Springer
the language of science

Dear Author/Editor,

Here are the proofs of your chapter as well as the metadata sheets.

Metadata

- Please carefully proof read the metadata, above all the names and address.
- In case there were no abstracts for this book submitted with the manuscript, the first 10-15 lines of the first paragraph were taken. In case you want to replace these default abstracts, please submit new abstracts with your proof corrections.

Page proofs

- Please check the proofs and mark your corrections either by
  - entering your corrections online
    or
  - opening the PDF file in Adobe Acrobat and inserting your corrections using the tool "Comment and Markup"
    or
  - printing the file and marking corrections on hardcopy. Please mark all corrections in dark pen in the text and in the margin at least ¼" (6 mm) from the edge.
- You can upload your annotated PDF file or your corrected printout on our Proofing Website. In case you are not able to scan the printout , send us the corrected pages via fax.
- Please note that any changes at this stage are limited to typographical errors and serious errors of fact.
- If the figures were converted to black and white, please check that the quality of such figures is sufficient and that all references to color in any text discussing the figures is changed accordingly. If the quality of some figures is judged to be insufficient, please send an improved grayscale figure.

Metadata of the chapter that will be visualized online

Book Title		Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing
Chapter Title		Relative Pitch Perception and the Detection of Deviant Tone Patterns
Copyright		The Author(s) 2016
Corresponding Author	Prefix	
	Family name	**Denham**
	Particle	
	Given name	**Susan L.**
	Suffix	
	Division	Cognition Institute and School of Psychology
	Organization	Plymouth University
	Address	Plymouth, UK
	Email	sdenham@plymouth.ac.uk
Author	Prefix	
	Family name	**Coath**
	Particle	
	Given name	**Martin**
	Suffix	
	Division	Cognition Institute and School of Psychology
	Organization	Plymouth University
	Address	Plymouth, UK
	Email	martin.coath@plymouth.ac.uk
Author	Prefix	
	Family name	**Háden**
	Particle	
	Given name	**Gábor P.**
	Suffix	
	Division	Institute of Cognitive Neuroscience and Psychology
	Organization	Research Centre for Natural Sciences, Hungarian Academy of Sciences
	Address	Budapest, Hungary
	Email	haden@cogpsyphy.hu
Author	Prefix	
	Family name	**Murray**
	Particle	
	Given name	**Fiona**
	Suffix	
	Division	School of Psychology
	Organization	Plymouth University
	Address	Plymouth, UK
	Email	fiona.murray@students.plymouth.ac.uk
Author [Aff 1]	Prefix	
	Family name	**Winkler**
	Particle	
	Given name	**István**
	Suffix	
	Division	Institute of Cognitive Neuroscience and Psychology
	Organization	Research Centre for Natural Sciences, Hungarian Academy of Sciences
	Address	Budapest, Hungary

	Email	winkler.istvan@ttk.mta.hu
Author [Aff 2]	Prefix	
	Family name	**Winkler**
	Particle	
	Given name	**István**
	Suffix	
	Division	Institute of Psychology
	Organization	University of Szeged
	Address	Szeged, Hungary

Abstract

Most people are able to recognise familiar tunes even when played in a different key. It is assumed that this depends on a general capacity for relative pitch perception; the ability to recognise the pattern of inter-note intervals that characterises the tune. However, when healthy adults are required to detect rare deviant melodic patterns in a sequence of randomly transposed standard patterns they perform close to chance. Musically experienced participants perform better than naïve participants, but even they find the task difficult, despite the fact that musical education includes training in interval recognition.

To understand the source of this difficulty we designed an experiment to explore the relative influence of the size of within-pattern intervals and between-pattern transpositions on detecting deviant melodic patterns. We found that task difficulty increases when patterns contain large intervals (5–7 semitones) rather than small intervals (1–3 semitones). While task difficulty increases substantially when transpositions are introduced, the effect of transposition size (large vs small) is weaker. Increasing the range of permissible intervals to be used also makes the task more difficult. Furthermore, providing an initial exact repetition followed by subsequent transpositions does not improve performance. Although musical training correlates with task performance, we find no evidence that violations to musical intervals important in Western music (i.e. the perfect fifth or fourth) are more easily detected. In summary, relative pitch perception does not appear to be conducive to simple explanations based exclusively on invariant physical ratios.

Keywords

Relative pitch perception - Musical intervals - Oddball paradigm - Pattern detection - Deviant detection - Translation-invariant perception

Relative Pitch Perception and the Detection of Deviant Tone Patterns

Susan L. Denham, Martin Coath, Gábor P. Háden, Fiona Murray and István Winkler

Abstract Most people are able to recognise familiar tunes even when played in a different key. It is assumed that this depends on a general capacity for relative pitch perception; the ability to recognise the pattern of inter-note intervals that characterises the tune. However, when healthy adults are required to detect rare deviant melodic patterns in a sequence of randomly transposed standard patterns they perform close to chance. Musically experienced participants perform better than naïve participants, but even they find the task difficult, despite the fact that musical education includes training in interval recognition.

To understand the source of this difficulty we designed an experiment to explore the relative influence of the size of within-pattern intervals and between-pattern transpositions on detecting deviant melodic patterns. We found that task difficulty increases when patterns contain large intervals (5–7 semitones) rather than small intervals (1–3 semitones). While task difficulty increases substantially when transpositions are introduced, the effect of transposition size (large vs small) is weaker. Increasing the range of permissible intervals to be used also makes the task more difficult. Furthermore, providing an initial exact repetition followed by subsequent

---

S. L. Denham (✉) · M. Coath
Cognition Institute and School of Psychology, Plymouth University, Plymouth, UK
e-mail: sdenham@plymouth.ac.uk

M. Coath
e-mail: martin.coath@plymouth.ac.uk

G. P. Háden · I. Winkler
Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary
e-mail: haden@cogpsyphy.hu

F. Murray
School of Psychology, Plymouth University, Plymouth, UK
e-mail: fiona.murray@students.plymouth.ac.uk

I. Winkler
Institute of Psychology, University of Szeged, Szeged, Hungary
e-mail: winkler.istvan@ttk.mta.hu

P. van Dijk et al. (eds.), *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing,* Advances in Experimental Medicine and Biology 894,
DOI 10.1007/978-3-319-25474-6_43

transpositions does not improve performance. Although musical training correlates with task performance, we find no evidence that violations to musical intervals important in Western music (i.e. the perfect fifth or fourth) are more easily detected. In summary, relative pitch perception does not appear to be conducive to simple explanations based exclusively on invariant physical ratios.

Keywords Relative pitch perception · Musical intervals · Oddball paradigm · Pattern detection · Deviant detection · Translation-invariant perception

1 Introduction

Most people easily recognise well known melodies even when they are transposed to a different key. The invariant property of transposed melodies is the preserved pitch ratio relationship between notes of the melody; i.e. pitch intervals of the melody remain the same despite changes in absolute pitch. For this reason, it is assumed that the ability to recognise pitch relationships (relative pitch perception) is rather robust and commonly found in the population. Recognition of preserved pitch interval patterns irrespective of absolute pitch is an auditory example of translation-invariant object perception (Kubovy and Van Valkenburg 2001; Griffiths and Warren 2004; Winkler et al. 2009).

The robustness of the ability to recognise tone patterns has been supported by recent findings showing that listeners can detect random tone patterns very quickly (after ca. 1.5 repetitions) within rapidly presented tone sequences, even if the patterns are quite long (up to 20 tones in a pattern) (Barascud 2014). The human brain is also sensitive to pattern violations, with regular to random transitions (Chait et al. 2007) being detected within about 150 ms (~3 tones) from deviation onset (Barascud 2014). However, in these examples tone patterns were always repeated exactly, i.e. without transposition, so it is not clear whether listeners were remembering absolute pitch sequences or relative pitch relationships.

In support of the assumed generality of relative pitch perception, it has been shown that violations of transposed pitch patterns elicit discriminative brain responses in neonates (Stefanics et al. 2009) and young infants (Tew et al. 2009). So it is surprising that relative pitch perception can be rather poor (e.g. see (Foster and Zatorre 2010; McDermott et al. 2010)), especially if contour violations and tonal melodies are excluded (Dowling 1986). McDermott et al. (2010), commenting on the poor pitch interval discrimination threshold they found, suggested that the importance of pitch as an expressive musical feature may rest more on an ability to detect pitch differences between tones, rather than an ability to recognise complex patterns of pitch intervals.

Some years ago, in a pilot experiment we noticed that an oddball interval (e.g. a tone pair separated by 7 semitones) did not pop out as expected within a randomly transposed series of standard intervals (e.g. 3 semitones). We subsequently ran a series of experiments in which we maintained a standard pitch contour, but varied the number of repetitions of the standard phrase (2 or 3), the number of tones in a phrase (2–6), the size of the deviance (1–3 semitones), and the tonality of the short melodies (Coath 2008). Most listeners, including those with musical education, found it very difficult to detect an oddball melodic phrase in a sequence of randomly transposed standard phrases, performing close to chance. The source of the surprising difficulty of the task was not clarified by this experiment, as the variables tested only weakly influenced performance. Here we report another attempt to discover what makes this task so hard.

Consistent with Gestalt grouping principles (Köhler 1947), auditory streaming experiments show that featural separation (such as pitch differences) promote segregation and conversely that featural similarity promotes integration (Bregman 1990; Moore and Gockel 2012). It is also known that within-stream (within-pattern) comparisons are far easier to make than between stream comparisons; (e.g. (Bregman 1990; Micheyl and Oxenham 2010)). Therefore, we hypothesized that if the standard pattern satisfied Gestalt grouping principles and could thus be more easily grouped, this would facilitate pattern comparisons, and that deviations within such patterns would be easier to detect. Another possibility is that confusion between within-pattern intervals and between-pattern transpositions may make individual patterns less distinctive, and so increase the task difficulty. Therefore, we also investigated the effects of transposition size and interactions between transposition size and within-phase intervals. Finally, the predictive coding account of perception (Friston 2005) suggests that the precision with which perceptual discriminations can be made is inversely related to stimulus variance, suggesting that task difficulty would increase with variance of standard phrase pitch intervals.

Our specific hypotheses were:

1. Small within-pattern intervals will promote grouping and thus improve performance (Gestalt proximity/similarity);
2. Small transpositions, especially when within-pattern intervals are large, may make individual patterns less distinctive, and thus impair performance;
3. Exact repetitions with no transposition will result in very good performance;
4. One exact repeat (i.e. pattern 1 = pattern 2) before introducing transpositions may allow a better pattern representation to be built and used as a template for subsequent patterns, and so improve task performance;
5. Smaller variance in the intervals within a pattern (either only small or only large intervals) will increase the predictability of the pattern and allow the formation of a more precise representation of the pattern. Therefore, task performance will decrease with increasing interval variance.
6. Musical training and experience will facilitate task performance.

2 Methods

The study was approved by the ethical review board of Plymouth University. Participants either received credits in a university course for their participation, or volunteered to take part.

2.1 Participants

Data were collected from 54 participants in total (32 females, 22 males; age range 19–65 years, median 20.5 years). The majority were undergraduate Psychology students at Plymouth University. Additional participants recruited from a doctoral programme and the University orchestra. All participants confirmed they had normal hearing. Details of musical training (years of formal tuition) and playing experience (years playing) were recorded for each participant. Four participants' data were excluded from the analysis as they achieved less than 30 % in at least one experimental block (chance level being 50 %), suggesting that they may not have understood the task correctly.

2.2 Materials

The experiment was conducted using a bespoke Matlab programme. Participants listened to the stimuli using Sennheiser HD215 headphones, individually adjusted to a comfortable sound level during the initial practice trial. The absolute level selected by each participant was not recorded.

2.2.1 Stimuli

Each trial consisted of four patterns separated by 700 milliseconds (ms) silence, and each pattern consisted of six tones. Three of the patterns had the same sequence of pitch intervals (standard pattern); the last pitch interval of either the final or the penultimate pattern of the trial deviated from the other three. A different standard pattern was delivered on each trial and no pattern was used more than once in the experiment. Patterns were generated by randomly selecting a set of five intervals, with the restrictions that each interval should only occur once within a pattern, and two intervals with same magnitude but opposite sign should not follow each other immediately in the sequence (to prevent the occurrence of repeated tones in the pattern).

All tones making up the pitch sequences were harmonic complexes, consisting of the first four harmonics of the nominal pitch, exponentially decreasing in amplitude (1:1/2:1/4:1/8) to give an oboe-like timbre. Tone duration was 110 ms, with 5 ms onset and offset linear ramps and 40 ms silence between tones, giving a tone onset to onset interval of 150 ms. Deviant intervals were always four semitones. Since standard pattern intervals were chosen from the set {1, 2, 3, 5, 6, 7 semitones}, depending on the condition (see Table 1), the difference between the standard and the deviant pitch interval was always 1, 2 or 3 semitones. The first tone of the first pattern always had a pitch of 450 Hz. To avoid the use of pitches which may not be clearly audible to everyone despite reporting normal hearing, all pitches were restricted to lie between 100 and 3200 Hz.

Table 1 Details of the within-pattern and transposition intervals used and the number of trials in each test block

Block	Within-pattern intervals	Transposition intervals	Number of trials
1	Big	None	10
2	Big	One exact repeat, then two big transpositions	10
3	Small	Small	20
4	Small	Big	20
5	Big	Small	20
6	Big	Big	20
7	All: 1, 2,3,5,6,7 ST	Big	20

The experiment consisted of one practice block and seven test blocks, each distinguished by the set of intervals used, as detailed in Table 1. Intervals were nominally divided into two sets: *small* {1, 2, 3} semitones, and *big* {5, 6, 7} semitones.

The practice block consisted of 10 trials. The first four were very easy with no transpositions and small within pattern intervals. The next four were slightly harder with two exact repeats of the pattern before two transpositions, with small within-pattern intervals and small transpositions. The final two examples were similar to trials in block 3 with small within-pattern intervals and small transpositions. Participants were given feedback after each trial (the response button briefly turned green for correct and red for incorrect) and a final practice score.

2.2.2 Procedure

Participants were required to indicate using two on-screen response buttons (labelled '2nd Last' and 'Last') whether the penultimate or last pattern was different from the rest. They were told that any difference was in the last interval of the pattern.

Participants began by entering their personal details and then continued with the practice block. They were encouraged to repeat the practice block as many times as they needed to familiarize themselves with the task; 1–3 repetitions were judged to suffice in all cases.

Following the practice block, participants were presented with seven test blocks, with no feedback. Blocks as detailed in Table 1 were presented in random order. Once they had completed all the test blocks, participants were presented with a bar graph showing their score in each block. Each 20-trial block took 3–4 min to complete and the experiment lasted roughly 30 min.

2.2.3 Analysis

In all cases confidence was assessed at the.05 level. Score distributions in each test block were compared against chance using the t-test. The effect of block was assessed using a 1-way ANOVA with all test blocks. The effect of transposition was assessed by contrasting block 1 with the average of blocks 5 and 6. The effect of one exact repetition was assessed by contrasting block 2 with block 6. The effect of variance in interval range was assessed by contrasting block 7 with the average of blocks 4 and 6. The effects of within-phrase intervals and between-phrase transpositions on performance were assessed using a two way ANOVA on data from blocks 3–6. The effect of interval variance was also tested using correlation analysis on data from blocks 3–7. The effect of final interval size of performance was tested using correlation analysis on data from all test blocks. The influence of musical experience was tested using correlation analysis on data from all test blocks. Correlation analysis was carried out using Spearman's correlation coefficient as the data were not normally distributed.

3 Results

Figure 1 shows the score distributions for each block for the participants.

Performance in all blocks was found to be significantly different from chance (shown by dotted line in the figure; $p<0.05$).

There was a main effect of block ($F(6,294)=41.61$, $p<0.001$, $\varepsilon=0.790$, partial $\eta^2=0.459$). The effect of transposition (contrasting block 1 with the average of blocks 5 and 6) was significant ($t=-10.36$, $p<0.001$). The effect of one exact repetition (contrasting block 2 with block 6) was not significant ($t=0.59$, $p=0.559$). The effect of variance in interval range (contrasting block 7 with the average of blocks 4 and 6) was significant ($t=3.20$, $p=0.002$). The more detailed trial-level correlation analysis showed performance correlated negatively with the variance of the pattern intervals (correlation coefficient $=-0.336$, $p<0.001$). There was no significant correlation between the magnitude of the final interval and performance (correlation coefficient -0.298, $p=0.147$). Posthoc multiple comparison analysis showed performance for musically important final intervals (perfect fourth and fifth, 5 and 7 semitones, respectively) was significantly lower than that for 1 semitone.

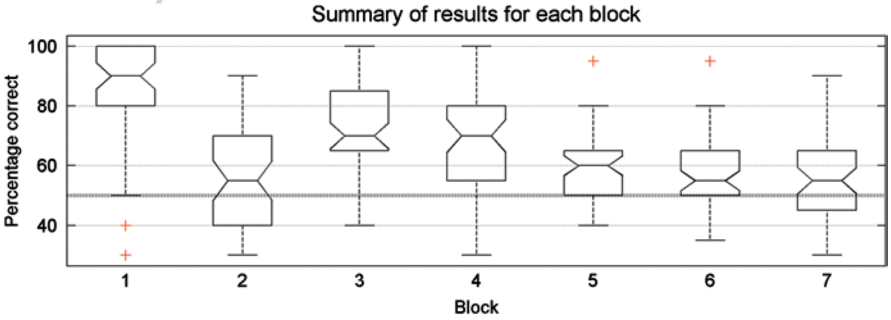

Fig. 1 Distribution of percentage correct scores in each block for all participants

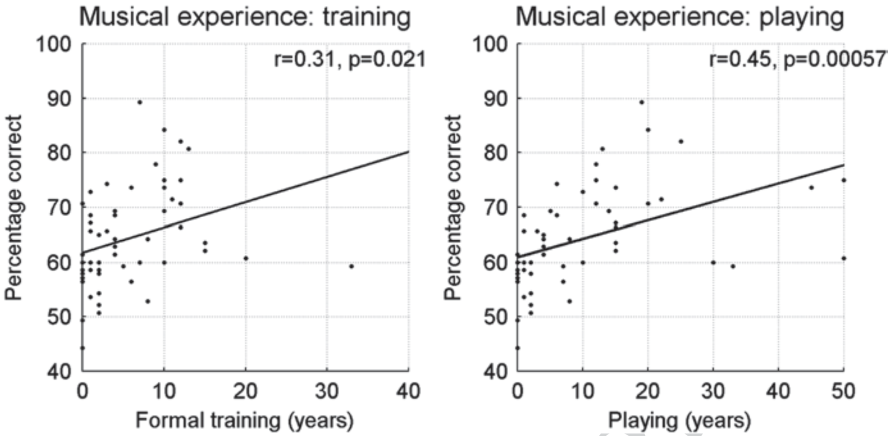

Fig. 2 The influence of musical experience on task performance

The two-way ANOVA assessing the effect of within-pattern and transposition intervals showed a significant main effect of within-pattern intervals ($F(1,49)=37.45$, $p<0.001$, partial $\eta^2=0.433$) and transposition size ($F(1,49)=12.16$, $p=0.001$, partial $\eta^2=0.199$) but their interaction was not significant ($F(1,49)=1.45$, $p=0.235$, partial $\eta^2=0.029$). A posthoc multiple comparison analysis showed that there was a tendency for large transpositions to impair performance more than small transpositions.

Performance correlated positively with musical experience; years of formal training (correlation coefficient $=0.342$, $p=0.015$), as well as years of playing (correlation coefficient $=0.435$, $p=0.002$).

The influence of musical training on task performance is illustrated in Fig. 2.

4 Discussion

In this study we investigated some of the potential sources of difficulty in detecting a pattern with a deviant pitch interval amongst transposed repetitions of a standard pattern, a task that is assumed to depend on relative pitch perception. Our results are consistent with a number of previous studies, e.g. (McDermott and Oxenham 2008), showing that relative pitch perception may be more limited than is commonly assumed. Performance is best when the standard phrase is repeated exactly with no transpositions (block 1), but falls substantially when transpositions are introduced (block 1 versus the average of blocks 3–7). Without transpositions, the task can be performed by direct comparisons between pitches, rather than using the interval relationships between successive pitches. Performance is not helped by one exact repetition of the standard pattern (block 2 versus block 6). This shows that although listeners may become sensitive to a repeating pattern after only 1.5

repetitions (Barascud 2014), they are unable to use this pattern for comparison with transposed versions of the pattern.

When patterns are transposed, then performance is best for standard patterns consisting of small intervals. This is consistent with the notion that grouping is promoted by featural similarity, and that representations of phrases consisting of small intervals are more easily formed, suggesting that comparisons between patterns may be facilitated by having a more coherent representation of the standard. With transpositions, large within-pattern intervals make the task very difficult. However, contrary to our hypothesis, large transpositions impaired performance more than small transpositions. This suggests that comparisons between pitch interval patterns are facilitated by proximity in pitch space. Increasing the variance in the pattern intervals, as predicted, impairs performance.

The idea that relative pitch perception depends solely on detecting a pattern of invariant pitch intervals is not supported by our results. Although the invariant property of the patterns in each trial is the sequence of pitch intervals defining the standard, listeners often could not use this information in the current experiment. Our results are compatible with the notion that in constructing object representations, the tolerance of the representation is a function of the variance in the pattern, i.e. increasing variance in object components lead to more permissive representations. This makes sense when the general problem of perceptual categorisation is considered; e.g. the variability of the spoken word.

Relative pitch perception has been likened to translation invariant object recognition in vision (Kubovy and Van Valkenburg 2001). Interestingly the literature on visual perceptual learning has shown that learning can be surprisingly specific to the precise retinal location of the task stimulus (Fahle 2005). The most influential model of translation invariant object recognition is the so-called trace model (Stringer et al. 2006), which assumes that this ability actually depends on learning the activity caused by the same stimulus being shown at many different locations; invariant recognition then emerges at a higher level by learning that these different activations are caused by the same object. Perhaps this is what happens when we learn a tune. The categorisation of the tune depends on hearing it at many different pitch levels within a context that provides clear links between the various repetitions (e.g. within the same piece of music, or same social context).

Acknowledgments We would like to thank Dávid Farkas for help with the statistical analysis. IW was supported by the Hungarian Academy of Sciences (Lendület project, LP36/2012). GPH was supported by a post-doctoral research fellowship of the Hungarian Academy of Sciences.

References

Barascud N (2014). Auditory pattern detection. PhD thesis, University College London (University of London), London

Bregman AS (1990) Auditory scene analysis: the perceptual organization of sound. MIT Press, Cambridge

Chait M, Poeppel D, de Cheveigne A, Simon JZ (2007) Processing asymmetry of transitions between order and disorder in human auditory cortex. J Neurosci 27(19):5207–5214. doi:10.1523/JNEUROSCI.0318-07.2007

Coath M, Háden GP, Winkler I, Denham SL (2008). Are short tone sequences memorable? Paper presented at the British Society of Audiology Annual Conference, Leeds

Dowling WJ (1986) Context effects on melody recognition: scale-step versus interval representations. Music Percept 3(3):281–296

Fahle M (2005) Perceptual learning: specificity versus generalization. Curr Opin Neurobiol 15(2):154–160. doi:10.1016/j.conb.2005.03.010

Foster NE, Zatorre RJ (2010) A role for the intraparietal sulcus in transforming musical pitch information. Cereb Cortex 20(6):1350–1359. doi:10.1093/cercor/bhp199

Friston K (2005) A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci 360(1456):815–836. doi:10.1098/rstb.2005.1622

Griffiths TD, Warren JD (2004) What is an auditory object? Nat Rev Neurosci 5(11):887–892. doi:10.1038/nrn1538

Köhler W (1947) Gestalt psychology: an introduction to new concepts in modern psychology. Liveright Publishing Corporation, New York

Kubovy M, Van Valkenburg D (2001) Auditory and visual objects. Cognition 80(1–2):97–126

McDermott JH, Oxenham AJ (2008) Music perception, pitch, and the auditory system. Curr Opin Neurobiol 18(4):452–463. doi:10.1016/j.conb.2008.09.005

McDermott JH, Keebler MV, Micheyl C, Oxenham AJ (2010) Musical intervals and relative pitch: frequency resolution, not interval resolution, is special. J Acoust Soc Am 128(4):1943–1951. doi:10.1121/1.3478785

Micheyl C, Oxenham AJ (2010) Objective and subjective psychophysical measures of auditory stream integration and segregation. J Assoc Res Otolaryngol 11(4):709–724. doi:10.1007/s10162-010-0227-2

Moore BC, Gockel HE (2012) Properties of auditory stream formation. Philos Trans R Soc Lond B Biol Sci 367(1591):919–931. doi:10.1098/rstb.2011.0355

Stefanics G, Haden GP, Sziller I, Balazs L, Beke A, Winkler I (2009) Newborn infants process pitch intervals. Clin Neurophysiol 120(2):304–308. doi:10.1016/j.clinph.2008.11.020

Stringer SM, Perry G, Rolls ET, Proske JH (2006) Learning invariant object recognition in the visual system with continuous transformations. Biol Cybern 94(2):128–142. doi:10.1007/s00422-005-0030-z

Tew S, Fujioka T, He C, Trainor L (2009) Neural representation of transposed melody in infants at 6 months of age. Ann N Y Acad Sci 1169:287–290. doi:10.1111/j.1749-6632.2009.04845.x

Winkler I, Denham SL, Nelken I (2009) Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends Cogn Sci 13(12):532–540. doi:10.1016/j.tics.2009.09.003