Comments

Comment by Gockel:

Could you comment on why, in your second experiment, FDTs increase when dF_{AB} is increased to 10 semitones? Do you think it is because, in this condition, tones A and B are separated by something coming closer to an octave?

Reply:

At this point, we have no clear explanation for the trend that you point out. All we can offer are the following suggestions: The observed increase in FDTs between 7 and 10 ST might reflect the fact that as dF_{AB} was approaching one octave, the A and B tones were becoming more similar along the chroma dimension, and thus, more easily confused by listeners. However, at the same time, the A and B tones were also becoming increasingly dissimilar in pitch height. This, together with the fact that the A and B tones had substantially different durations (200 and 100 ms, respectively), makes it unlikely that listeners mistook the A tones for B or B' tones. A second possible explanation is that when the A and B tones became more similar in chroma, they tended to fall into the same stream. To the extent that - as we propose in the chapter - the integration of the A and B tones into the same stream interferes with the extraction and/or the comparison of the frequency information in/between the B and B' tones, this may explain the finding that FDTs tended to increase as dF_{AB} was increasing toward 1 octave. However, to our knowledge, there is no indication in the existing streaming literature that stream segregation is reduced when the frequency ratio between the A and B tones comes close to an integer multiple of an octave. In fact, based on the currently-available literature, we feel that a frequency difference of 10 ST should have been largely sufficient to produce stream segregation, especially since the two streams had different tempi, which provided an additional cue for teasing them apart. Yet another possible explanation relates to the idea that at least for intermediate dFs, (the build-up of) streaming requires attention (Carlyon *et al*. 2002). So, it is possible that at dF_{AB}s < 10 ST, the listeners were attending hard to the tones in order to stream them apart, whereas at $dF_{AB} = 10$ ST, less attention being required for stream segregation, the listeners devoted less sustained attention to the stimuli, which might explain the increase in thresholds. Further experiments are needed to clarify this issue.

Carlyon R.P., Cusack R., Foxton J.M., Robertson I.H. (2001) Effects of attention and unilateral neglect on auditory stream segregation. J. Exp. Psychol. Hum. Percept. Perform. 27, 115-127.

Comment by McAdams:

In your Experiment 2, listeners must discriminate the frequency of the high tone at the end of the stimulus sequence (B') from the other high tones (B). The results show an improvement in discrimination performance as a function of the A-B frequency difference in your gallop stimuli, ABA—ABA—ABA—. I would like to propose another possible interpretation of the data. When a two-stream percept is prominent (large A-B difference), listeners can focus on the B stream and do a simple frequency discrimination between B and B'. However, when a one-stream percept is prominent (small A-B difference), the A and B tones are integrated into a single stream and the perceived pattern ABA has to be discriminated from AB'A since it is more difficult to separate out the B tone from the A tones perceptually. In this case, what listeners may be judging is the difference in pitch interval of A-B compared to A-B'. It would be interesting to check whether simple interval discrimination has thresholds equivalent to those of your small A-B difference. If so, one might conclude that listeners are in fact performing different tasks in the two end regions of your Fig. 5 (interval discrimination at small A-B difference and frequency discrimination at large A-B difference), and are switching between the two in the transition region.

Reply:

The general idea in your comment - that listeners were performing different tasks at small and large frequency separations - is entirely consistent with our interpretation of the results. We suggest that the perceptual organization of the sound sequences into one or two streams constrained what listeners could do: when the A and B or B' tones were segregated into different streams, it was probably difficult or impossible for the listeners to combine or compare information between these tones; when the tones were integrated into the same stream, it was probably difficult or impossible for listeners to focus on just the B tones and ignore the A tones. As a result of this, listeners may have been led to use to different task-performance strategies at small and large frequency separations, indeed.

Your specific suggestion - that what listeners were perhaps doing at small frequency separations was to compare the frequency intervals formed by successive AB and AB' pairs – is certainly be worth exploring in future studies. However, at this stage, other possible interpretations appear equally likely. For instance, it could be that at small frequency separations, the pitch sensations evoked by temporally-neighboring A and B (or B') tones somehow interfered with one another in short-term memory (either at the encoding or at the retrieval stage), so that although listeners were really trying to compare the frequencies of the B and B' tones, the internal representations of these frequencies being altered, performance was worse.