FACTORS AFFECTING STREAM SEGREGATION IN COCHLEAR IMPLANT USERS

Nicholas R. Haywood[*] **Ben Williges** **Marina Salorio-Corbetto** **Deborah A. Vickers**
Department of Clinical Neurosciences, The University of Cambridge, United Kingdom

ABSTRACT

Aspects of stream segregation in cochlear implant (CI) users remain poorly understood. In normal hearing (NH), segregation increases as the frequency separation (ΔF) between alternating tones is increased and/or the inter-stimulus interval (ISI) between tones is decreased. However, while stream segregation in CI listeners appears to be influenced by ΔF, ISI has not been found to affect segregation judgements.

In this on-going research, we asked CI listeners to report perceived segregation in stimuli where both ΔF (i.e., targeted electrode) and ISI were varied – the range of ISIs tested extend beyond those tested previously. In the preliminary dataset, all listeners show an effect of ΔF, and some do show ISI effects. A second task required listeners to detect a temporal delay imposed on a single tone. Stimuli were arranged so that any obligatory stream segregation should impair performance.

Preliminary results are varied, all listeners showed an influence of ΔF indicative of stream segregation, but the way a segregation-promoting precursor sequence affected performance varied across listeners.

We will compare these measures of stream segregation to other aspects of auditory performance – such as speech in noise.

Keywords: *cochlear implant(s), auditory stream segregation, psychophysics.*

1. INTRODUCTION

The on-going research project described in the abstract concerns the perception of auditory stream segregation in cochlear implant (CI) listeners. This paper summarizes one component of this broader project - the demonstration of stream segregation in a threshold-based task. Note that data collection is incomplete at the time of writing (n = 5).

In normal hearing (NH), a sequence of sounds that alternate between low (L) and high (H) frequency may be perceived as a single auditory object, which is commonly referred to as a "perceptual stream" (integration, Figure 1A) or as two separate streams, corresponding to the low- and high-frequency subsets (segregation, Figure 1B) [1, 2].

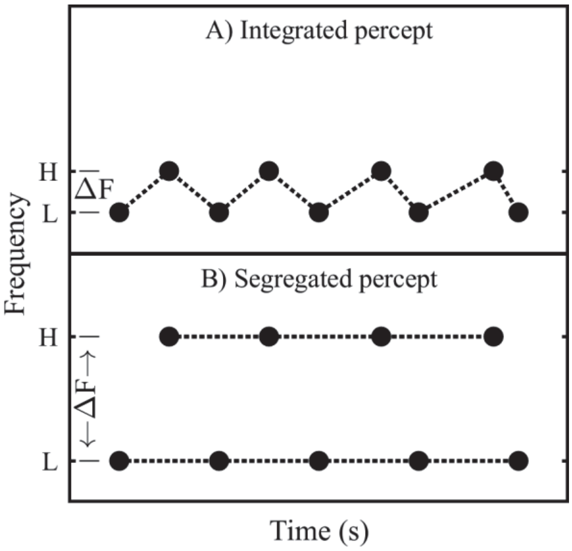

Figure 1. An illustration of steam segregation. Panel A illustrates a stimulus in which alternating low ('L') and high ('H') frequency sounds are perceptually grouped together as a single object, or stream (as illustrated by the connecting lines). Panel B illustrates a stimulus with a larger frequency separation (ΔF), in which the frequency subsets are heard as segregated into two streams. Note that in both panels, a delay is imposed on the final two H sounds. This delay is easier to detect when subsets are grouped together (i.e., Panel A).

---

[*] ***Corresponding author**: nh572@cam.ac.uk*

Stream segregation can impact performance in auditory tasks. Of relevance here is ‘temporal discrimination’, where, for example, a listener must detect a H frequency sound (or sounds) which is delayed relative to neighboring L frequency sounds. As it is difficult to compare the precise timing of sounds across different streams [3], so the detection of a delayed H sound(s) is challenging when the L & H subsets are heard as separate streams [4, 5, and see Figure 1]. As such, poor performance is considered an indication of ‘obligatory’ stream segregation – segregation that occurs even though detrimental to performance.

In NH, temporal discrimination performance corresponds well with the subjective perception of stream segregation – as both are influenced by the frequency separation (ΔF) between subsets, and the build up of segregation over time. Considering these factors in turn, firstly, subjective segregation increases and temporal discrimination performance deteriorates as the frequency difference between L and H subsets is increased [1, 2]. Secondly, the tendency to perceive segregation increases over the first several seconds of an on-going sequence [6]. This ‘build up’ is also evident in temporal discrimination measures, where performance is impaired when the target delay is preceded by a precursor/induction sequence designed to promote prior build up of segregation. Critically, a precursor impairs performance to a greater extent when ΔF is large. In this regard, temporal discrimination performance reflects the increased rate and extent of build up perceived at larger ΔFs [6].

Previous temporal discrimination studies with CI listeners have yielded mixed results. Hong & Turner [7] tested eight CI listeners with a long sequence comprising 12 LH tone cycles. A target delay was imposed progressively over the final six H tones. As such, the initial portion of this sequence can be considered a precursor, during which time segregation would be expected to build up. As would be predicted from NH, CI thresholds increased with ΔF – which in the context of CI listening can be considered as a difference in stimulating electrode(s); ΔE. A subset of three listeners were also tested with a sequence comprising a single LHL triplet (i.e., no precursor). Thresholds were overall considerably higher when the precursor was absent as opposed to present. This is opposite to what would be expected from build-up occurring during the precursor in NH, and so suggests a factor besides stream segregation influenced performance – potentially that the no precursor conditions contained a single delayed H tone, in contrast to the precursor conditions which contained six. Nonetheless, threshold elevation due to ΔE was relatively greater across the precursor conditions than across the no-precursor conditions, suggesting some degree of build-up did occur during the precursor. Cooper & Roberts [8] tested six CI listeners with similar stimuli, and like [7], observed poorer performance with increasing ΔE and poorer performance in the absence of a precursor. However, the authors found no interaction between ΔE and precursor. As no evidence for build up was found, the authors concluded that temporal discrimination performance may not reflect stream segregation processes in CI listeners (see also [9]).

Given the similar experimental approach between Hong & Turner [7] and Cooper & Roberts [8], it is challenging to reconcile their different findings. Perhaps the most straightforward conclusion is simply that the majority (6 vs. 3) of CI listeners did not show evidence for build up across these studies. Lastly, Wijetilake and colleagues [10] also measured temporal discrimination with CI listeners, but did not test for build-up effects.

Here, we re-examine ΔE and build up effects in a temporal discrimination task. All previous CI studies [7, 8, 10] tested sequences with ~40 ms silences between L and H sounds - a relatively rapid inter-stimulus interval (ISI) that is common in NH studies [4, 5, 13, 14]. We explored whether stream segregation effects in CI listeners are better observed at a longer ISI of 200 ms. We also aimed to make performance directly comparable between the no-precursor and precursor conditions, by presenting an identical ‘target’ LHL triplet in each. When present, the precursor comprised only L tones, an arrangement which has previously been shown to promote subsequent LH stream segregation in NH studies [11, 12]. Although this form of ‘same-frequency’ precursor differs from the alternating-frequency precursors described previously, both types have been shown to cause relatively greater impairment to temporal discrimination performance at larger values of ΔF in NH [13, 14]. To this end, if obligatory stream segregation does occur for CI listeners, the following should be observed: 1) performance should deteriorate with increasing ΔE, 2) the precursor should impair performance, and 3) the precursor should cause greatest impairment to performance at larger ΔEs.

2. EXPERIMENT

2.1 Listeners

Five CI listeners have taken part in the experiment to date - four with postlingual and one [CI-01] with prelingual hearing loss [CI-01 = 33 years old, CI-05 = 70, CI-12 = 73, CI10 = 67, CI14 = 61]. All used Cochlear® devices, and during the experiment wore a processor programmed with their everyday map, but with any AGCs, noise reduction and automatic scene classification (SCAN) features disabled as far as clinically possible.

2.2 Stimuli

All stimuli comprised a ‘test’ sequence of three tones in an LHL arrangement. All tones were 50-ms sinusoids with 10-ms raised-cosine ramps at onset and offset. There was a 200-ms silent interval between each tone when the H-tone was in a ‘centred’ position (Figure 2A). To minimize any potential systematic loudness cues, the level of each tone was set randomly to one uniformly distributed integer value between 61-65 dB SPL (i.e., a 4 dB ‘jitter’).

The L tones were set to the centre frequency (CF) of electrode 17, and the H tone was set to the CF of either electrode 17, 14, or 12 (ΔEs of 0, 3, or 5). These CFs correspond respectively to frequencies of 875, 1250, and 1688 Hz under default Cochlear® electrode frequency allocation. Tone frequencies were adjusted for two listeners with non-standard electrode frequency allocation due to deactivated remote electrodes [CI-01 & CI-12].

The test sequence could be preceded by a precursor sequence, which comprised four L tones, each followed by a 450-ms silence. This matched the interval between the two L tones of the test sequence, such that the six L tones across the precursor and test sequences were presented with a regular interval (Figure 2B).

Stimuli were presented acoustically via a RME Fireface UCX soundcard (Haimhausen, Germany) and Sennheiser HD-600 headphones (Hannover, Germany). Output levels were calibrated using a Tektronix MDO3024 oscilliscope (Beaverton, OR, USA).

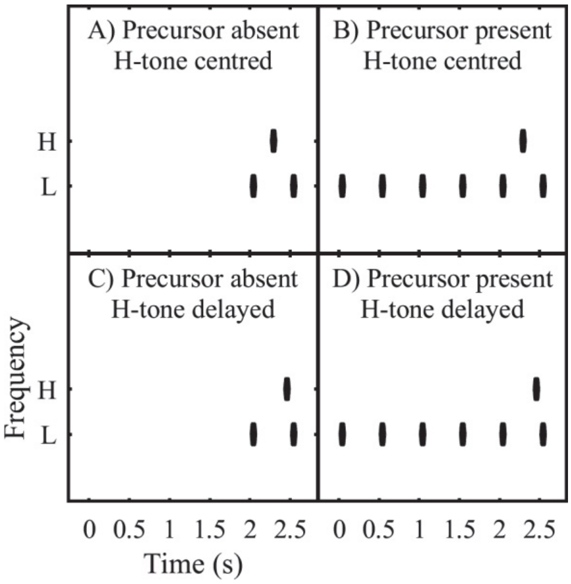

Figure 2. An illustration of the experimental stimuli. Panels A & B illustrate ‘reference’ stimuli in which the H tone is centred between surrounding L tones. Panels C & D illustrate ‘target’ stimuli in which the H tone is delayed by +200 ms from centre. Listeners were tasked to detect delayed stimuli, and the size of the delay was varied adaptively. Refer to the main text for further details.

2.3 Method

Listeners were required to discriminate a target sequence, in which the H tone was delayed (Figure 2C & 2D), from two reference sequences in which the H tone was centred (Figure 2A & 2B). In an adaptive 3I-2AFC procedure, the target interval was randomly assigned to either the 2nd or 3rd interval. There was a 1-s pause between each interval, and listeners responded via keyboard or mouse after the end of the final interval. Visual feedback was provided (‘Correct’ or ‘Incorrect’).

The size of the delay imposed on the H tone was varied adaptively, starting at +200 ms from the centred position (the maximum possible delay without causing tones to overlap). An illustration of the starting delay is shown in Figure 2C & 2D. The step size was a factor of 1.189, and the experiment used a weighted one-up, one down method. A correct response decreased the size of the delay by one step, and an incorrect response increased the size of the delay by four steps. This procedure estimates the 80% correct point on the psychometric function [15]. Each trial ran until six turn-points were observed, and the geometric mean of the final four turn-points was taken as the estimate of threshold. The experiment contained contingencies for listeners who were unable to perform the task, but all listeners tested to date could easily perform at (and near) to maximum delay.

Each adaptive run tested one of the six experimental conditions (three ΔEs: 0/3/5 electrodes × two precursor arrangements: absent/present). Listeners completed two runs for each experimental condition, and the geometric mean of the two threshold estimates was taken as their final threshold estimate. In cases where the standard deviation of the log values of those estimates exceeded 0.2, a third run took place, and the outlying estimate was removed from the average (13.33% of cases required a third run). Note that listener CI-10 is yet to complete the third runs required for two conditions (precursor present, ΔE = 3 & 5). At present, the threshold estimate is calculated from the two available runs.

2.4 Results

Both individual and group-averaged results are shown in Figure 3. From the mean data (bottom-right panel), it is apparent that thresholds increased with ΔE (i.e., the targeted electrode separation), and that thresholds were higher when the precursor was present. There also appears a trend for the precursor to cause greatest impairment at larger values of ΔE. These observations are reflected in a repeated-measures two-way ANOVA, which observes a main effect of ΔE [F(2,8) = 6.97, $p<0.05$, $\eta_p^2 = 0.63$] and precursor [F(1,4) = 14.30, $p<0.05$, $\eta_p^2 = 0.78$]. The interaction between these two factors approaches significance [F(2,8) = 4.13, $p = 0.059$, $\eta_p^2 = 0.508$]. Some caution is needed when considering these statistics, as the analysis is limited at present by the small sample size, and the fact the CI-10 has not fully completed the experiment.

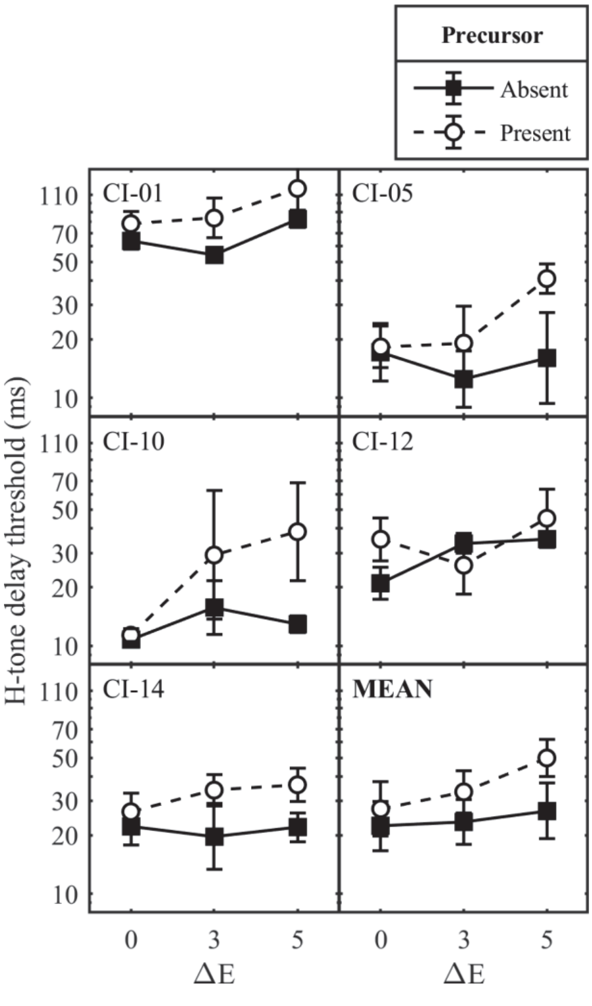

Figure 3. Experiment results. Panels displaying individual results are labelled by participant code, and error bars indicate ±1 standard deviation of two threshold estimates per condition. Subject CI-10 had not completed the experiment (refer to section 2.3 for details). The bottom-right panel shows the group mean, and error bars indicate ±1 standard error.

3. DISCUSSION

Although results may change with further testing, the emerging evidence from this experiment suggests performance may capture and reflect stream segregation processes in CI listeners. The evidence available suggests that 1) performance was poorer at larger values of ΔE, 2) the addition of a precursor impaired performance, and 3) there was a trend for the precursor to cause greatest impairment at larger ΔEs – although this effect is not statistically significant in the sample tested. If a significant interaction could be observed in a complete data set, we propose this would offer evidence that the precursor promoted subsequent segregation, and the overall pattern of results would be consistent with established stream segregation effects in NH [13, 14].

Previous CI temporal discrimination studies have not observed build-up effects in the majority of listeners tested [7, 8]. Why then, are build up effects seemingly more apparent for listeners in the present study? Firstly, the present study used a single-frequency precursor ("L-L-L'") rather than an alternating-frequency precursor ("LHLHLH"). Whilst both precursor types can promote subsequent segregation in NH, alternating-frequency precursors may be less effective in CI hearing. This possibility would require further testing to verify. Secondly, the present study maintained an identical test sequence across precursor conditions. Previous studies [7, 8] tested no-precursor conditions comprising one "LHL" triplet, and precursor conditions comprising a progressive delay to the final six H tones of an on-going "LHLH…" sequence. The fact that listeners were presented with a greater number of delayed H tones in the precursor conditions may account for the better performance in the former set of conditions. Thirdly, the current experiment presented stimuli with a considerably longer ISI (200 ms) than tested previously (40 ms). It may be that stream segregation effects are better observed at relatively slow presentation rates for CI listeners.

In NH, effects of ΔF, presentation rate, and build up on stream segregation have been attributed to adaptation and inhibition in the brainstem and auditory cortex [16, 17].

Any presentation rate differences between NH and CI could therefore reflect differences in these retrocochlear processes. To explore this question further, we are currently testing a version of this experiment with shorter ISIs of 100 and 50 ms. We are also testing a NH control group across different ISIs (200, 100, and 50 ms), and obtaining reports of subjective stream segregation in NH & CI listeners (see also [9]). From these endeavors, we hope to further quantify and understand the effect of presentation rate on stream segregation in CI listeners.

4. ACKNOWLEDGMENTS

This research was supported by an MRC Senior Fellowship in Hearing Research (MR/S002537/1). We thank our CI volunteers for their valuable time and help with this study.

5. REFERENCES

[1] G. A. Miller, and G. A. Heise: "The trill threshold," *The Journal of the Acoustical Society of America*, vol. 22, no. 5, pp. 637-638, 1950.

[2] L. P. A. S. van Noorden: *Temporal coherence in the perception of tone sequences*.' Ph.D. thesis, Eindhoven University of Technology, 1975.

[3] A. S. Bregman, and J. Campbell: "Primary auditory stream segregation and perception of order in rapid sequences of tones," *Journal of experimental psychology*, vol. 89, no. 2, pp. 244-249, 1971.

[4] J. Vliegen, B. C. J. Moore, and A. J. Oxenham: "The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task," *The Journal of the Acoustical Society of America*, vol. 106, no. 2, pp. 938-945, 1999.

[5] B. Roberts, B. R. Glasberg, and B. C. J Moore: "Primitive stream segregation of tone sequences without differences in fundamental frequency or passband," *The Journal of the acoustical society of America*, vol. 112, no. 5, pp. 2074-2085, 2002.

[6] S. Anstis, and S. Saida: "Adaptation to auditory streaming of frequency-modulated tones," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 11, no. 3, pp. 257-271, 1985.

[7] R. S. Hong, and C. W. Turner: "Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients," *The Journal of the Acoustical Society of America*, vol. 120, no. 1, pp. 360-374, 2006.

[8] H. R. Cooper, and B. Roberts: "Auditory stream segregation in cochlear implant listeners: Measures based on temporal discrimination and interleaved melody recognition," *The Journal of the Acoustical Society of America*, vol. 126, no. 4, pp. 1975-1987, 2009.

[9] H. R. Cooper, and B. Roberts: "Auditory stream segregation of tone sequences in cochlear implant listeners," *Hearing research*, vol. 225, no. 1-2, pp. 11-24, 2007.

[10] A. A. Wijetillake, R. J. van Hoesel, and R. Cowan: "Sequential stream segregation with bilateral cochlear implants," *Hearing Research*, vol. 383, article 107812, 2019.

[11] W. L. Rogers, and A. S. Bregman: "An experimental evaluation of three theories of auditory stream segregation," *Perception & psychophysics*, vol. 53, no. 2, pp. 179-189, 1993.

[12] N. R. Haywood, and B. Roberts: "Build-up of auditory stream segregation induced by tone sequences of constant or alternating frequency and the resetting effects of single deviants," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 39, no. 6, pp. 1652-1666, 2013.

[13] B. Roberts, B. R. Glasberg, and B. C. J Moore: "Effects of the build-up and resetting of auditory stream segregation on temporal discrimination," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 34, no. 4, pp. 992-1006, 2008.

[14] N. R. Haywood, and B. Roberts: "Build-up of the tendency to segregate auditory streams: resetting effects evoked by a single deviant tone," *The Journal of the Acoustical Society of America*, vol. 128, no. 5, pp. 3019-3031, 2010.

[15] C. Kaernbach: "Simple adaptive testing with the weighted up-down method," *Perception & psychophysics*, vol. 49, no. 3, pp. 227-229, 1991.

[16] D. Pressnitzer, M. Sayles, C. Micheyl, and I. M. Winter: "Perceptual organization of sound begins in the auditory periphery," *Current Biology*, vol. 18, no. 15, pp. 1124-1128, 2008.

[17] C. Micheyl, B. Tian, R. P. Carlyon, and J. P. Rauschecker: "Perceptual organization of tone sequences in the auditory cortex of awake macaques," *Neuron*, vol. 48, no. 1, pp. 139-148, 2005.