
- Miller, G. A., & Heise, G. A. (1950). The trill threshold. *The Journal of the Acoustic Society of America*, 129, 637–638. doi:10.1121/1.1906663
- Moore, B. C. J., & Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acustica United With Acustica*, 88, 320–333. doi:10.1093/jhered/esr029
- Moore, B. C. J., & Gockel, H. E. (2012). Properties of auditory stream formation. *Philosophical Transactions of the Royal*

- Society B: Biological Science*, 367, 919–931. doi:10.1098/rstb.2011.0355
- Nelson, P. B., Jin, S. -H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *The Journal of the Acoustic Society of America*, 113, 961–968. doi:10.1121/1.1531983
- Nie, Y., & Nelson, P. (2015). Auditory stream segregation using amplitude modulated bandpass noise. *The Journal of the Acoustic Society of America*, 127, 1809. doi:10.1121/1.3384104
- Nie, Y., Zhang, Y., & Nelson, P. B. (2014). Auditory stream segregation using bandpass noises: Evidence from event-related potentials. *Frontiers in Neuroscience*, 8, 1–12. doi:10.3389/fnins.2014.00277
- Paredes-Gallardo, A., Madsen, S. M. K., Dau, T., & Marozeau, J. (2018). The role of place cues in voluntary stream segregation for cochlear implant users. *Trends in Hearing*, 22, 1–13. doi: 10.1177/2331216517750262
- Rader, T., Döge, J., Adel, Y., Weissgerber, T., & Baumann, U. (2016). Place dependent stimulation rates improve pitch perception in cochlear implantees with single-sided deafness. *Hearing Research*, 354, 109. doi:10.1016/j.heares.2017.09.009
- Roberts, B., Glasberg, B. R., & Moore, B. C. J. (2002). Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. *The Journal of the Acoustic Society of America*, 112, 2074–2085. doi:10.1121/1.1508784
- Roberts, B., Glasberg, B. R., & Moore, B. C. J. (2008). Effects of the build-up and resetting of auditory stream segregation on temporal discrimination. *Journal of Experimental Psychology Human Perception & Performance*, 34, 992–1006. doi: 10.1037/0096-1523.34.4.992
- Rose, M. M., & Moore, B. C. J. (2005). The relationship between stream segregation and frequency discrimination in normally hearing and hearing-impaired subjects. *Hearing Research*, 204, 16–28. doi:10.1016/j.heares.2004.12.004
- Sach, A. J., & Bailey, P. J. (2004). Some characteristics of auditory spatial attention revealed using rhythmic masking release. *Perception & Psychophysics*, 66, 1379–1387. doi:10.3758/BF03195005
- Shannon, R. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. *Hearing Research*, 11, 157–189. doi:10.1016/0378-5955(83)90077-1
- Singh, P. G., & Bregman, A. S. (1997). The influence of different timbre attributes on the perceptual segregation of complex-tone sequences. *The Journal of the Acoustic Society of America*, 102, 1943–1952. doi:10.1121/1.419688
- Stainsby, T. H., Fu, C., Flanagan, H. J., Waldman, S. K., & Moore, B. C. J. (2011). Sequential streaming due to manipulation of interaural time. *The Journal of the Acoustic Society of America*, 130, 904–914. doi:10.1121/1.3605540
- Stickney, G. S., Zeng, F. -G., Litovsky, R., & Assmann, P. (2004). Cochlear implant speech recognition with speech maskers. *The Journal of the Acoustic Society of America*, 116, 1081–1091. doi:10.1121/1.1772399
- Tejani, V. D., Schwartz-Leyzac, K. C., & Chatterjee, M. (2017). Sequential stream segregation in normally-hearing and cochlear-implant listeners. *The Journal of the Acoustic Society of America*, 141, 50–64. doi:10.1121/1.4973516
- Thompson, S. K., Carlyon, R. P., & Cusack, R. (2011). An objective measurement of the build-up of auditory streaming and of its modulation by attention. *Journal of Experimental Psychology Human Perception & Performance*, 37, 1253–1262. doi:10.1037/a0021925
- Tong, Y. C., Blamey, P. J., Dowell, R. C., & Clark, G. M. (1983). Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant. *The Journal of the Acoustic Society of America*, 74, 73–80. doi:10.1121/1.389620
- Tong, Y. C., & Clark, G. M. (1985). Absolute identification of electric pulse rates and electrode positions by cochlear implant patients. *The Journal of the Acoustic Society of America*, 77, 1881–1888. doi:10.1121/1.391939
- Townshend, B., Cotter, N., Van Compernelle, D., & White, R. L. (1987). Pitch perception by cochlear implant subjects. *The Journal of the Acoustic Society of America*, 82, 106–115. doi:10.1121/1.395554
- van Hoesel, R. J., & Clark, G. M. (1997). Psychophysical studies with two binaural cochlear implant subjects. *The Journal of the Acoustic Society of America*, 102, 495–507. doi:10.1121/1.419611
- Van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Eindhoven, Holland: Institute for Perceptual Research. doi: 10.1007/s13398-014-0173-7.2
- Vliegen, J., Moore, B. C. J., & Oxenham, A. J. (1999). The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task. *The Journal of the Acoustic Society of America*, 106, 938–945. doi:10.1121/1.427140
- Vliegen, J., & Oxenham, A. J. (1999). Sequential stream segregation in the absence of spectral cues. *The Journal of the Acoustic Society of America*, 105, 339–346. doi:10.1121/1.424503
- Zeng, F. G. (2002). Temporal pitch in electric hearing. *Hearing Research*, 174, 101–106. doi:10.1016/S0378-5955(02)00644-5
- Zeng, F. G., Rebscher, S., Harrison, W., Sun, X., & Feng, H. (2008). Cochlear implants: System design, integration, and evaluation. *IEEE Reviews in Biomedical Engineering*, 1, 115–142. doi:10.1109/RBME.2008.2008250.