The role of temporal cues on voluntary stream segregation in cochlear implant users

ANDREU PAREDES GALLARDO*, SARA MIAY KIM MADSEN, TORSTEN DAU, AND JEREMY MAROZEAU

Hearing Systems, Department of Electrical Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark

Cochlear implant (CI) listeners experience difficulties in complex listening scenarios, where the auditory system is required to segregate a target signal from the competing sound sources. The present study investigated segregation abilities of CI listeners as a function of temporal cues and examined whether a two-stream percept occurs instantaneously or needs time to build up. CI users participated in a detection task where a sequence of regularly presented bursts of pulses ("B") on a single electrode interleaved with an irregular sequence ("A") presented on the same electrode with a different pulse rate. The pulse rate difference and the duration of the sequences were varied between trials. In half of the trials, a delay was added to the last burst of the regular A sequence and the listeners were asked to detect this delay. As the period between consecutive B bursts was jittered, time judgments between the A and B sequences did not provide a reliable cue to perform the task such that the segregation of A and B should improve performance. The results showed that performance improved with increasing rate differences and increasing sequence duration, suggesting that CI listeners can segregate sounds based on temporal cues and that this percept builds up over time.

INTRODUCTION

The cochlear implant (CI) is probably among the most successful neural prosthesis (Zeng *et al.*, 2008), making it possible for severely hearing-impaired listeners to achieve relatively high levels of speech intelligibility in quiet environments. However, CI listeners experience difficulties when listening in complex listening situations, where the auditory system is required to segregate the target signal from other competing sounds. To understand the role of different factors and cues on the segregation process an "auditory streaming" paradigm has been proposed (e.g., Bregman, 1990; Carlyon, 2004; Moore and Gockel, 2012; Van Noorden, 1975). In this paradigm, two perceptually different sounds (A and B) are presented sequentially to the listener who might fuse them into a single stream or segregate them into two separate streams, depending on the difference between the sounds. In normal-hearing (NH) listeners, large perceptual differences between the sounds facilitate segregation, whereas small differences promote fusion (integration).

*Corresponding author: apaga@elektro.dtu.dk

Proceedings of the International Symposium on Auditory and Audiological Research (Proc. ISAAR), Vol. 6: Adaptive Processes in Hearing, August 2017, Nyborg, Denmark. Edited by S. Santurette, T. Dau, J. C.-Dalsgaard, L. Tranebjærg, T. Andersen, and T. Poulsen. The Danavox Jubilee Foundation, 2017.
 ISBN: 978-87-990013-6-1.

The duration of the sequence of A and B sounds is another important factor of the auditory streaming paradigm, since the probability of achieving a segregated percept has been reported to increase over time (Anstis and Saida, 1985; Bregman, 1990; Moore and Gockel, 2012). This phenomenon is often referred as the build-up effect.

In electric hearing, perceptual differences can be elicited by varying the place or the rate of stimulation (e.g., Landsberger *et al.*, 2016). Most of the previous studies in CI listeners assessed the role of place cues on stream segregation (Böckmann-Barthel *et al.*, 2014; Chatterjee *et al.*, 2006; Cooper and Roberts, 2009, 2007; Hong and Turner, 2006; Tejani *et al.*, 2017) and little attention has been given to the role of rate or temporal periodicity cues.

Chatterjee *et al.* (2006), Hong and Turner (2009) and Duran *et al.* (2012) assessed the effect of temporal periodicity cues on stream segregation in CI listeners. The results from these studies suggest that larger differences in the temporal envelope or pulse rate between the A and the B sounds facilitate stream segregation. Chatterjee *et al.* (2006) also reported an effect of sequence duration on stream segregation. However, only one listener participated in this preliminary experiment. These studies have presented some evidence that CI listeners can use temporal periodicity cues to segregate sounds. However, there is limited evidence about the effect of the sequence duration on stream segregation (i.e., the build-up effect) in CI listeners, a well-documented phenomenon in NH listeners.

The present study investigated the role of temporal periodicity on stream segregation in CI listeners. Streaming abilities were assessed with a temporal detection task that does not rely on direct reports of perception from the listeners. Temporal periodicity cues were induced by changing the pulse rate at a fixed cochlear location. This was done to evaluate whether CI listeners can use this cue to segregate the streams, as proposed in previous studies. The effect of sequence duration was also investigated to determine whether a two-stream percept builds up over time.

METHODS

Listeners

Seven bilateral CI listeners (one male) participated in this experiment. None of the listeners had residual hearing. 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).

Stimuli and conditions

The stimulation paradigm is illustrated in Fig. 1. A sequence of 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. The onset-to-onset interval in the B sequence was 340 ms, and a random jitter of ± 220 ms was added to the onset-to-onset in the A sequence. Consecutive A and B sounds were always separated by a minimum interval of 10 ms. In half of the trials, a small delay (Δt) was added to the last burst of the B sequence, which the listeners were asked to detect. To optimize performance, the listener needed to compare the time interval between the last two B-tones to those between previous B-tones, since time judgments between successive A and B tones was an unreliable cue for performing the task. Thus, the task became easier if the A and B sequences were segregated (Micheyl and Oxenham, 2010; Nie *et al.*, 2014; Nie and Nelson, 2015).

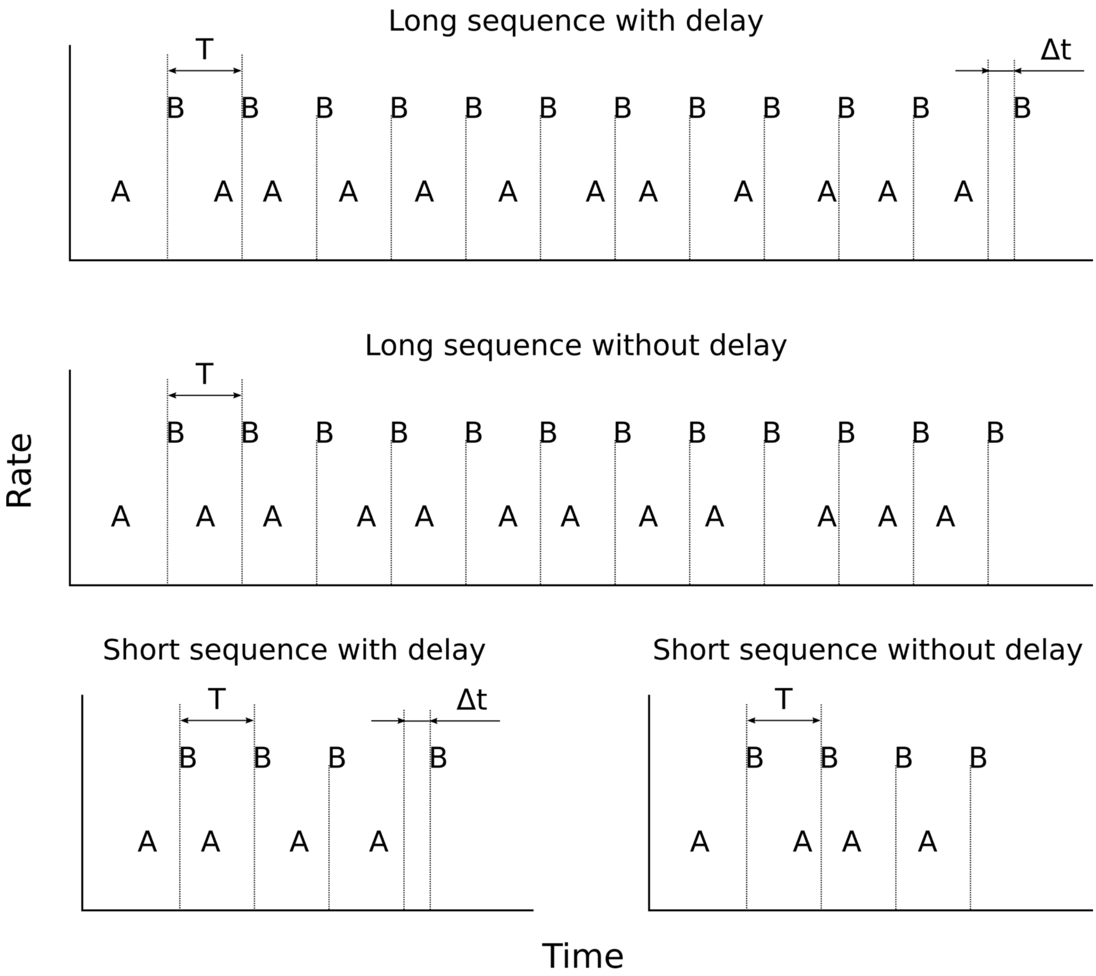

Fig. 1: Graphical representation of the experimental paradigm. The onset-to-onset interval is represented by T and the delay of the last B sound by Δt. The Δrate between A and B sounds varied across conditions.

The B sequence was played with a constant rate of 300 pps, while the A sequence was played with a pulse rate of either 80 or 260 pps, leading to a pulse rate difference (Δrate) between the streams of either 220 or 40 pps. Two sequence durations were tested. The long sequence consisted of 12 AB duplets and the short sequence of 4 AB duplets. All sequences started with the A sequence.

Each A and B sound consisted of a 50-ms biphasic pulse burst presented through electrode 11 with the corresponding rate in monopolar mode. Each biphasic pulse had a phase width of 25 µs and phase gap of 8 µs. The stimuli were presented through the Nucleus Implant Communicator research interface (NIC v2, Cochlear Limited, Sydney).

Temporal detection performance for the long and short sequences was also measured with the B sequence alone. These conditions were easier than the test conditions and, thus, a different (shorter) Δt was used to avoid ceiling effects.

For each combination of rate difference 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

Categorical loudness scaling was performed for each pulse rate in order to find the most comfortable level (MCL) for each listener. Thereafter, all stimuli were loudness matched to the 300 pps stimulus by the listener using a simple user interface (±0.15 dB).

Δt adjustment procedure

Δt were chosen such that all listeners would be equally sensitive to the delayed target in a given condition. The individual delay adjustment procedure was part of a prior study where listeners performed the temporal detection task with sequences consisting of 12 duplets of AB sounds. In that study, A and B were 50-ms bursts of pulses presented at 900 pps to electrodes 19 and 11, respectively. The sensitivity to the delayed target was measured for four different delays based on 60 presentations of each delayed sequence and 60 presentations of the non-delayed sequences. Psychometric functions were fitted to the data of each listener and the individual Δt was defined as the delay leading to *d'* = 2. Individual Δt were always smaller than the jitter applied to each A sound and ranged from 35 to 80 ms.

The same delay adjustment procedure was used to find the individual Δt to be used in the control conditions. In this case, the long sequence without distractor stream was used to fit the psychometric function. The delay leading to *d'* = 3 was chosen as Δt for the control condition. This *d'* value was chosen to keep the control conditions relatively easy while avoiding ceiling effects.

Procedure

A one-interval, two-alternative, forced-choice procedure was used, where the listeners were asked to report whether a given sequence contained a delayed target or not. A total of eight different sequences were presented to the listeners, resulting from the combination of two possible A-sequence pulse rates, two sequence durations and two different Δt (delayed or non-delayed). Short and long sequences were presented in different blocks. In each block, each of the four possible sequences was repeated 12 times in pseudo-random order, ensuring that the distractor electrode alternated from one sequence to the next one. Thus, the first sound of each sequence alternated between a pulse rate of 80 and 260 pps, contributing to reset the build-up of a two-stream percept after each presentation (Roberts *et al.*, 2008). Each block was repeated five times in a random order.

The control conditions were tested in four blocks (two with long sequences and two with short sequences) containing 30 repetitions of the delayed and 30 repetitions of the non-delayed sequences.

Statistical analysis

A mixed-effects linear model was fitted to the computed *d'* scores with the experimental factors as fixed effects terms and the listener-related effects as random effects. 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). Post-hoc analysis was performed through contrasts of least-square means. *p*-values were corrected for multiple comparisons using the Tukey method.

RESULTS AND DISCUSSION

Figure 2 shows the *d'* scores for all combinations of sequence duration and Δrate. Results from the post-hoc analysis are shown with asterisks. Both sequence duration [$F(1,18) = 27.902$, $p < 0.001$], Δrate [$F(2,18) = 13.523$, $p < 0.001$] and their interaction [$F(2,18) = 4.804$, $p < 0.021$] were found to be significant factors in the statistical model.

For the long sequence, greater *d'* scores were achieved for a Δrate of 220 pps than for a Δrate of 40 pps [$t(26.41) = 4.363$, $p = 0.002$, difference estimate = 1.436], implying that CI listeners benefitted from the larger Δrate to perform the temporal detection task. These findings are consistent with earlier work suggesting that larger differences between the temporal periodicity of the A and the B sounds facilitated a segregated percept, both in CI listeners (Chatterjee *et al.*, 2006; Duran *et al.*, 2012; Hong and Turner, 2009) and NH listeners (e.g., Grimault *et al.*, 2002; Roberts *et al.*, 2002; Vliegen *et al.*, 1999; Vliegen and Oxenham, 1999). The effect of Δrate was smaller for the short sequence [$t(26.41) = 2.194$, $p = 0.274$, difference estimate = 0.722], where no significant difference was observed between the *d'* scores achieved for the large and small Δrate conditions. Listeners achieved larger *d'* scores with the long sequence than with the short sequence for the large Δrate condition [$t(18.00) = 5.554$, $p < 0.001$, difference estimate = 1.152] but not for the small Δrate condition [$t(18.00) = 2.113$, $p = 0.324$, difference estimate = 0.428] or for the no distractor condition [$t(18.00) = 1.482$, $p = 0.679$, difference estimate = 0.307].

These results suggest that the combination of both a large Δrate and a long sequence facilitated the segregation of the streams, as reflected by the larger *d'* scores obtained for this condition. Results from the "no distractor" condition demonstrated that the sequence duration itself did not affect the temporal detection task. Thus, the difference between the d' scores achieved with the long and the short sequences, for the large Δrate condition, are likely to represent the build-up of a two stream percept.

These findings are consistent with the results from a preliminary experiment by Chatterjee *et al.* (2006) with a single CI listener despite the fact that they relied on direct reports of perception from the listener, which can be problematic with CI listeners (Cooper and Roberts, 2007; Hong and Turner, 2009). The results presented here are also consistent with the findings from Nie and Nelson (2015), who investigated the effect of amplitude modulation (AM) rate and sequence duration in NH listeners. In both studies, a significant interaction was found between AM or pulse rate and the sequence duration, suggesting that CI listeners experience a similar build-up process as NH listeners do (e.g., Anstis and Saida, 1985; Bregman, 1990; Moore and Gockel, 2012).

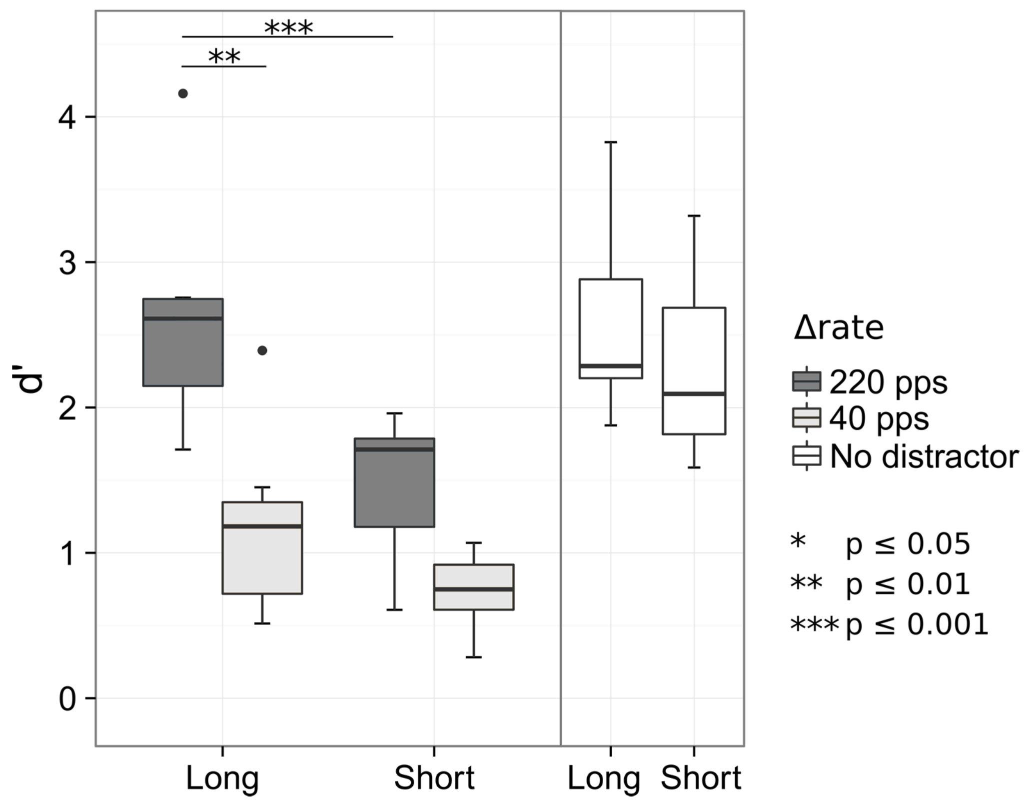

Fig. 2: Sensitivity to the delayed sound (*d'*) for each Δrate and sequence duration. The long and short sequences consisted of 12 and 4 duplets of AB sounds, respectively.

The similarity between the trends observed in NH and CI listeners supports the idea that CI listeners might experience stream segregation in a similar way as NH listeners. However, shorter Δt were needed in the "no distractor" condition to avoid ceiling effects. This reflects the increased difficulty experienced by CI listeners to perform the temporal detection task in the presence of a distractor stream, even when a large Δrate was 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 might not be able to completely ignore a competing stream.

SUMMARY AND CONCLUSIONS

The present study demonstrated that temporal periodicity cues elicited by changes in the stimulus rate can facilitate the segregation of sequential sounds for CI listeners, given that enough time is provided to build up a two-stream percept. Overall, the findings reported here are consistent with earlier work with CI and NH listeners. The similarity in the trends observed between CI and NH listeners suggests that both groups of listeners might experience stream segregation in a similar way. However, these findings are based on the results from a relatively simple task and may not be generalizable to more complex and realistic environments.

ACKNOWLEDGMENTS

Work supported by the Oticon Centre of Excellence for Hearing and Speech Sciences and the Carlsberg foundation. Research equipment provided by Cochlear.

REFERENCES

Anstis, S.M., and Saida, S. (**1985**). "Adaptation to auditory streaming of frequency-modulated tones," J. Exp. Psychol. Hum. Percept. Perform., **11**, 257-271. doi:10.1037/0096-1523.11.3.257

Bregman, A.S. (**1990**). *Auditory Scene Analysis : The Perceptual Organization of Sound.* The MIT Press.

Böckmann-Barthel, M., Deike, S., Brechmann, A., Ziese, M., and Verhey, J.L. (2014). "Time course of auditory streaming: do CI users differ from normal-hearing listeners?" Front. Psychol., **5**, 775. doi: 10.3389/fpsyg.2014.00775

Carlyon, R.P. (**2004**). "How the brain separates sounds," Trends Cogn. Sci., **8**, 465-471. doi: 10.1016/j.tics.2004.08.008

Chatterjee, M., Sarampalis, A., and Oba, S.I. (**2006**). "Auditory stream segregation with cochlear implants: A preliminary report," Hear. Res., **222**, 100-107. doi: 10.1016/j.heares.2006.09.001

Cooper, H.R., and Roberts, B. (**2007**). "Auditory stream segregation of tone sequences in cochlear implant listeners," Hear. Res., **225**, 11-24. doi: 10.1016/j.heares.2006.11.010

Cooper, H.R., and Roberts, B. (**2009**). "Auditory stream segregation in cochlear implant listeners: measures based on temporal discrimination and interleaved melody recognition," J. Acoust. Soc. Am., **126**, 1975-1987. doi: 10.1121/1.3203210

Duran, S.I., Collins, L.M., and Throckmorton, C.S. (**2012**). "Stream segregation on a single electrode as a function of pulse rate in cochlear implant listeners," J. Acoust. Soc. Am., **132**, 3849-3855. doi: 10.1121/1.4764875

Grimault, N., Bacon, S.P., and Micheyl, C. (**2002**). "Auditory stream segregation on the basis of amplitude-modulation rate," J. Acoust. Soc. Am., 111, 1340-1348. doi: 10.1121/1.1452740

Hong, R.S., and Turner, C.W. (**2006**). "Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients," J. Acoust. Soc. Am., **120**, 360-374. doi: 10.1121/1.2204450

Hong, R.S., and Turner, C.W. (**2009**). "Sequential stream segregation using temporal periodicity cues in cochlear implant recipients," J. Acoust. Soc. Am., 126, 291-299. doi: 10.1121/1.3140592

Kuznetsova, A., Christensen, R.H.B., Bavay, C., and Brockhoff, P.B. (**2015**). "Automated mixed ANOVA modeling of sensory and consumer data," Food Qual. Prefer., **40**, 31-38. doi: 10.1016/j.foodqual.2014.08.004

Landsberger, D.M., Vermeire, K., Claes, A., Van Rompaey, V., and Van de Heyning, P. (**2016**). "Qualities of single electrode stimulation as a function of rate and place of stimulation with a cochlear implant. Ear Hearing, **37**, e149-e159. doi: 10.1097/AUD.0000000000000250

Micheyl, C., and Oxenham, A.J. (**2010**). "Objective and subjective psychophysical measures of auditory stream integration and segregation," J. Assoc. Res. Otolaryngol., **11**, 709-724. doi: 10.1007/s10162-010-0227-2

Moore, B.C.J., and Gockel, H.E. (**2012**). "Properties of auditory stream formation," Philos. Trans. R. Soc. B Biol. Sci., **367**, 919-931. doi: 10.1098/rstb.2011.0355

Nie, Y., Zhang, Y., and Nelson, P.B. (**2014**). "Auditory stream segregation using bandpass noises: evidence from event-related potentials," Front. Neurosci., **8**, 1-12. doi: 10.3389/fnins.2014.00277

Nie, Y., and Nelson, P. (**2015**). "Auditory stream segregation using amplitude modulated bandpass noise," J. Acoust. Soc. Am., **127**, 1809. doi: 10.1121/1.3384104

Roberts, B., Glasberg, B.R., and Moore, B.C.J. (**2002**). "Primitive stream segregation of tone sequences without differences in fundamental frequency or passband," J. Acoust. Soc. Am., **112**, 2074-2085. doi: 10.1121/1.1508784

Roberts, B., Glasberg, B.R., and Moore, B.C.J. (**2008**). "Effects of the build-up and resetting of auditory stream segregation on temporal discrimination," J. Exp. Psychol. Hum. Percept. Perform., **34**, 992-1006. doi: 10.1037/0096-1523.34.4.992

Tejani, V.D., Schvartz-Leyzac, K.C., and Chatterjee, M. (**2017**). "Sequential stream segregation in normally-hearing and cochlear-implant listeners," J. Acoust. Soc. Am., **141**, 50-64. doi: 10.1121/1.4973516

Van Noorden, L.P.A.S. (**1975**). *Temporal Coherence in the Perception of Tone Sequences.* Institute for Perceptual Research.

Vliegen, J., Moore, B.C.J., and Oxenham, A.J. (**1999**). "The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task," J. Acoust. Soc. Am., **106**, 938-945. doi: 10.1121/1.427140

Vliegen, J., and Oxenham, A J. (**1999**). "Sequential stream segregation in the absence of spectral cues," J. Acoust. Soc. Am., **105**, 339-346. doi: 10.1121/1.424503

Zeng, F.G., Rebscher, S., Harrison, W., Sun, X., and Feng, H. (**2008**). "Cochlear implants: system design, integration, and evaluation," IEEE Rev. Biomed. Eng., **1**, 115-142. doi: 10.1109/RBME.2008.2008250