



## Effect of harmonic rank on sequential sound segregation

Madsen, Sara Miay Kim; Dau, Torsten; Moore, Brian C.J.

*Published in:*  
Hearing Research

*Link to article, DOI:*  
[10.1016/j.heares.2018.06.002](https://doi.org/10.1016/j.heares.2018.06.002)

*Publication date:*  
2018

*Document Version*  
Peer reviewed version

[Link back to DTU Orbit](#)

*Citation (APA):*  
Madsen, S. M. K., Dau, T., & Moore, B. C. J. (2018). Effect of harmonic rank on sequential sound segregation. *Hearing Research*, 367, 161-168. <https://doi.org/10.1016/j.heares.2018.06.002>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Accepted Manuscript

Effect of harmonic rank on sequential sound segregation

Sara M.K. Madsen, Torsten Dau, Brian C.J. Moore

PII: S0378-5955(18)30115-1

DOI: [10.1016/j.heares.2018.06.002](https://doi.org/10.1016/j.heares.2018.06.002)

Reference: HEARES 7570

To appear in: *Hearing Research*

Received Date: 24 March 2018

Revised Date: 1 June 2018

Accepted Date: 8 June 2018



Please cite this article as: Madsen, S.M.K., Dau, T., Moore, B.C.J., Effect of harmonic rank on sequential sound segregation, *Hearing Research* (2018), doi: 10.1016/j.heares.2018.06.002.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1 **Effect of harmonic rank on sequential sound segregation**

2 Sara M.K. Madsen<sup>1,\*</sup>, Torsten Dau<sup>1</sup>, Brian C.J. Moore<sup>2</sup>

3

4 <sup>1</sup> *Hearing Systems Group, Department of Electrical Engineering, Technical University of*  
5 *Denmark, DK-2800 Lyngby, Denmark*

6 <sup>2</sup> *Department of Psychology, University of Cambridge, Cambridge, UK*

7

8

9 \* Corresponding author. Hearing Systems Group, Department of Electrical Engineering,  
10 Technical University of Denmark, DK-2800 Lyngby, Denmark.

11 *E-mail address:* samkma@elektro.dtu.dk (S.M.K. Madsen)

12

13 Revised version submitted June 2018

14

15 *Keywords:* Stream segregation, Fundamental frequency, Fundamental frequency  
16 discrimination

17

18

## 19 ABSTRACT

20 The ability to segregate sounds from different sound sources is thought to depend on the  
21 perceptual salience of differences between the sounds, such as differences in frequency or  
22 fundamental frequency (F0). F0 discrimination of complex tones is better for tones with low  
23 harmonics than for tones that only contain high harmonics, suggesting greater pitch salience  
24 for the former. This leads to the expectation that the sequential stream segregation (streaming)  
25 of complex tones should be better for tones with low harmonics than for tones with only high  
26 harmonics. However, the results of previous studies are conflicting about whether this is the  
27 case. The goals of this study were to determine the effect of harmonic rank on streaming and  
28 to establish whether streaming is related to F0 discrimination. Thirteen young normal-hearing  
29 participants were tested. Streaming was assessed for pure tones and complex tones containing  
30 harmonics with various ranks using sequences of ABA triplets, where A and B differed in  
31 frequency or in F0. The participants were asked to try to hear two streams and to indicate  
32 when they heard one and when they heard two streams. F0 discrimination was measured for  
33 the same tones that were used as A tones in the streaming experiment. Both streaming and F0  
34 discrimination worsened significantly with increasing harmonic rank. There was a significant  
35 relationship between streaming and F0 discrimination, indicating that good F0 discrimination  
36 is associated with good streaming. This supports the idea that the extent of stream segregation  
37 depends on the salience of the perceptual difference between successive sounds.

38

39 Keywords: stream segregation, pitch, fundamental frequency discrimination, perceptual  
40 differences

41

## 42 **1 Introduction**

43 The ability to segregate sounds from different sound sources is thought to depend on the  
44 perceptual salience of differences between the sounds, such as differences in frequency or  
45 fundamental frequency (F0) (Moore and Gockel, 2002; Paredes-Gallardo et al., 2018). It is  
46 therefore easier to understand speech produced by a female speaker in the presence of one or  
47 more male speakers than when in the presence of other female speakers (Brungart et al.,  
48 2001). The ability to segregate sounds into different auditory objects is used constantly in  
49 daily life and is essential for understanding speech in the presence of background sounds. The  
50 ability also makes it possible to hear out individual instruments or voices in music.

51 Speech and music are complex signals and sound segregation is often investigated  
52 using simpler, more controlled, stimuli such as sequences of interleaved A and B sounds  
53 where A and B differ in some way (e.g., Bregman, 1990; van Noorden, 1975). These  
54 sequences can be heard either as one stream (integrated) or as two streams (segregated). The  
55 perceptual construction of two streams is called sequential stream segregation or streaming.  
56 Several studies have used such sequences to explore the effect of differences between the A  
57 and B sounds in frequency (pure tones) or in F0 (complex tones) and have shown that the  
58 ability to segregate increases with increasing frequency or F0 difference (e.g., Grimault et al.,  
59 2000; Grimault et al., 2001; Rose and Moore, 1997; van Noorden, 1975; Vliegen and  
60 Oxenham, 1999; Vliegen et al., 1999).

61 Studies of F0 discrimination have shown that F0 difference limens (F0DLs) are  
62 relatively small when the tones contain low harmonics (with harmonic numbers, also called  
63 ranks, up to about 8), but increase when the rank of the lowest harmonic increases above  
64 about 8, indicating that pitch salience decreases when only high-rank harmonics are present  
65 (e.g., Bernstein and Oxenham, 2006a; Hoekstra and Ritsma, 1977; Houtsma and Smurzynski,  
66 1990; Shackleton and Carlyon, 1994). The increase in F0DLs with increasing harmonic rank

67 might be explained by better resolution of lower than of higher harmonics (Bernstein and  
68 Oxenham, 2006b; Shackleton and Carlyon, 1994). However, some lines of evidence suggest  
69 that resolution of harmonics is not the key factor. Firstly, for very low F0s, the harmonics that  
70 dominate the pitch percept are not the lowest resolved harmonics (Jackson and Moore, 2013).  
71 Secondly, Bernstein and Oxenham (2003) compared F0DLs for tones with all harmonics  
72 presented to both ears (diotic) and tones with odd harmonics presented to one ear and even  
73 harmonics to the opposite ear (dichotic). If F0 discrimination were governed by the degree of  
74 resolution of the harmonics, performance should have been better for the dichotic condition,  
75 since the frequency separation of harmonics within each ear was twice as large as for the  
76 diotic condition. In fact, F0DLs were similar for the diotic and dichotic conditions. The  
77 results suggest that harmonic rank *per se* is important. The effect of harmonic rank has been  
78 explained by ‘place dependence’, i.e. for each place in the cochlea (corresponding to a  
79 specific auditory filter with a certain center frequency) there is a limited range of periodicities  
80 that can be analyzed, and this range is closely tied to the center frequency of that filter  
81 (Bernstein and Oxenham, 2005; Moore, 2003).

82         If stream segregation depends on the salience of the perceptual differences between  
83 successive sounds (Hartmann and Johnson, 1991; Moore and Gockel, 2002; Paredes-Gallardo  
84 et al., 2018), one might expect that the ease with which a sequence of complex tones (tones A  
85 and B, differing in F0) can be segregated into streams would be affected by pitch salience  
86 (strength). If so, then for a fixed difference in F0 between successive tones, stream  
87 segregation should be more likely to occur for tones containing low harmonics than for tones  
88 containing only high harmonics. A few studies have investigated the effect of harmonic rank  
89 on streaming, but with differing results. Vliegen and Oxenham (1999) measured sequential  
90 stream segregation for pure tones, complex tones with low harmonics, and complex tones  
91 with only high harmonics. For each of these, the F0 of the B tone was between one and 11

92 semitones higher than the F0 of the fixed A tone. The listeners were instructed to try to hear  
93 the sequence as segregated and to indicate whether they heard each sequence as one or two  
94 streams. The proportion of trials that were perceived as segregated was similar for all  
95 conditions, indicating no effect of harmonic rank. Grimault et al. (2000) measured streaming  
96 for complex tones with fixed F0s for the A and B tones. The tones were filtered into three  
97 regions (low, mid, and high) to vary the ranks of the harmonics in the tones. They found that  
98 the percentage of segregation decreased with increasing harmonic rank and argued that this  
99 was an effect of the resolvability of the harmonics in the tones. They did not instruct the  
100 listeners to try to hear the streams as segregated or integrated, as in the study of Vliegen and  
101 Oxenham (1999). The instruction to try to segregate used by Vliegen and Oxenham might  
102 have increased the proportion of segregation, especially for the difficult conditions with only  
103 high harmonics, and Grimault et al. (2000) suggested that the difference in instructions might  
104 explain the difference between studies. Also, they proposed that the difference across studies  
105 might be explained by their conditions being more extreme in terms of resolvability than the  
106 ones used by Vliegen and Oxenham (1999). If so, this would indicate that large differences in  
107 harmonic rank are required to reveal differences in stream segregation.

108 The aims of the present study were: (1) to determine the effect of harmonic rank when the  
109 listeners were instructed to try to hear the sequence as segregated; 2) to establish whether  
110 there is a relationship between F0DLs and streaming. Sequential stream segregation was  
111 investigated for pure tones and complex tones with harmonic ranks ranging from low (with  
112 well resolved harmonics) to high (with all harmonics clearly unresolved), i.e. representing  
113 conditions with harmonic rank less than 8 or larger than 14, respectively (Moore and Gockel,  
114 2011). Preliminary data from this study was previously presented in a conference paper  
115 (Madsen et al., 2015).

116

## 117 2 General method

### 118 2.1 Listeners

119 Thirteen normal-hearing listeners (audiometric thresholds  $\leq 20$  dB HL at octave  
120 frequencies between 250 and 8000 Hz; five females, eight males) between 21 and 27 years of  
121 age (mean = 23.6 years, SD = 1.6 years) were tested. The listeners had no musical training.  
122 All experiments were approved by the Science-Ethics Committee for the Capital Region of  
123 Denmark.

124

### 125 2.2 Stimulus generation and presentation

126 The stimuli were generated in MATLAB at a sampling rate of 44100 and presented via a  
127 Fireface UCX sound card (RME, Haimhausen Germany) and Sennheiser HD 650 headphones  
128 (Sennheiser, Wedemark, Germany). All stimuli were presented monaurally at a sound  
129 pressure level (SPL) of 80 dB to the ear with the lowest audiometric threshold averaged  
130 across the frequencies 2, 3, and 4 kHz. This level was chosen since this study was meant to be  
131 the first in a series of experiments in which hearing-impaired listeners would also be tested.  
132 This allows the comparison of results for normal-hearing and hearing-impaired listeners at the  
133 same sound pressure level. All measurements were made in an acoustically shielded booth.

134

## 135 3 Experiment 1: Sequential stream segregation

### 136 3.1 Rationale

137 The goal of this experiment was to determine whether subjective sequential stream  
138 segregation is affected by harmonic rank. F0 discrimination is better for tones with low  
139 harmonic rank and it was therefore hypothesized that the presence of low harmonics would  
140 facilitate the segregation of sequences of complex tones.

141

### 142 3.2 Method

143

144 [Insert Fig. 1 approximately here]

145

146 *3.2.1 Stimuli*

147 The stimuli consisted of sequences of ABA-ABA tones where A and B are different tones  
148 and “-” represents a brief pause. This type of stimulus has been used in many experiments on  
149 stream segregation (e.g., Bregman, 1990; van Noorden, 1975). As illustrated in Fig. 1A, such  
150 a sequence can be perceived as one stream (upper panel; integration) that is heard as having a  
151 galloping rhythm or as two separate streams, one twice as fast as the other (lower panel;  
152 segregation). As in the study of Vliegen and Oxenham (1999), each tone had a duration of 90  
153 ms including 20-ms raised-cosine ramps. The time interval between tones within each triplet  
154 was 10 ms and consecutive triplets were separated by 110 ms. Each tone sequence consisted  
155 of 19 triplets and had a duration of approximately 8 s.

156 Both the A and B tones were either complex tones or pure tones. As illustrated in Fig.  
157 1B, the complex tones were initially generated to contain all harmonics with equal amplitude,  
158 added in sine phase. The tones were then bandpass filtered between 2 and 4 kHz (3-dB down  
159 points), using a filter slope of 30 dB/octave for the first 100 Hz on each side of the flat  
160 passband and 50 dB/octave beyond that range. The edge frequencies of the passband were  
161 2125 and 3798 Hz. The filter slope was chosen to avoid abrupt changes in level of individual  
162 harmonics as they passed into and out of the passband when the F0 was changed. The  
163 harmonic rank was varied by varying the F0; the higher the F0 the lower was the harmonic  
164 rank. For the pure-tone stimuli, the frequency of the A tone was 2000 Hz. For the complex  
165 tones, the A-tone F0 was 80, 100, 150, 250, or 500 Hz. Hence, the rank of the lowest  
166 harmonic in the passband varied from 27 (F0 = 80 Hz) to 5 (F0 = 500 Hz). The B-tone  
167 frequency or F0 was always higher than that of the A tone. The frequency or F0 difference  
168 between the A and B tones ( $\Delta F_0$ ) was 1, 3, 4, 5, 7, or 11 semitones (ST), resulting in 36  
169 conditions. The frequencies or the F0s of the A and B tones were fixed within each trial.

170 A threshold-equalizing noise (TEN) (Moore et al., 2000) was used to mask combination  
171 tones and to limit the audibility of stimulus components falling on the filter skirts. According  
172 to Oxenham et al. (2009) the  $2f_1$ - $f_2$  combination tone produced by interaction of the two  
173 lowest components in the passband may just be audible when the component level is 15 dB

174 higher than the TEN level, expressed as dB SPL/ERB<sub>N</sub>, where ERB<sub>N</sub> is the average value of  
175 the equivalent rectangular bandwidth of the auditory filter for listeners with normal hearing  
176 (Glasberg and Moore, 1990). The component level needs to be about 30 dB higher than the  
177 TEN level for the next lower combination tone to be audible. The present study used a TEN  
178 level of 55 dB SPL/ERB<sub>N</sub>, which meant that the level of each component in the complex  
179 tones was 20-24 dB higher than the level/ERB<sub>N</sub> of the TEN. Hence, the 2f<sub>1</sub>-f<sub>2</sub> combination  
180 tone may have been just audible, but no lower combination tones were audible. This does not  
181 create a problem in the interpretation of the results presented here, since the only consequence  
182 of the 2f<sub>1</sub>-f<sub>2</sub> combination tone being audible would be to lower the harmonic rank by one.  
183 This would not affect whether the tones in the different conditions were resolved or  
184 unresolved.

185

### 186 3.2.2 Procedure

187 The aim was to assess the proportion of time that two streams were perceived when  
188 listeners were actively trying to segregate the sequence. The listeners were therefore asked to  
189 try to hear the sequence as segregated and to press one key when they heard one stream and a  
190 different key when they heard two streams. They could switch between the two keys during  
191 presentation of a sequence if the percept appeared to change. The listeners were trained for at  
192 least two hours and tested in four 2-hour sessions. Each condition was tested 36 times for  
193 each listener in blocks that each contained one presentation of each condition. The conditions  
194 were randomized such that the order of conditions within a block was always different across  
195 blocks for each listener. The order was different for each listener. Nine blocks were tested in  
196 each session.

197 To ensure that the listeners had been sufficiently trained, the standard deviation of the  
198 streaming scores (percentage of time that two streams were reported) for each condition was  
199 calculated across each set of three successive blocks and then averaged across conditions. If  
200 the mean standard deviation was larger than or equal to 20% for at least one of the three sets  
201 of three blocks tested in the first test session, these blocks were considered as training and  
202 they were repeated in the following session.

203

204 *3.2.3 Statistical analysis*

205 Due to large deviations from normality, the data were transformed using the aligned rank  
206 transform (Wobbrock et al., 2011) and then analyzed with a linear mixed-effects model with  
207 harmonic rank and  $\Delta F0$  as fixed factors and listener as a random factor, using the ARTool  
208 library (Wobbrock et al., 2011) in R. Post-hoc analysis was performed using the  
209 lsmeans library (Lenth, 2016) and Tukey corrections were used to correct for multiple  
210 comparisons.

211

212 *3.3 Results*

213 Subjective sequential stream segregation was assessed as the proportion of time that the  
214 listeners indicated hearing two streams (no galloping rhythm), assessed over the whole  
215 duration of the sequence. Figure 2 shows the individual data and Fig. 3 shows the mean data.  
216 Results for the complex tones are plotted on the left as a function of  $F0$  and results for the  
217 pure tones are plotted on the right. While there were large individual differences, the  
218 streaming scores generally increased with increasing  $\Delta F0$  or  $\Delta F$  (pure tones) and with  
219 increasing  $F0$ , i.e. decreasing harmonic rank. All conditions, including the pure tone  
220 conditions, were included in the analysis. Both main effects were significant ( $\Delta F0$ :  $F(5, 420)$   
221  $= 142.77, p < 0.001$ ; harmonic rank:  $F(5, 420) = 205.34, p < 0.001$ ) and the interaction was  
222 also significant ( $F(25, 420) = 9.96, p < 0.001$ ). Pairwise comparison of conditions with  
223 different  $F0$  (harmonic rank) showed that the differences between all pairs were significant ( $p$   
224  $< 0.01$ ) except between  $F0 = 500$  and  $250$  Hz. Similarly, pairwise comparison of conditions  
225 with different  $\Delta F0$  showed that all differences were significant ( $p < 0.01$ ) except between  $\Delta F0$   
226  $= 4$  and  $5$  ST.

227

228 *3.4 Discussion*

229 It is possible that the listeners judged the perceptual difference between the A and B  
230 tones rather than judging stream segregation per se. To assess this possibility, it was  
231 determined whether a build up of “two-stream” responses occurred over time, since build up

232 is generally regarded as a key characteristic of stream segregation (e.g., Anstis and Saida,  
233 1985; Bregman, 1978b). This was done for a condition leading to an intermediate percentage  
234 of two-stream responses ( $\Delta F0 = 7$  and  $F0 = 250$  Hz) to avoid floor and ceiling effects. Figure  
235 4 shows the percentage of time where the two-stream key was pressed after every half second  
236 for each of the listeners. As expected, the proportion of two-stream responses increased with  
237 time for most listeners, confirming that judgements were based on stream segregation rather  
238 than on the perceptual difference between the A and B tones.

239

240 [Please insert Fig. 4 approximately here]

241

242 The significant increase in segregation with increasing  $\Delta F0$  is consistent with the results of  
243 many other studies (e.g., Grimault et al., 2000; Grimault et al., 2001; Rose and Moore, 1997;  
244 van Noorden, 1975; Vliegen and Oxenham, 1999; Vliegen et al., 1999) and with the idea that  
245 the extent of stream segregation increases with increasing perceptual difference between  
246 successive sounds (Moore and Gockel, 2002). This idea is also supported by the decrease in  
247 stream segregation with increasing harmonic rank. The harmonic rank was varied by varying  
248 the  $F0$ . In theory, therefore, the observed effects could be a result of variations in  $F0$  rather  
249 than variations in harmonic rank. However, this seems unlikely, since F0DLs for sounds with  
250 fixed harmonic content (e.g. harmonics 1-5 or 6-12) are similar (when expressed as Weber  
251 fractions) for  $F0$ s within the range tested in this study (e.g., Moore and Moore, 2003).

252 The effect of harmonic rank found here differs from that reported by Vliegen and Oxenham  
253 (1999) but is consistent with the findings of Grimault et al. (2000). However, the results from  
254 the present study are not consistent with the suggestions made by Grimault et al. (2000) to  
255 explain the difference between their results and those of Vliegen and Oxenham (1999).

256 Firstly, the present results showed an effect of harmonic rank when the listeners were  
257 instructed to try to segregate, as was done by Vliegen and Oxenham (1999) but not by  
258 Grimault et al. (2000). Secondly, streaming differed between conditions that did not differ  
259 greatly in terms of the resolvability of the harmonics in the complex tones. For example, the  
260 harmonics can be assumed to be mostly completely unresolved for the  $F0$ s of 80 and 100 Hz

261 (lowest harmonics in the passband were 27 and 22, respectively, for the A tones and 15 and  
262 12, respectively, for the B tones for  $\Delta F0 = 11$  ST), but streaming differed significantly for  
263 these two conditions.

264 One difference between the present study and that of Vliegen and Oxenham (1999) is that  
265 segregation here was quantified as the percentage of time that the listeners indicated that they  
266 heard two streams, measured over the whole duration of the sequence, while Vliegen and  
267 Oxenham (1999) obtained a single response for each sequence, presumably made towards the  
268 end of the sequence or after the sequence was finished. Stream segregation tends to build up  
269 over time for stimuli with small perceptual differences between successive sounds (Anstis and  
270 Saida, 1985; Bregman, 1978a) but can occur very rapidly when there are large perceptual  
271 differences. Using the measure of segregation of the present study, this build-up effect might  
272 have had a greater influence for stimuli where the build up was slow (small perceptual  
273 differences) than for stimuli where the build up was rapid (large perceptual differences), thus  
274 increasing differences across conditions. In the study of Vliegen and Oxenham (1999), the  
275 build up was probably near-complete for all stimuli. This might have contributed to the  
276 difference across studies.

277 To assess this possibility, the percentage of trials for which the two-streams key was the  
278 last key pressed was determined, giving a measure similar to that of Vliegen and Oxenham  
279 (1999). Analysis with a logistic generalized mixed-effects model for binary data using the  
280 lme4 library in R (Bates et al., 2015) showed significant effects of harmonic rank ( $\chi^2(5) =$   
281  $100.04, p < 0.001$ ) and of  $\Delta F0$  ( $\chi^2(5) = 822.96, p < 0.001$ ) and a significant interaction ( $\chi^2(25)$   
282  $= 402.52, p < 0.001$ ). Thus, it does not seem that the measure used here to assess the amount  
283 of segregation can explain the difference in results across studies.

284 Another difference between the two studies is the sequence length. In this study, each  
285 sequence contained 19 triplets whereas Vliegen and Oxenham (1999) used 12 triplets per  
286 sequence. However, since segregation builds up slowly over time when perceptual differences  
287 are small (Anstis and Saida, 1985; Bregman, 1978a), it would have been less likely to occur  
288 for conditions with high harmonic rank in the study of Vliegen and Oxenham (1999) than in

289 the present study, so this factor also cannot explain the difference between the results of the  
290 two studies.

291 Another factor that might have influenced the results is combination tones. The present  
292 study used a TEN to mask combination tones while Vliegen and Oxenham (1999) did not use  
293 any noise in their main experiment. Vliegen and Oxenham (1999) presented a preliminary  
294 experiment showing a small deleterious effect of masking noise on the stream segregation of  
295 complex tones containing only high harmonics, but a similar effect occurred for pure tones.  
296 They argued that combination tones were unlikely to explain their results. The results of the  
297 preliminary experiment, did, however, generally show more segregation for pure tones than  
298 for the complex tones, which is similar to the findings of this study but different from the  
299 results of their main study. Vliegen and Oxenham (1999) argued that this difference “may  
300 simply illustrate the large inter-subject variability”. In the present study, the results also  
301 varied markedly across listeners, so it is possible that inter-listener variability can explain the  
302 difference between studies. However, the fact that all listeners in the present study showed  
303 some effect of harmonic rank and the fact that Grimault et al. (2000) also found a significant  
304 effect of harmonic rank indicate that stream segregation does worsen with increasing  
305 harmonic rank, at least for most listeners.

306 The results of the present study are also consistent with studies that investigated F0  
307 discrimination for pairs of tones preceded and followed by complex tones with fixed F0  
308 (fringes) (Gockel et al., 1999; Micheyl and Carlyon, 1998). In these studies, it was proposed  
309 that the fringes interfere with F0 discrimination when the fringes and target tones are  
310 perceived as a single stream, but that interference is small when the fringes are perceived as  
311 being in a separate stream from the target tones. The results showed that when the mean F0 of  
312 the fringes differed from that of the target, there was more interference when both fringes and  
313 target tones contained unresolved harmonics than when they both contained resolved  
314 harmonics. This suggests that stream segregation of the fringes and target was more likely to  
315 occur when they both contained resolved harmonics, which is consistent with the results  
316 presented here.

317 The results from the present study confirm that stream segregation is possible for complex  
318 tones without resolved components. This is consistent with results from several studies  
319 showing that stream segregation can be induced using temporal cues alone, without any  
320 excitation-pattern cues (e.g., Dannenbring and Bregman, 1976; Grimault et al., 2002; Hong  
321 and Turner, 2009; Paredes-Gallardo et al., 2018; Roberts et al., 2002; Stainsby et al., 2004;  
322 Vliegen et al., 1999).

323 The present study found that the percentage of segregation increased with decreasing  
324 harmonic rank (increasing F0; Fig. 2) except that there was no difference between F0s of 250  
325 and 500 Hz. The mean streaming scores were very similar for those two conditions. Most  
326 individual scores were also similar for these conditions, but a few listeners (L2, L6 and L11)  
327 showed consistent decreases in streaming when the F0 was increased from 250 to 500 Hz.  
328 These decreases may be explained by the relatively small number of harmonics in the  
329 conditions with the A-tone F0 = 500 Hz. Assuming that all harmonics with a level of 55 dB  
330 SPL or above (which was the level/ERB<sub>N</sub> of the TEN) were audible, the A tone had six  
331 audible harmonic components and the number of audible harmonics in the B tone decreased  
332 with increasing  $\Delta F0$ . For the conditions with  $\Delta F0 = 7$  and 11 ST, the B tones had only four  
333 audible harmonics. Due to the limited number of well-resolved harmonics, a few listeners  
334 may have heard individual harmonics (spectral pitch) instead of the fundamental pitch of the  
335 tone complex (Schneider et al., 2005). They may have focused their attention on non-  
336 corresponding harmonics in the A and B tones. For example, when  $\Delta F0 = 11$  ST they may  
337 have attended to the 4th harmonic of the A tone (2000 Hz) and the second harmonic of the B  
338 tone (1879 Hz), which might have led to reduced segregation, since these harmonics differ in  
339 frequency by only slightly more than 1 ST.

340

## 341 **4 Experiment 2: Relation between stream segregation and discrimination of pure** 342 **tones and complex tones**

### 343 *4.1 Rationale*

344 FODLs were measured to determine the relationship between streaming and the salience of  
345 the F0 differences between the A and B tones. Furthermore, FODLs were measured for tones

346 whose harmonics were added in sine phase or in random phase, to provide an indirect  
347 measure of the resolvability of the harmonics. It is generally assumed that harmonic phase has  
348 an influence on FODLs only when the harmonics interfere, and therefore are at least partly  
349 unresolved (Houtsma and Smurzynski, 1990; Moore, 1977; Wang et al., 2012). The outputs  
350 of auditory filters in response to tones with unresolved harmonics have a higher peak factor  
351 for sine-phase tones than for random-phase tones. This is expected to affect FODLs based on  
352 the use of envelope cues. Therefore, FODLs are expected to be smaller for sine-phase than for  
353 the random-phase tones when all harmonics are unresolved.

354

## 355 4.2 Method

### 356 4.2.1 Stimuli

357 FODLs were measured for pure tones and complex tones similar to the ones used in  
358 experiment 1. The tones were bandpass filtered between 2 and 4 kHz and the nominal F0s of  
359 the reference tones were the same as the F0s of the A tones in experiment 1. Each tone had a  
360 duration of 500 ms including 10-ms raised-cosine onset and offset ramps. The interval  
361 between the three tones in each trial was 250 ms. The stimuli were presented at the same  
362 level and in the same TEN as for experiment 1. The TEN had the same purposes as for  
363 experiment 1. In addition, it was intended to promote synthetic rather than analytic listening  
364 (listening to the pitch corresponding to the missing F0 rather than to individual harmonics),  
365 since background noise promotes synthetic listening (Hall and Peters, 1981; Houtgast, 1976).

366

### 367 4.2.2 Procedure

368 FODLs were measured using a 3-alternative-forced-choice (AFC) weighted up-down  
369 paradigm (Kaernbach, 1991) to estimate the 75% point on the psychometric function. The  
370 listeners were asked to indicate which of the intervals contained the tone with the different  
371 pitch (the deviant). The F0 of this tone was always higher than the reference F0. The  
372 reference F0 was roved by  $\pm 5\%$  between trials using a uniform distribution around the  
373 nominal value. For each run, the initial F0 difference between the reference and the deviant  
374  $(F0_{\text{deviant}} - F0_{\text{reference}})/F0_{\text{reference}}$  was 20%. In the following trials, the F0 difference was

375 decreased logarithmically by a step size that decreased after every second reversal. The F0DL  
376 was calculated as the geometric mean of the F0 difference at the last six out of 10 reversals.  
377 Each condition was tested twice during training and the final F0DLs were calculated from  
378 values obtained over five runs. To check whether the F0DLs had stabilized after training, a  
379 straight line was fitted to the five F0DLs for each condition and one more block for each  
380 condition was added if the slope of the line was significantly lower than 0 for more than two  
381 conditions.

382

383 

### 4.3 Results

384 As shown in Fig. 5, the F0DLs decreased (improved) with increasing F0 (decreasing  
385 harmonic rank) for conditions with both sine and with random phase and were larger for  
386 random than for sine phase for the conditions with high harmonic rank but similar across  
387 phases for the conditions with lower harmonic rank. Analysis using a mixed-effects model  
388 with F0 and phase as fixed factors and listener as a random factor confirmed that both main  
389 effects (F0:  $F(4, 628) = 223.9, p < 0.001$ , phase:  $F(1, 628) = 32.77, p < 0.001$ ) and the  
390 interaction ( $F(4, 628) = 10.61, p < 0.001$ ) were significant. Comparisons of pairs of conditions  
391 with different phase but the same F0 (adjusted for multiple comparisons controlling the false  
392 discovery rate (Benjamini and Hochberg, 1995)) showed significant effects of phase for F0s  
393 of 80 Hz ( $t(628) = -6.25, p < 0.001$ ), 100 Hz ( $t(628) = -5.46, p < 0.001$ ) and 150 Hz ( $t(628) =$   
394  $-2.23, p = 0.031$ ) but not for the higher F0s, consistent with the idea that an effect of  
395 component phase occurs when the harmonics are unresolved.

396 Both stream segregation and F0 discrimination improved with increasing F0 (decreasing  
397 harmonic rank). The left panel of Fig. 6 illustrates this relationship. In this scatter plot, the  
398 mean percentage segregation for each A-tone F0 (averaged geometrically across  $\Delta F0$ s and  
399 across listeners) is plotted against the mean F0DL (across listeners) obtained for the same F0.  
400 There was a strong negative Pearson correlation between the two measures ( $r = -0.95, p =$   
401  $0.002$ , one tailed, since a negative correlation was hypothesized), indicating that small F0DLs  
402 are associated with greater streaming. To investigate the relationship between stream  
403 segregation and F0DLs for the individual listeners, for each listener the mean segregation

404 score was plotted against the mean F0DL (right panel of Fig. 6). There was a general  
405 tendency for stream segregation to decrease with increasing F0DL indicating that good F0  
406 discrimination for an individual is associated with greater segregation for that individual. The  
407 Pearson correlation was moderate but significant ( $r = -0.54$ ,  $p = 0.03$ , one tailed).

408

#### 409 4.4 Discussion

410 The increase in F0DLs with increasing harmonic rank and the effect of phase seen in Fig. 5  
411 are consistent with what has been found in earlier studies (e.g., Bernstein and Oxenham,  
412 2006a; Bernstein and Oxenham, 2006b; Houtsma and Smurzynski, 1990; Wang et al., 2012).  
413 The better performance for sine than for random phase for tones containing only high  
414 harmonics is thought to reflect the use of envelope cues resulting from the interference of  
415 harmonics in the cochlea. No phase effects are expected when one or more resolved  
416 harmonics are present, since performance is then dominated by the resolved harmonic(s). The  
417 results therefore suggest that the complex tones with F0s of 80, 100 and 150 Hz did not  
418 contain any resolved harmonics whereas the tones with higher F0s did. However, the F0DLs  
419 for the random-phase tones did increase significantly as the F0 decreased from 150 to 80 Hz  
420 ( $t(628) = 5.25$ ,  $p < 0.001$ ), suggesting that F0 discrimination worsens with increasing harmonic  
421 rank even when all harmonics are unresolved. This is consistent with the idea that the  
422 worsening of F0DLs with increasing harmonic rank reflects an effect of harmonic rank *per se*,  
423 rather than an effect of resolvability.

424 Figure 6 shows a clear relationship between stream segregation and F0 discrimination,  
425 supporting the idea that the extent of stream segregation depends on the salience of the  
426 perceptual difference between successive sounds. This is consistent with result from two  
427 recent studies that both showed a relationship between pitch salience and sequential stream  
428 segregation performance (Paredes-Gallardo et al., 2018; Shearer et al., 2018).

429 Some studies have shown a significant relationship between speech-in-speech perception  
430 and performance in a stream segregation task (Gaudrain et al., 2012; Hong and Turner, 2006;  
431 Mackersie et al., 2001). This raises the possibility that F0 discrimination might be related to  
432 speech-in-speech perception. Furthermore, musical training is associated with enhanced

433 frequency discrimination and F0 discrimination (e.g., Bianchi et al., 2016; Brown et al., 2017;  
434 Madsen et al., 2017; Micheyl et al., 2006; Ruggles et al., 2014), so it is possible that musical  
435 training would be associated with better stream segregation and better speech in speech  
436 perception. However, two recent studies have shown that musicians are not better than non-  
437 musicians at using F0 differences between competing voices to understand speech (Deroche  
438 et al., 2017; Madsen et al., 2017) and the latter specifically found no relationship between  
439 F0DLs and speech-in-speech perception.

440

## 441 **5 Overall summary and conclusions**

442 Experiment 1 investigated the effect of harmonic rank on the subjective sequential stream  
443 segregation of complex tones in a task where the listeners were instructed to try to segregate.  
444 Stream segregation scores were compared to F0DLs measured using similar stimuli in  
445 experiment 2.

446 The results of experiment 1 showed that: (1) segregation increased with decreasing  
447 harmonic rank; (2) the effect of harmonic rank was continuous and progressive and even  
448 small differences in harmonic rank led to differences in segregation.

449 In experiment 2, F0DLs were measured for pure tones and complex tones similar to the A-  
450 tones used in experiment 1. F0DLs increased with increasing harmonic rank. Significant  
451 correlations were found between the mean percentage segregation for each A-tone F0  
452 (averaged across  $\Delta F0$ s and across listeners) and the mean F0DL (across listeners) obtained for  
453 the same F0 and between the mean percentage segregation for each listener (averaged across  
454 all conditions) and the mean F0DL for that listener (averaged across all F0s). This supports  
455 the idea that the extent of stream segregation of successive sounds depends on the salience of  
456 the perceptual difference between those sounds.

457

## 458 **Acknowledgements**

459 This study was supported by the Carlsberg Foundation. We thank Andrew Oxenham, Bob  
460 Carlyon, Federica Bianchi, Sebastián Santurette and two reviewers for useful discussions and  
461 comments.

462

463 **References**

- 464 Anstis, S., Saida, S. 1985. Adaptation to auditory streaming of frequency-modulated tones. *J.*  
465 *Exp. Psychol. Hum. Percept. Perform.* 11, 257-271.
- 466 Bates, D., Mächler, M., Bolker, B.M., Walker, S.C. 2015. Fitting linear mixed-effects models  
467 using lme4. *J. Stat. Softw.* 67, 1-48.
- 468 Benjamini, Y., Hochberg, Y. 1995. Controlling the false discovery rate - a practical and  
469 powerful approach to multiple testing *J. R. Stat. Soc. Ser. B-Methodol.* 57, 289-300.
- 470 Bernstein, J.G., Oxenham, A.J. 2003. Pitch discrimination of diotic and dichotic tone  
471 complexes: harmonic resolvability or harmonic number? *J. Acoust. Soc. Am.* 113,  
472 3323-3334.
- 473 Bernstein, J.G., Oxenham, A.J. 2005. An autocorrelation model with place dependence to  
474 account for the effect of harmonic number on fundamental frequency discrimination.  
475 *J. Acoust. Soc. Am.* 117, 3816-3831.
- 476 Bernstein, J.G., Oxenham, A.J. 2006a. The relationship between frequency selectivity and  
477 pitch discrimination: effects of stimulus level. *J. Acoust. Soc. Am.* 120, 3916-3928.
- 478 Bernstein, J.G., Oxenham, A.J. 2006b. The relationship between frequency selectivity and  
479 pitch discrimination: sensorineural hearing loss. *J. Acoust. Soc. Am.* 120, 3929-3945.
- 480 Bianchi, F., Santurette, S., Wendt, D., Dau, T. 2016. Pitch discrimination in musicians and  
481 non-musicians: effects of harmonic resolvability and processing effort. *Assoc. Res.*  
482 *Otolaryngol.* 17, 69-79.
- 483 Bregman, A.S. 1978a. Auditory streaming is cumulative. *J. Exp. Psychol. Hum. Percept.*  
484 *Perform.* 4, 380-387.
- 485 Bregman, A.S. 1978b. The formation of auditory streams. In: Requin, J., (Ed.), *Attention and*  
486 *Performance VII.* Erlbaum, Hillsdale, NJ.
- 487 Bregman, A.S. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound.*  
488 Bradford Books, MIT Press, Cambridge, Mass. pp. 455-528.
- 489 Brown, C.J., Jeon, E.K., Driscoll, V., Mussoi, B., Deshpande, S.B., Gfeller, K., Abbas, P.J.  
490 2017. Effects of long-term musical training on cortical auditory evoked potentials. *Ear*  
491 *Hear.* 38, E74-E84.
- 492 Brungart, D.S., Simpson, B.D., Ericson, M.A., Scott, K.R. 2001. Informational and energetic  
493 masking effects in the perception of multiple simultaneous talkers. *J. Acoust. Soc.*  
494 *Am.* 110, 2527-2538.

- 495 Dannenbring, G.L., Bregman, A.S. 1976. Stream segregation and the illusion of overlap. *J.*  
496 *Exp. Psychol. Hum. Percept. Perform.* 2, 544-555.
- 497 Deroche, M.L.D., Limb, C.J., Chatterjee, M., Gracco, V.L. 2017. Similar abilities of  
498 musicians and non-musicians to segregate voices by fundamental frequency. *J.*  
499 *Acoust. Soc. Am.* 142, 1739-1755.
- 500 Gaudrain, E., Grimault, N., Healy, E.V., Bera, J.C. 2012. The Relationship Between  
501 Concurrent Speech Segregation, Pitch-Based Streaming of Vowel Sequences, and  
502 Frequency Selectivity. *Acta Acust. United Acust.* 98, 317-327.
- 503 Glasberg, B.R., Moore, B.C.J. 1990. Derivation of auditory filter shapes from notched-noise  
504 data. *Hear. Res.* 47, 103-138.
- 505 Gockel, H., Carlyon, R.P., Micheyl, C. 1999. Context dependence of fundamental-frequency  
506 discrimination: Lateralized temporal fringes. *J. Acoust. Soc. Am.* 106, 3553-3563.
- 507 Grimault, N., Bacon, S.P., Micheyl, C. 2002. Auditory stream segregation on the basis of  
508 amplitude-modulation rate. *J. Acoust. Soc. Am.* 111, 1340-1348.
- 509 Grimault, N., Micheyl, C., Carlyon, R.P., Arthaud, P., Collet, L. 2000. Influence of peripheral  
510 resolvability on the perceptual segregation of harmonic complex tones differing in  
511 fundamental frequency. *J. Acoust. Soc. Am.* 108, 263-271.
- 512 Grimault, N., Micheyl, C., Carlyon, R.P., Arthaud, P., Collet, L. 2001. Perceptual auditory  
513 stream segregation of sequences of complex sounds in subjects with normal and  
514 impaired hearing. *Br. J. Audiol.* 35, 173-182.
- 515 Hall, J.W., Peters, R.W. 1981. Pitch for nonsimultaneous successive harmonics in quiet and  
516 noise. *J. Acoust. Soc. Am.* 69, 509-513.
- 517 Hartmann, W.M., Johnson, D. 1991. Stream segregation and peripheral channeling. *Music*  
518 *Perception* 9, 155-184.
- 519 Hoekstra, A., Ritsma, R.J. 1977. Perceptive hearing loss and frequency selectivity. In: Evans,  
520 E.F., Wilson, J.P., (Eds.), *Psychophysics and Physiology of Hearing*. Academic,  
521 London, England. pp. 263-271.
- 522 Hong, R.S., Turner, C.W. 2006. Pure-tone auditory stream segregation and speech perception  
523 in noise in cochlear implant recipients. *J. Acoust. Soc. Am.* 120, 360-374.
- 524 Hong, R.S., Turner, C.W. 2009. Sequential stream segregation using temporal periodicity  
525 cues in cochlear implant recipients. *J. Acoust. Soc. Am.* 126, 291-299.
- 526 Houtgast, T. 1976. Subharmonic pitches of a pure tone at low S/N ratio. *J. Acoust. Soc. Am.*  
527 60, 405-409.

- 528 Houtsma, A.J.M., Smurzynski, J. 1990. Pitch identification and discrimination for complex  
529 tones with many harmonics. *J. Acoust. Soc. Am.* 87, 304-310.
- 530 Jackson, H.M., Moore, B.C.J. 2013. The dominant region for the pitch of complex tones with  
531 low fundamental frequencies. *J. Acoust. Soc. Am.* 134, 1193-1204.
- 532 Kaernbach, C. 1991. Simple adaptive testing with the weighted up-down method. *Percept.*  
533 *Psychophys.* 49, 227-229.
- 534 Lenth, R.V. 2016. Least-Squares Means: The R Package lsmeans. *J. Stat. Softw.*, 1-33.
- 535 Mackersie, C.L., Prida, T.L., Stiles, D. 2001. The role of sequential stream segregation and  
536 frequency selectivity in the perception of simultaneous sentences by listeners with  
537 sensorineural hearing loss. *J. Speech. Lang. Hear. Res.* 44, 19-28.
- 538 Madsen, S.M.K., Dau, T., Moore, B.C.J. 2015. Effect of harmonic rank on the streaming of  
539 complex tones. In: Santurette S., D.T., Christensen-Dalsgaard J., Tranebjærg L.,  
540 Andersen T, (Ed.), *Proceedings of the International Symposium on Auditory and*  
541 *Audiological Research: Individual Hearing Loss - Characterization, modelling,*  
542 *compensation strategies*, Vol. 5.
- 543 Madsen, S.M.K., Whiteford, K.L., Oxenham, A.J. 2017. Musicians do not benefit from  
544 differences in fundamental frequency when listening to speech in competing speech  
545 backgrounds. *Sci. Rep.* 7.
- 546 Micheyl, C., Carlyon, R.P. 1998. Effects of temporal fringes on fundamental-frequency  
547 discrimination. *J. Acoust. Soc. Am.* 104, 3006-3018.
- 548 Micheyl, C., Delhommeau, K., Perrot, X., Oxenham, A.J. 2006. Influence of musical and  
549 psychoacoustical training on pitch discrimination. *Hear. Res.* 219, 36-47.
- 550 Moore, B.C.J. 1977. Effects of relative phase of the components on the pitch of three-  
551 component complex tones. In: Evans, E.F., Wilson, J.P., (Eds.), *Psychophysics and*  
552 *Physiology of Hearing*. Academic Press, London. pp. 349-358.
- 553 Moore, B.C.J. 2003. *An Introduction to the Psychology of Hearing*, 5th Ed. Emerald, Bingley,  
554 UK.
- 555 Moore, B.C.J., Gockel, H. 2002. Factors influencing sequential stream segregation. *Acta*  
556 *Acust. United Acust.* 88, 320-333.
- 557 Moore, B.C.J., Moore, G.A. 2003. Discrimination of the fundamental frequency of complex  
558 tones with fixed and shifting spectral envelopes by normally hearing and hearing-  
559 impaired subjects. *Hear. Res.* 182, 153-163.
- 560 Moore, B.C.J., Gockel, H. 2011. Resolvability of components in complex tones and  
561 implications for theories of pitch perception. *Hear. Res.* 276, 88-97.

- 562 Moore, B.C.J., Huss, M., Vickers, D.A., Glasberg, B.R., Alcántara, J.I. 2000. A test for the  
563 diagnosis of dead regions in the cochlea. *Br. J. Audiol.* 34, 205-224.
- 564 Oxenham, A.J., Micheyl, C., Keebler, M.V. 2009. Can temporal fine structure represent the  
565 fundamental frequency of unresolved harmonics? *J. Acoust. Soc. Am.* 125, 2189-  
566 2199.
- 567 Paredes-Gallardo, A., Madsen, S.M.K., Dau, T., Marozeau, J. 2018. The role of temporal cues  
568 in voluntary stream segregation for cochlear implant users. *Trends. Hear.* 22.
- 569 Roberts, B., Glasberg, B.R., Moore, B.C.J. 2002. Primitive stream segregation of tone  
570 sequences without differences in F0 or passband. *J. Acoust. Soc. Am.* 112, 2074-2085.
- 571 Rose, M.M., Moore, B.C.J. 1997. Perceptual grouping of tone sequences by normally hearing  
572 and hearing-impaired listeners. *J. Acoust. Soc. Am.* 102, 1768-1778.
- 573 Ruggles, D.R., Freyman, R.L., Oxenham, A.J. 2014. Influence of musical training on  
574 understanding voiced and whispered speech in noise. *PLoS One* 9, e86980
- 575 Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H.J., Dosch, H.G.,  
576 Bleeck, S., Stippich, C., Rupp, A. 2005. Structural and functional asymmetry of lateral  
577 Heschl's gyrus reflects pitch perception preference. *Nat. Neurosci.* 8, 1241-1247.
- 578 Shackleton, T.M., Carlyon, R.P. 1994. The role of resolved and unresolved harmonics in pitch  
579 perception and frequency modulation discrimination. *J. Acoust. Soc. Am.* 95, 3529-  
580 3540.
- 581 Shearer, D.E., Molis, M.R., Bennett, K.O., Leek, M.R. 2018. Auditory stream segregation of  
582 iterated rippled noises by normal-hearing and hearing-impaired listeners. *J. Acoust.*  
583 *Soc. Am.* 143, 378-387.
- 584 Stainsby, T.H., Moore, B.C.J., Medland, P.J., Glasberg, B.R. 2004. Sequential streaming and  
585 effective level differences due to phase-spectrum manipulations. *J. Acoust. Soc. Am.*  
586 115, 1665-1673.
- 587 van Noorden, L.P.A.S. 1975. Temporal coherence in the perception of tone sequences. Ph.D.  
588 Thesis, Eindhoven University of Technology.
- 589 Vliegen, J., Oxenham, A.J. 1999. Sequential stream segregation in the absence of spectral  
590 cues. *J. Acoust. Soc. Am.* 105, 339-346.
- 591 Vliegen, J., Moore, B.C.J., Oxenham, A.J. 1999. The role of spectral and periodicity cues in  
592 auditory stream segregation, measured using a temporal discrimination task. *J. Acoust.*  
593 *Soc. Am.* 106, 938-945.

- 594 Wang, J., Baer, T., Glasberg, B.R., Stone, M.A., Ye, D.T., Moore, B.C.J. 2012. Pitch  
595 perception of concurrent harmonic tones with overlapping spectra. *J. Acoust. Soc.*  
596 *Am.* 132, 339-356.
- 597 Wobbrock, J.O., Findlater, L., Gergle, D., Higgins, J.J. 2011. The Aligned Rank Transform  
598 for nonparametric factorial analysis using only ANOVA procedures, Proceedings of  
599 the ACM Conference on Human Factors in Computing Systems (CHI '11). New York:  
600 ACM Press, Vancouver, British Columbia. pp. 143-146.
- 601
- 602

603 Figure captions:

604 **Fig. 1.** Schematic illustration of the stimuli. A) Illustration of the ABA-ABA sequences used.

605 A and B tones with a small difference in F0 ( $\Delta F0$ ) are likely to be perceived as one stream

606 (integration; upper panel) whereas A and B tones with large  $\Delta F0$  are likely to be perceived as

607 two streams (segregation; lower panel). B) Schematic spectra of complex tones. Tones were

608 initially generated with many harmonics, and were bandpass filtered between 2 and 4 kHz.

609 The harmonic rank was varied by varying the F0. Examples are shown of tones with low

610 harmonics (F0 = 500 Hz, left) and high harmonics (F0 = 150 Hz, right).

611 **Fig. 2.** Percentage of time the sequences were indicated as being perceived as two streams for

612 each listener. The main boxes show the percentages for the complex tones, plotted as a

613 function of the F0 of Tone A, with  $\Delta F0$  as parameter. The smaller panels to the right show

614 results obtained with pure tones. Different symbols refer to different frequency differences or

615 F0 differences between the A and B tones. Error bars indicate  $\pm 1$  SE across trials.

616 **Fig. 3.** As Fig. 2, but showing the mean across listeners. Error bars indicate  $\pm 1$  standard error.

617 **Fig. 4.** Percentage of trials indicated as being heard as two streams after a given number of

618 seconds for the condition with  $\Delta F0 = 7$  and F0 = 250 Hz. Each line shows results for one

619 listener.

620 **Fig. 5.** Frequency difference limens for pure tones (right) and F0DLs for complex tones added

621 in sine phase (filled diamonds) and in random phase (open circles) (left), plotted as a function

622 of the reference F0. Error bars indicate  $\pm 1$  SE.

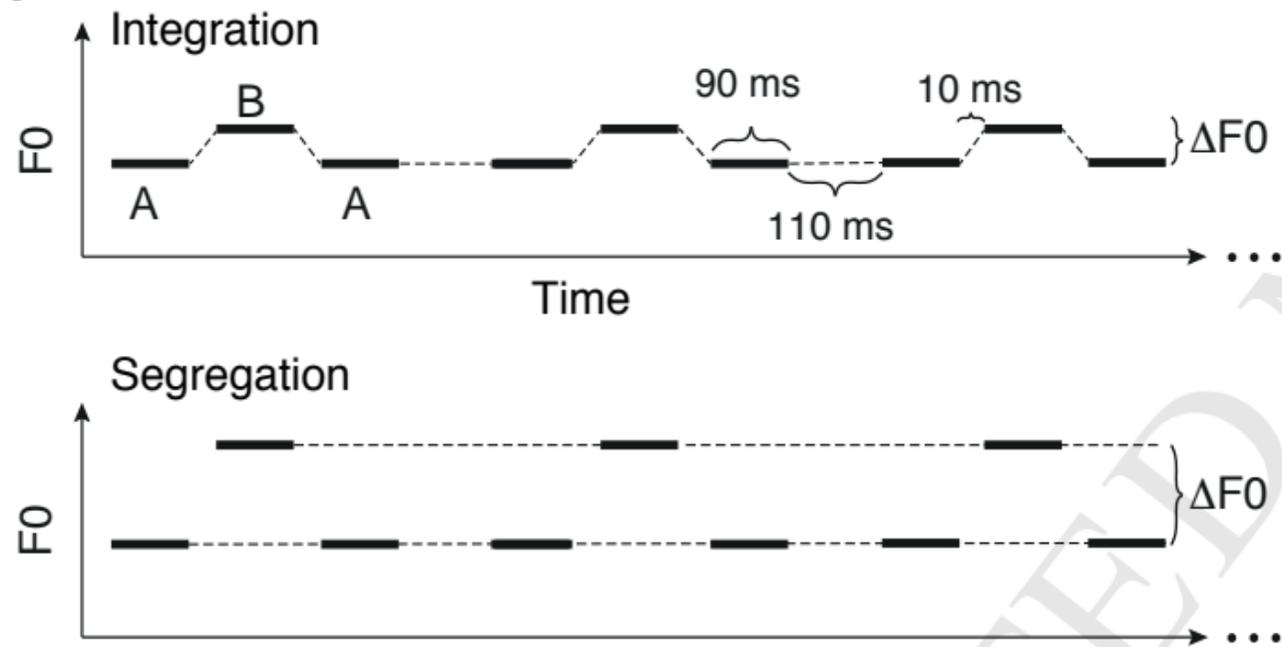
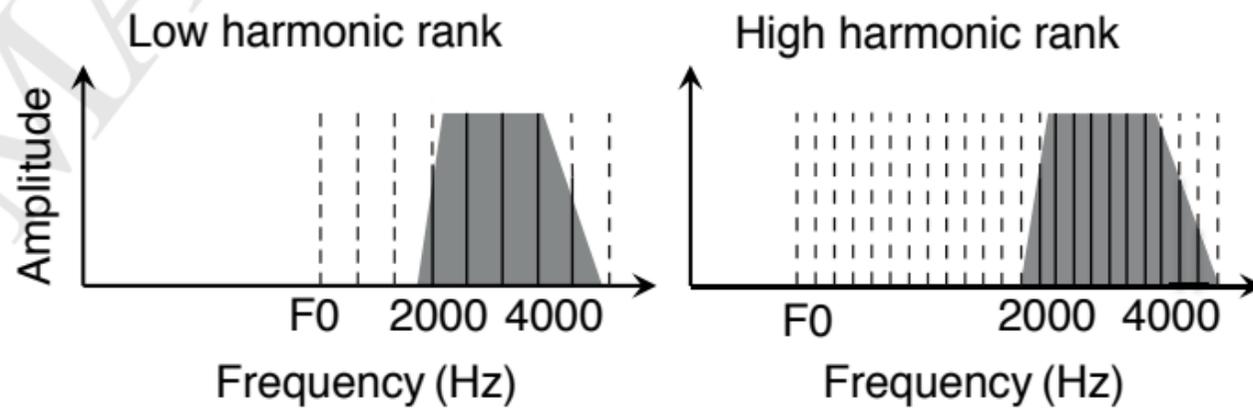
623 **Fig. 6.** Scatter plots showing the relation between stream segregation and F0DLs. Left panel:

624 the mean percentage segregation for each A-tone F0 (averaged across  $\Delta F0$ s and across

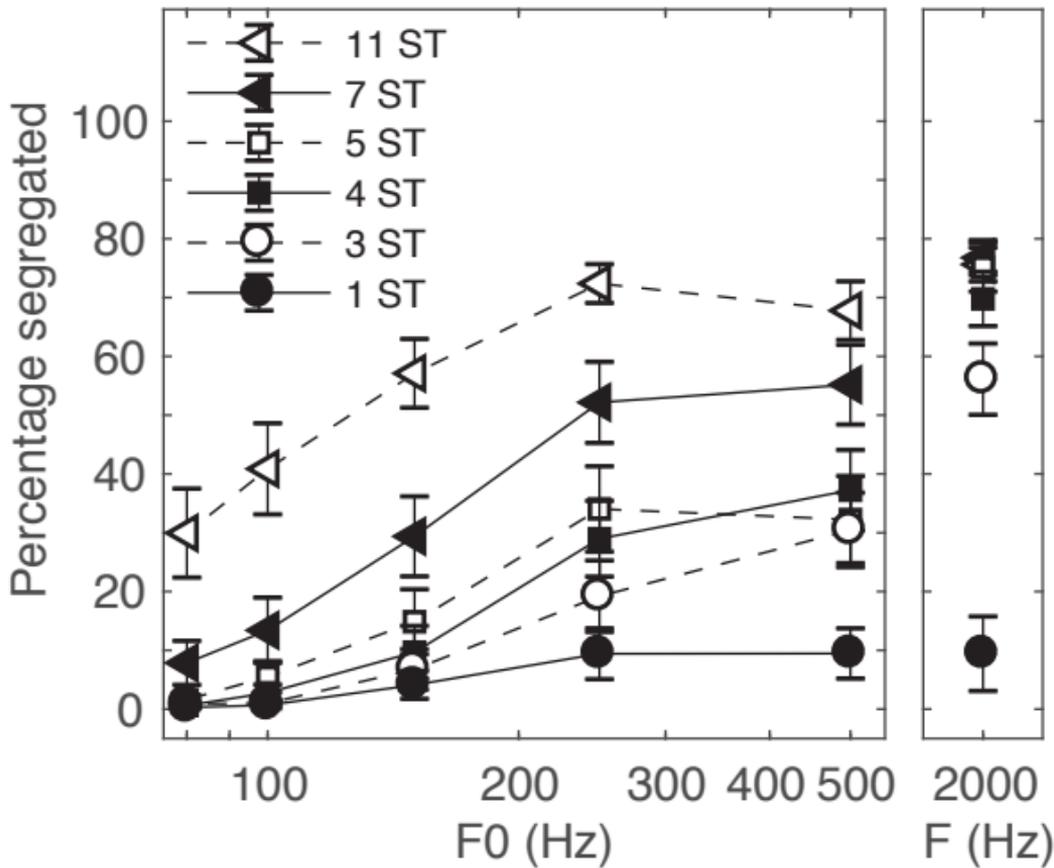
625 listeners) is plotted against the mean F0DL (across listeners) obtained for the same F0. Right

626 panel: the mean percentage segregation score for each listener (averaged across all conditions)

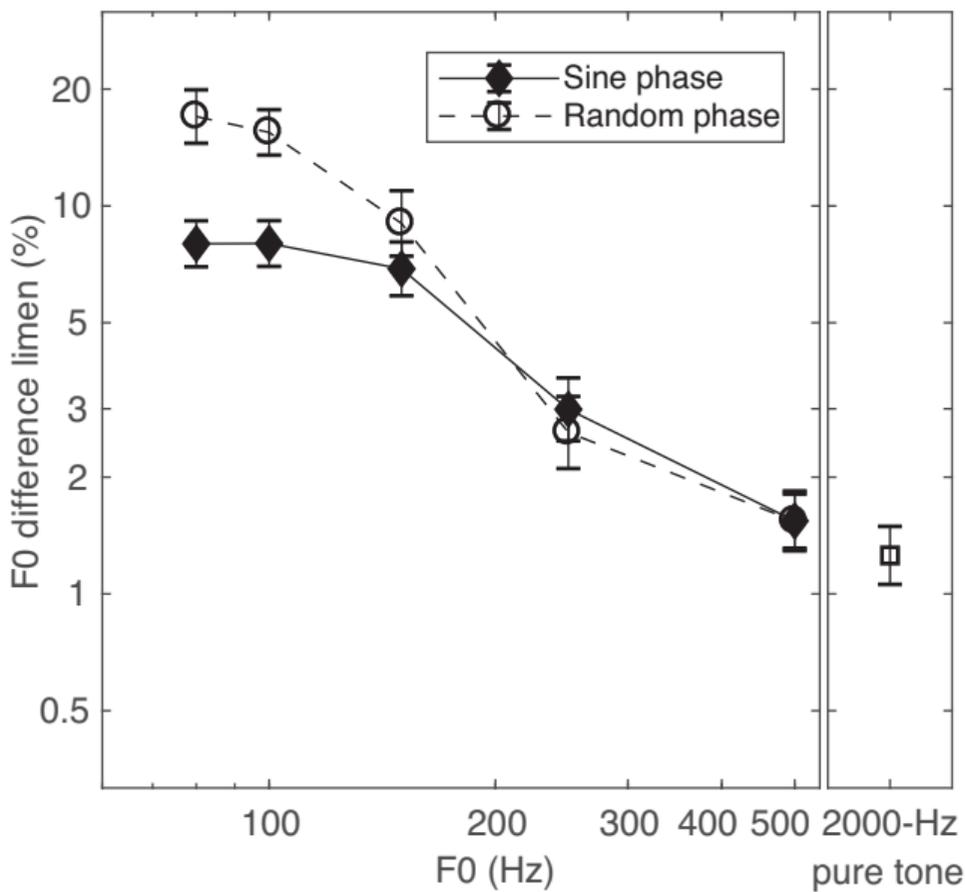
627 is plotted against the mean F0DL for that listener (averaged across all F0s).

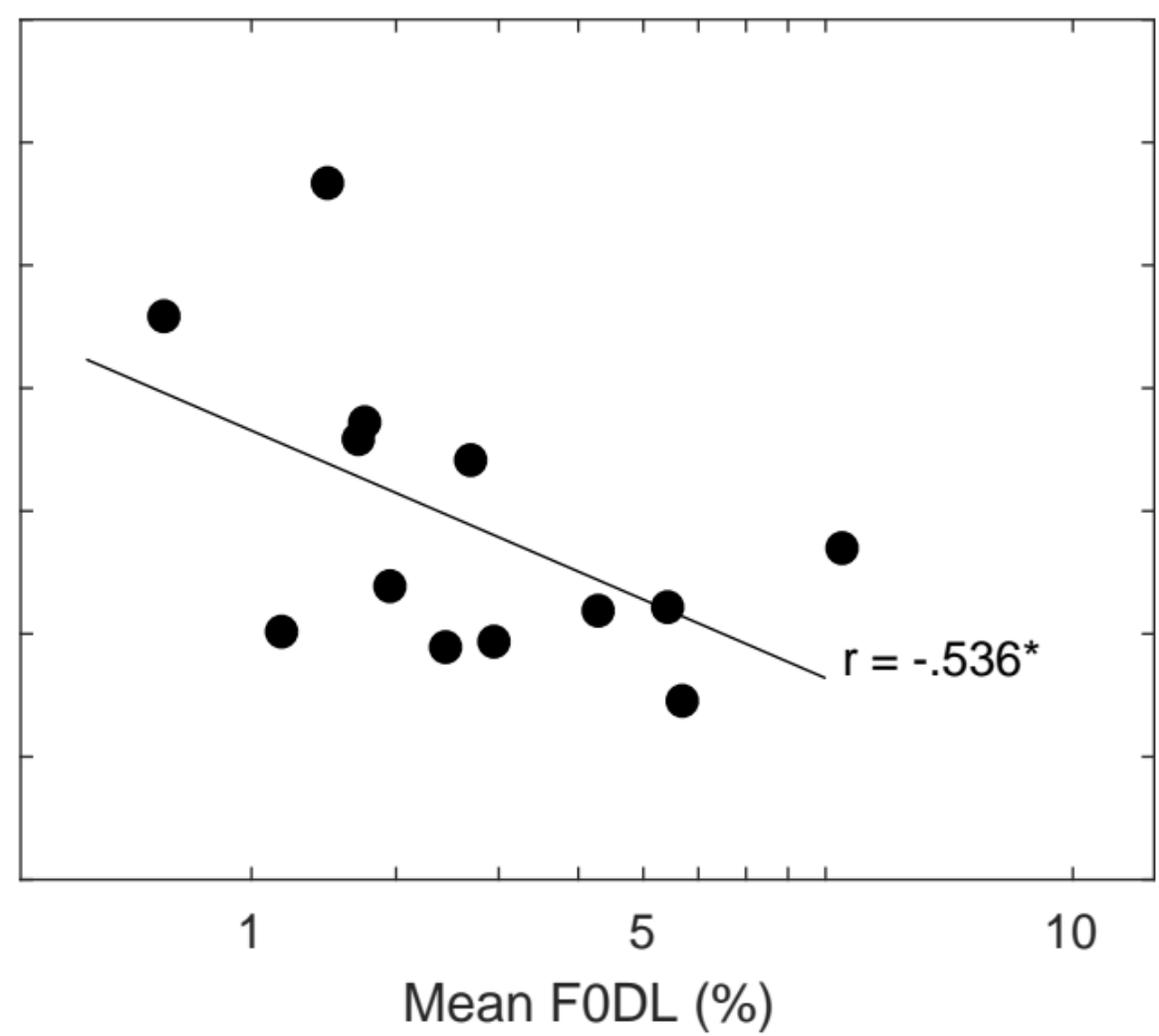
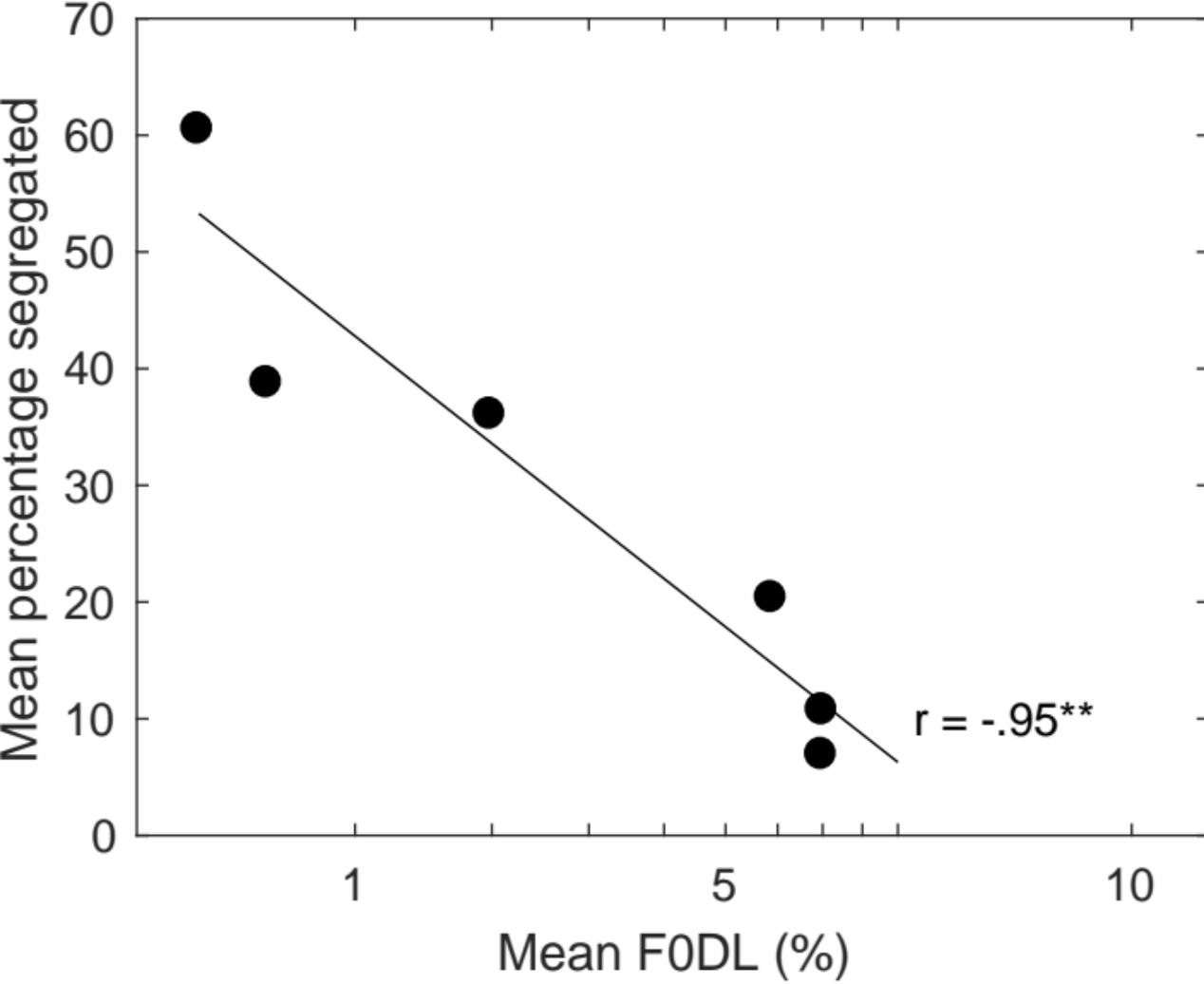
**A****B**











- Sequential stream segregation and fundamental frequency difference limens (FODLs) were measured for pure tones and complex tones with varying harmonic rank
- Segregation increased with decreasing harmonic ranks
- The effect of harmonic rank was continuous and progressive and even small differences in harmonic rank led to differences in segregation.
- Correlations were found between stream segregation and FODLs supporting the idea that the extent of stream segregation depends on the salience of the perceptual difference between sounds