The relationship between stream segregation of complex tones and frequency selectivity

SARA M. K. MADSEN[1,*], TORSTEN DAU[1], AND BRIAN C. J. MOORE[2]

[1] *Hearing Systems group, Department of Electrical Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark*

[2] *Department of Psychology, University of Cambridge, Cambridge, UK*

The discrimination of changes in fundamental frequency (F0) is better for complex tones with low than with high harmonics, perhaps because the low harmonics are spectrally resolved. The reduced frequency selectivity of hearing-impaired (HI) participants may lead to poorer resolution of low and medium harmonics. This may adversely affect F0 discrimination and in turn reduce the extent of perceptual segregation (streaming) of a rapid sequence of complex tones. We assessed how the streaming of complex tones is affected by harmonic rank and whether HI listeners are less able to segregate tones with low and medium harmonics than near normal-hearing (NH) participants. Subjective streaming was assessed for complex tones that were bandpass filtered between 2 and 4 kHz. Harmonic rank was varied by changing the baseline F0 (with differences in F0 from 5 to 11 semitones). Auditory filter shapes were estimated from notched-noise masking using a 2-kHz signal. The auditory filters were wider for the HI than for the NH participants. Streaming decreased with increasing harmonic rank but was similar for the two groups. Streaming scores were not correlated with auditory filter bandwidths. The results suggest that the effects of harmonic rank on streaming cannot be explained in terms of resolvability.

INTRODUCTION

The extent to which a rapid sequence of tones is perceived as one stream or two streams depends on the perceptual difference between successive tones, produced, for example, by differences in frequency or level; the greater the perceptual difference, the greater is the extent of stream segregation (Moore and Gockel, 2002). When successive complex tones differ in fundamental frequency (F0), the degree of segregation may therefore depend on the salience of the F0. Complex tones with low harmonics have higher pitch salience than tones with only high harmonics and, correspondingly, F0 discrimination limens increase (performance worsens) when the rank (also called the harmonic number) of the lowest harmonic is increased above about eight (Houtsma and Smurzynski, 1990; Bernstein and Oxenham, 2006a). Both of these effects may be related to the fact that lower harmonics are better resolved

*Corresponding author: samkma@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.

than higher harmonics (Bernstein and Oxenham, 2006b). It is therefore possible that, for a given difference in F0, ΔF0, stream segregation will be greater for complex tones with low resolved harmonics than for tones with only high unresolved harmonics.
It has been shown that participants can segregate streams of complex tones with only high unresolved harmonics (Vliegen and Oxenham, 1999; Vliegen *et al.*, 1999; Grimault *et al.*, 2000; Grimault *et al.*, 2001; Madsen *et al.*, 2015). However, whereas one study found that stream segregation was similar for pure tones, complex tones with low harmonics and complex tones with only high harmonics (Vliegen and Oxenham, 1999), two other studies found a significant decrease in perceived segregation with increasing harmonic rank (Grimault *et al.*, 2000; Madsen *et al.*, 2015). However, both of these studies tested only young NH participants and the results do not make it possible to determine whether the effect of harmonic rank on stream segregation was solely an effect of resolvability; harmonic rank itself may play a role, as has been argued to be the case for the F0 discrimination of complex tones (Bernstein and Oxenham, 2003).

Here, we measured stream segregation and estimated auditory filter bandwidths using notched-noise measurements for older age-matched near-NH and HI-impaired participants. The goals were (1) to establish how the sequential stream segregation of complex tones was affected by harmonic rank and (2) to determine if the pattern of the results could be explained by the resolvability of the harmonics in the tones.

PARTICIPANTS

Eight near-NH (three male) and 13 HI (seven male) participants were tested. The NH participants were 52-78 years old (mean = 63 years, SD = 9 years) and the HI participants were 49-80 years old (mean = 67 years, SD = 9 years). The pure-tone audiometric threshold averaged across 2, 3, and 4 kHz (PTA) was required to be ≤25 dB HL for the NH participants and was between 26 and 55 dB HL for the HI participants. Audiometric thresholds for the test ear (the ear with the lower PTA) of each participant are shown in Fig. 1.

STREAM-SEGREGATION EXPERIMENT

Method

The stimuli for the stream-segregation experiment consisted of sequences of ABA_ triplets where A and B are different tones and "_" represents a pause. These sequences are perceived as having a galloping rhythm when the A and B tones are similar to each other (integration; **Fig. 2**A, upper panel) and as being two streams – with one faster than the other – if the A and B tones are more different from each other (segregation; Fig. 2A, lower panel). The ABA_ sequences had an overall duration of about 8 s. Each tone had a duration of 90 ms, and tones within a triplet were separated by a 20-ms pause. Consecutive ABA_ triplets were separated by 110-ms pauses. The tones were gated on and off with 20-ms raised-cosine ramps.

All tones contained multiple harmonics, but they were bandpass filtered between 2 and 4 kHz (Fig. 2B) to limit the audible range of the harmonics. The harmonic rank

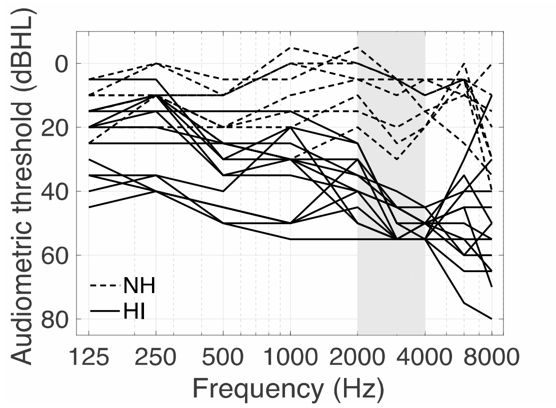

Fig. 1: Audiograms of the test ears of each NH and HI participant.

was varied by varying the F0. Four A-tone F0s were used, and the B-tone F0 was 5, 7, or 11 semitones (ST) higher than the A-tone F0. The F0s of the A- and B-tones were fixed within each trial. The overall level of each tone was 80 dB SPL, and the level of the threshold-equalising noise (TEN) used to mask combination tones and to limit the audibility of stimulus components falling in the filter skirts was 55 dB SPL/ERBn, where ERBn is the average value of the equivalent rectangular bandwidth of the auditory filter for NH listeners (Glasberg and Moore, 1990). The experiment aimed to assess the proportion of time that the streams were perceived as segregated when actively trying to segregate them. The participants were therefore asked to try to hear the sequence as segregated and to press one key when they heard one stream and a different key when they heard two streams.

Results

Participants varied markedly in the time that they took to first press the two-streams key and they rarely made more than one key press within a trial. Hence, the results are presented as the proportion of trials for which the two-streams key was the last key pressed. For brevity, we refer to this as “proportion segregated”. This measure is similar to the measure used in other studies (Vliegen and Oxenham, 1999; Grimault *et al.*, 2000; Grimault *et al.*, 2001). Increasing the F0 difference between the A and B tones increased the proportion segregated (Fig. 3). Also, generally, the proportion segregated increased with increasing F0 (i.e., with decreasing harmonic rank) as hypothesized. However, the results were similar for the NH and HI participants.

The streaming scores were averaged across ΔF0 values for each A-tone F0 and were analysed with a mixed linear model with harmonic rank and participant group (NH or HI) as fixed factors and participants as a random factor. There was a significant effect of harmonic rank [$F(3,60) = 6.46$, $p < 0.001$] but no effect of participant group [$F(1,19) = 0.11$, $p = 0.74$] and no interaction [$F(3,57) = 1.21$, $p = 0.31$].

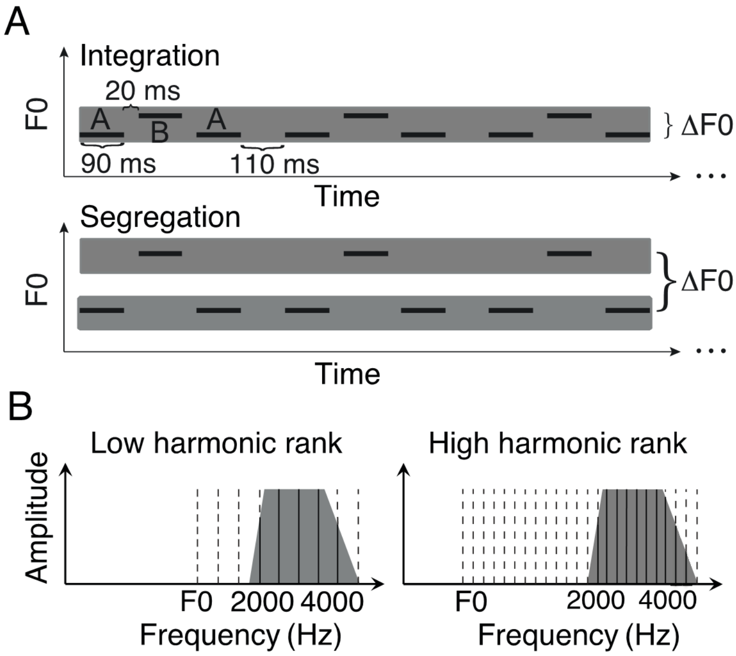

Fig. 2: Illustration of the stimuli used in the stream-segregation experiment. A: Illustration of the ABA_ tone sequences. Upper panel: When A and B were sufficiently similar to each other, all tones were perceived as being in one stream with a galloping rhythm (integration). Lower panel: When A and B were more different from each other, they were perceived as two separate streams. B: Schematic spectrum of the complex tones. Tones were filtered between 2 and 4 kHz, and the harmonic rank was varied by varying the F0.

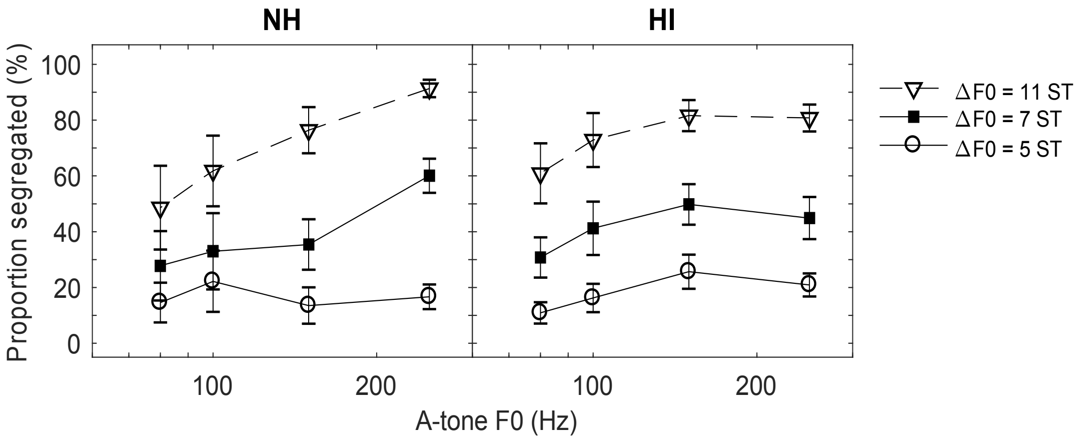

Fig. 3: Mean proportion segregated for the NH (left) and HI (right) participants as a function of the A-tone F0. The parameter is ΔF0, as indicated in the key. Error bars show ±1 standard error of the mean.

NOTCHED-NOISE EXPERIMENT

Regardless of whether stream segregation for complex tones depends on the resolvability of the harmonics or on harmonic rank *per se*, it seems reasonable to assume that performance in the stream-segregation task was dominated by the lowest-ranking harmonics in the complex tones. Therefore, auditory filter shapes were estimated at 2 kHz, the lower edge of the bandpass filter used in the stream-segregation experiment.

Method

The notched-noise method (Patterson, 1976; Rosen *et al.*, 1998) was used. The participants were asked to identify which of three successive noise bursts contained a 2-kHz pure tone. The noise had a duration of 500 ms, and the 400-ms pure tone was temporally centred in one randomly chosen noise burst. Raised-cosine ramps of 10 and 20 ms, respectively, were used to gate the noise and pure tone on and off. The signal level was fixed at 10 dB sensation level while the level of the noise was varied adaptively using a 2-up 1-down procedure. The outside edges of the noise were fixed at 400 and 3600 Hz. The notch width is specified as the deviation of each edge of the notch from the signal frequency, divided by the signal frequency. There were five symmetric conditions with notch widths of 0, 0.1, 0.2, 0.3, and 0.4, and two asymmetric conditions with notch widths of 0.2|0.4 and 0.4|0.2.

Results

A two-parameter roex(p_u, p_l) filter model (Stone *et al.*, 1992) was used to estimate the parameters p_u and p_l, which characterise the slopes of the upper and lower filter skirts, respectively. The equivalent rectangular bandwidth (ERB) of the auditory filter in Hz was estimated as $8*2000/(p_u + p_l)$ (Patterson *et al.*, 1982). As expected, the estimated ERB values were generally smaller for the NH than for the HI participants (Fig. 4). This difference was confirmed by a Welch's *t*-test [$t(14.31) = -2.38$, $p = 0.032$] and was even more pronounced when the ERB value for one NH participant with a very high ERB value was omitted [$t(16.52) = -4.74$, $p < 0.001$]. Also, results for the streaming experiment were similar, if omitting this participant [harmonic rank: $F(3,57) = 5.64$, $p = 0.002$; participant group: $F(1,18) = 0.0023$, $p = 0.96$].

DISCUSSION

While the HI participants had significantly broader auditory filters (greater ERB values) than the NH participants, stream segregation was similar for the two groups. This suggests that the effect of harmonic rank on stream segregation cannot be explained by differences in resolvability of the harmonics. If resolvability was critical, the proportion segregated should have been higher for the NH than the HI participants, at least for conditions where the lowest harmonics were only just resolved for the NH group, since the harmonics for these conditions would have been largely unresolved for the HI group. However, the ERB values varied markedly across participants, especially for the HI group. To further explore whether there was a relationship

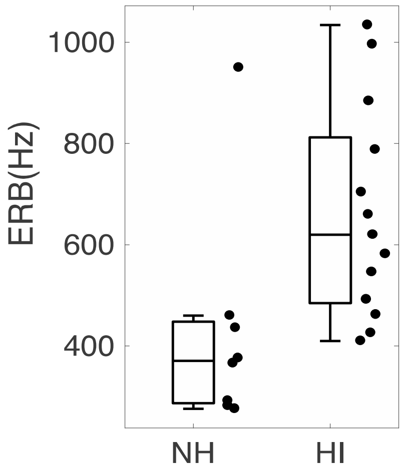

Fig. 4: Box plots of ERB values for each participant group.

between stream segregation and the resolvability of the harmonics, the average of the proportion segregated values across the conditions with the two highest F0s (250 and 150 Hz) was compared with the ERB value for each participant. If better resolvability leads to greater stream segregation, a negative correlation between these two measures would be expected. In fact, the two measures were not correlated (R^2 = 0.043, p = 0.37). A scatter plot of the proportion segregated values against the ERB values is shown in Fig. 5. This supports our conclusion that the effects of harmonic rank on the stream segregation of complex tones cannot be explained in terms of resolvability.

The finding that the stream segregation of complex tones is affected by harmonic rank is not consistent with the results of one of the three previous studies on this topic (Vliegen and Oxenham, 1999). However, Vliegen and Oxenham (1999) also presented some data from a preliminary experiment that, for some conditions and participants, showed results similar to those of the present study.

The conclusion that the effect of harmonic rank cannot be explained by resolvability contrasts with the conclusions of two earlier studies (Grimault *et al.*, 2000; Grimault *et al.*, 2001). The latter of these tested three groups: young NH, older with normal hearing for their age, and older HI. There was little difference in stream segregation between the two groups of older participants, consistent with the findings of the present study. However, they generally found more stream segregation for the young NH group than for the two other groups. This might have been due to differences in auditory filter bandwidth but is more likely to be an effect of age unrelated to frequency selectivity. Therefore, the results of Grimault *et al.* (2001) cannot be used to draw any firm conclusions about the relationship between the stream segregation of complex tones and the resolvability of the harmonics in the tones.

If it is accepted that the resolvability of the harmonics is not the critical factor governing the stream segregation of complex tones, then it appears that harmonic rank *per se* has an influence. This is consistent with the results of a study showing that F0

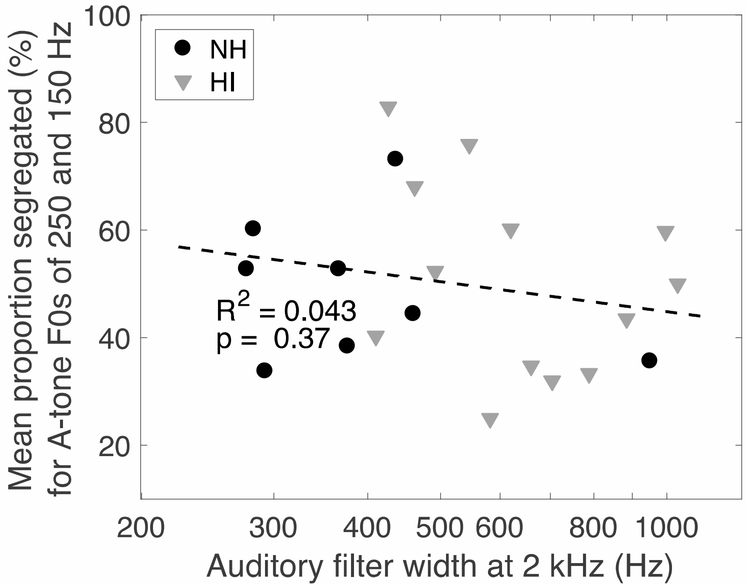

Fig. 5: Scatter plot of mean proportion segregated values across conditions with A-tone F0s of 250 and 150 Hz against ERB values at 2 kHz. The regression line and statistics were calculated from the data for all participants.

discrimination limens were similar for conditions where all harmonics were presented to both ears (diotic) and where odd harmonics were presented to one ear and even harmonics to the other ear (dichotic) (Bernstein and Oxenham, 2003). There were some conditions of this experiment where the harmonics would have been unresolved for diotic presentation but would have been resolved for dichotic presentation, because of the greater spacing of the harmonics within each ear. The lack of effect of presentation mode suggests that F0 discrimination is governed by harmonic rank and not by the resolvability of the harmonics. The effect of harmonic rank has been explained by ‘place dependence’, i.e., for each auditory filter there is a limited range of periodicities that can be analysed, and this range is closely tied to the centre frequency of that filter (Moore, 2003; Bernstein and Oxenham, 2005).

CONCLUSIONS

For both older near-NH and older HI participants, the proportion segregated in a sequence of complex tones increased with decreasing harmonic rank (increasing F0). Furthermore, the proportion segregated varied with harmonic rank in a similar way for the two groups. However, auditory filter bandwidths at 2 kHz estimated using the notched-noise method were, on average, greater for the HI than for the NH group. Also, the proportion segregated scores were not correlated with the auditory filter bandwidth estimates. These findings suggest that the effect of harmonic rank on stream segregation cannot be explained by the better resolution of lower than of higher harmonics, but rather reflects an effect of harmonic rank *per se*.

REFERENCES

Bernstein, J.G., and Oxenham, A.J. (**2003**). "Pitch discrimination of diotic and dichotic tone complexes: harmonic resolvability or harmonic number?," J. Acoust. Soc. Am., **113**, 3323-3334. doi: 10.1121/1.1572146

Bernstein, J.G., and Oxenham, A.J. (**2005**). "An autocorrelation model with place dependence to account for the effect of harmonic number on fundamental frequency discrimination," J. Acoust. Soc. Am., **117**, 3816-3831. doi: 10.1121/1.1904268

Bernstein, J.G., and Oxenham, A.J. (**2006a**). "The relationship between frequency selectivity and pitch discrimination: effects of stimulus level," J. Acoust. Soc. Am., **120**, 3916-3928. doi: 10.1121/1.2372451

Bernstein, J.G., and Oxenham, A.J. (**2006b**). "The relationship between frequency selectivity and pitch discrimination: sensorineural hearing loss," J. Acoust. Soc. Am., **120**, 3929-3945. doi: 10.1121/1.2372452

Glasberg, B.R., and Moore, B.C.J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res., **47**, 103-138.

Grimault, N., Micheyl, C., Carlyon, R.P., Arthaud, P., and Collet, L. (**2000**). "Influence of peripheral resolvability on the perceptual segregation of harmonic complex tones differing in fundamental frequency," J. Acoust. Soc. Am., **108**, 263-271. doi: 10.1121/1.429462

Grimault, N., Micheyl, C., Carlyon, R.P., Arthaud, P., and Collet, L. (**2001**). "Perceptual auditory stream segregation of sequences of complex sounds in subjects with normal and impaired hearing," Br. J. Audiol., **35**, 173-182.

Houtsma, A.J.M., and Smurzynski, J. (**1990**). "Pitch identification and discrimination for complex tones with many harmonics," J. Acoust. Soc. Am., **87**, 304-310. doi: 10.1121/1.399297

Madsen, S.M.K., Dau, T., and Moore, B.C.J. (**2015**). "Effect of harmonic rank on the streaming of complex tones," Proc. ISAAR, **5**, 477-483.

Moore, B.C.J., and Gockel, H. (**2002**). "Factors influencing sequential stream segregation," Acta Acust., **88**, 320-333.

Moore, B.C.J. (**2003**). *An Introduction to the Psychology of Hearing, 5th Ed.* (Emerald, Bingley, UK), pp. 413.

Patterson, R.D. (**1976**). "Auditory filter shapes derived with noise stimuli," J. Acoust. Soc. Am., **59**, 640-654. doi: 10.1121/1.380914

Patterson, R.D., Nimmo-Smith, I., Weber, D.L., and Milroy, R. (**1982**). "The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold," J. Acoust. Soc. Am., **72**, 1788-1803.

Rosen, S., Baker, R.J., and Darling, A. (**1998**). "Auditory filter nonlinearity at 2 kHz in normal hearing listeners," J. Acoust. Soc. Am., **103**, 2539-2550. doi: 10.1121/1.422775

Stone, M.A., Glasberg, B.R., and Moore, B.C.J. (**1992**). "Simplified measurement of impaired auditory filter shapes using the notched-noise method," Br. J. Audiol. **26**, 329-334. doi: 10.3109/03005369209076655

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.

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