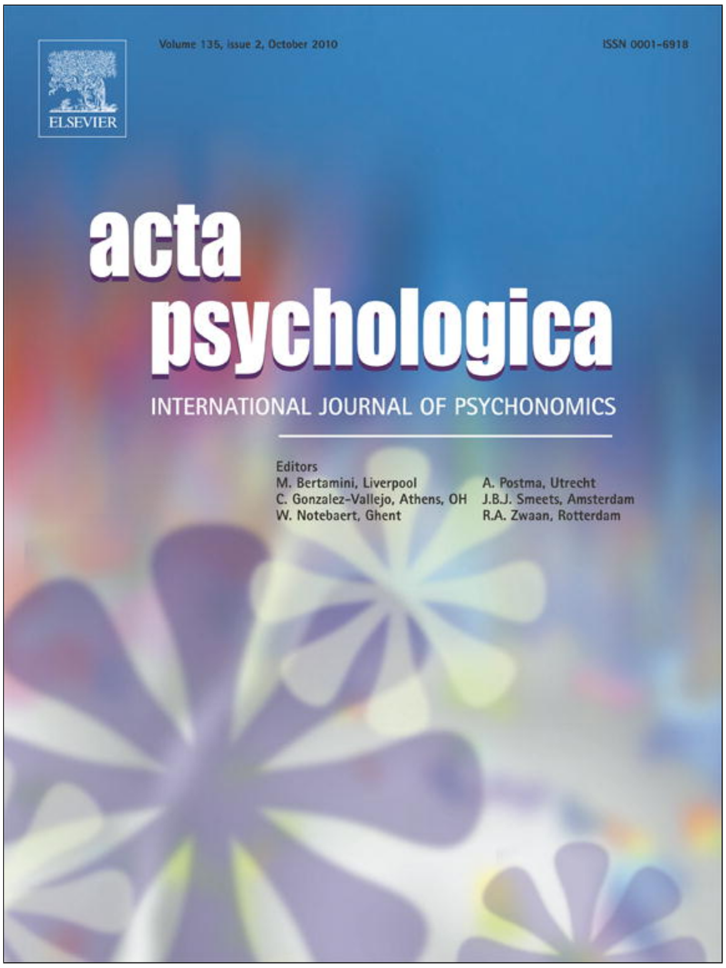
Volume 135, issue 2, October 2010
ISSN 0001-6918
ELSEVIER
acta psychologica
INTERNATIONAL JOURNAL OF PSYCHONOMICS
Editors
M. Bertamini, Liverpool
C. Gonzalez-Vallejo, Athens, OH
W. Notebaert, Ghent
A. Postma, Utrecht
J.B.J. Smeets, Amsterdam
R.A. Zwaan, Rotterdam

Acta Psychologica 135 (2010) 182–190

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Learning melodies from non-adjacent tones

Ansgar D. Endress

Massachusetts Institute of Technology, Room 46-4127, 43 Vassar St, 02139 Cambridge, MA, USA

ARTICLE INFO

Article history:
Received 31 August 2009
Received in revised form 1 March 2010
Accepted 11 June 2010

PsycINFO classification:
2340
2343
2720
2820

Keywords:
Statistical learning
Melody perception
Language acquisition
Transitional probabilities
Tonality

ABSTRACT

Language acquisition might heavily rely on statistical learning mechanisms. This has led many researchers to investigate the computational constraints that limit such learning. In particular, it has been argued that statistical relations among non-adjacent items cannot be tracked, as this might lead to a "computational explosion" making statistical learning intractable. In line with this view, previous research suggests that listeners cannot track relations among non-adjacent musical tones (Creel, Newport, & Aslin, 2004). Here I show that participants readily track non-adjacent tone relations when these are implemented in a musically meaningful way. Specifically, participants readily track non-adjacent tone relations in tonal melodies, but find it more difficult to track non-adjacent tone relations in random melodies, suggesting that non-adjacent relations are easier to track when listeners face "ecological", musically meaningful stimuli.

Statistical learning might play an important role in cognition, and especially during language acquisition. There is overwhelming evidence that human adults and infants, as well as other animals, have powerful statistical learning abilities (e.g., Aslin, Saffran, & Newport, 1998; Fiser & Aslin, 2001, 2002; Hauser, Newport, & Aslin, 2001; Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996; Saffran, Johnson, Aslin, & Newport, 1999; Toro & Trobalón, 2005). However, the very power of these abilities might be their Achilles' heel, as unconstrained, overly powerful learning mechanisms have the potential to learn one of the many spurious regularities in their input rather than the regularities they ought to learn. Powerful learning mechanisms thus have to be constrained to be of any use.

In the domain of statistical learning, one plausible constraint may be the complexity of the statistics that can be extracted. This can be illustrated by one of the most prominent uses of statistical learning in language acquisition, namely speech segmentation. Specifically, children have to learn the words of their native language from fluent speech, where acoustic cues to word boundaries are highly unreliable. Hence, before children can start acquiring the words of their native language, they have to cut the speech signal into its constituent words, a problem that is known as the segmentation problem. One prominent cue to word boundaries is co-occurrence statistics among syllables such as transitional probabilities (TPs)[1]; these reflect the possibility that syllables within words are more predictive of each other than syllables in different words. Hence, dips in TPs might be used as cues to word boundaries.

While it is well documented that human adults and infants can use TPs between adjacent syllables to segment fluent speech, many linguistic dependencies arise between non-adjacent items. The English progressive, for example, is constructed by a form of "to be", the "ing" suffix and an intervening verb-root; the regularity between the auxiliary and the "ing" suffix is, therefore, non-adjacent. Infants seem to know such relations at 18 months of age (Santelmann & Jusczyk, 1998).

Despite the importance of non-adjacent relations in language, the proposal that learners may be sensitive to TPs between non-adjacent syllables (e.g., Peña, Bonatti, Nespor, & Mehler, 2002) has met with considerable resistance, because some authors believe that learning non-adjacent dependencies would lead to a computational explosion of the learning problem (e.g., Newport & Aslin, 2004). This assumption, however, relies on a specific model of how TPs are computed: if they are stored in a table-like structure (whose entries indicate that, say, [pu] goes to [ki] with probability p), learners would need to construct a new table for each order of TPs they consider; this would lead to a linear

[1] Formally, TPs are conditional probabilities on syllable sequences. For two syllables σ_1 and σ_2, for example, $TP(\sigma_1 \rightarrow \sigma_2) = \frac{p(\sigma_1 \sigma_2)}{p(\sigma_1)}$, where the probabilities are estimated by the corresponding frequencies.

E-mail address: ansgar.endress@m4x.org.

0001-6918/$ – see front matter
doi:10.1016/j.actpsy.2010.06.005

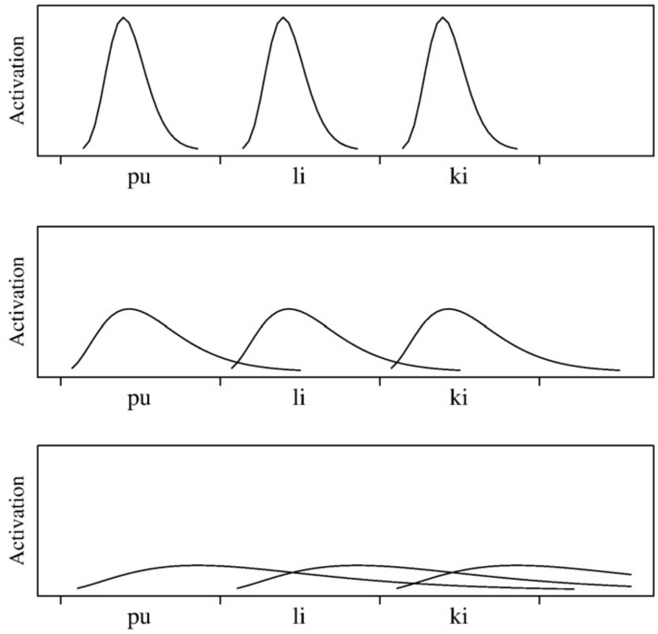

Fig. 1. Responses of three hypothetical "grand-mother" neurons responding to the syllables [pu], [li] and [ki]. (Top) Upon hearing the non-sense sequence [puliki], the neurons fire. As their activation decays quickly, the three neurons are never simultaneously active; hence, they cannot form any associations. (Middle) The activation of each neuron decays more slowly; as a result, the neurons representing adjacent syllables fire simultaneously and can form associations by some form of correlational learning. (Bottom) The responses of the neurons decay even more slowly. Hence, the neuron representing [pu] is active together with the neuron representing [ki]; this allows the neurons to form associations between non-adjacent syllables.

growth of the problem's complexity with the order of TPs. (More specifically, the memory space taken up by such tables might scale roughly linearly with the order of TPs; in contrast, the computational complexity of finding an entry in the table might well grow more slowly, depending on the algorithms employed.) However, Fig. 1 illustrates a simpler, alternative model of learning non-adjacent relations. Syllable representations might form associations with the syllable representations they are simultaneously active with, for example by some form of correlational learning (e.g., Hebbian learning, possibly with some renormalization mechanism), according to the old adage that "what fires together wires together". If syllable activations last longer than one syllable duration, one would expect them to form associations with non-adjacent syllables. By this account, associations among non-adjacent syllables would likely be weaker than association among adjacent syllables (because, due to decay, the product of their simultaneous activations would be smaller), but a sensitivity to non-adjacent TPs would certainly not lead to a computational explosion.

Evidence for the ability to learn non-adjacent statistical relations has been conflicting. While some researchers observed a sensitivity to TPs between non-adjacent syllables (e.g., Ebbinghaus, 1885/1913; Endress & Bonatti, 2007; Endress & Mehler, 2009; Gómez, 2002; Onnis, Monaghan, Richmond, & Chater, 2005; Peña et al., 2002), others did not (e.g., Newport & Aslin, 2004). Most relevant to the current paper, participants seem to experience difficulties tracking non-adjacent relations among tones in continuous tone sequences (Creel, Newport, & Aslin, 2004). While human adults readily track statistical relations among adjacent tones (e.g., Saffran et al., 1999), Creel et al. (2004) found them unable to track non-adjacent tone statistics unless the non-adjacent relations were made salient through other perceptual factors. Specifically, if tones were played in the same frequency range and with the same timbre, participants were unable to track non-adjacent relations among them. However, when every second tone was in another pitch register, or if tones had alternating timbres, participants tracked the "non-adjacent" relations. However, due to "auditory streaming", tones played in different pitch registers or with different timbres are perceived as coming from different sound sources (Bregman, 1990); in other words, the "non-adjacent" tones were perceptually adjacent due to auditory streaming principles, providing no evidence for non-adjacent learning.

While these results suggest that non-adjacent tone relations are difficult to track, this conclusion seems to clash with intuitions from everyday music experience suggesting that listeners do track non-adjacent tone relations. Melodies spanning non-adjacent tones are ubiquitous in music, for example in variations and ornaments. For instance, a transition between two tones can be ornamented by a turn (which involves inserting other notes between the original tones); if listeners were unable to track non-adjacent melodies, a turn should disrupt the melody. Since similar phenomena are widespread in music, one would expect listeners to be able to track such "interrupted" melodies.

Previous research suggests that people can recognize interrupted melodies at least when these are familiar. Dowling (1973) showed that they can recognize melodies interleaved with distracter tones when the target melodies are familiar and pre-specified (see also Dowling, Lung, & Herrbold, 1987). However, as the current experiments investigate the recognition of unfamiliar melodies interleaved with distracter tones (and participants were not informed about the underlying melodies either), it is unclear whether participants can use non-adjacent tone relations for learning *new* melodies.

These observations raise the question of why Creel et al. (2004) did not observe any sensitivity to non-adjacent relations unless using streaming cues, while Dowling's (1973) and Dowling et al.'s (1987) results (as well as everyday experience) appear to suggest that listeners do have such a sensitivity. One possible reason is that ornaments and so on typically occur in tonal melodies, and not in random melodies such as those used by Creel et al. (2004). (Tonal melodies are used in "classical" or popular music. Atonal melodies, in contrast, may seem melodically somewhat random to most listeners.) Indeed, it is well established that tonal structure has an important influence on how melodies are memorized (e.g., Croonen, 1994; Dowling, 1991; Dowling, Kwak, & Andrews, 1995; Saffran, 2003). I thus ask whether using tonal melodies would allow participants to track non-adjacent tone relations even in the absence of streaming cues.

Participants were familiarized with a continuous concatenation of three six-tone "words" whose design is shown in Fig. 2. The tone words were constructed by interspersing musically meaningful three-tone sequences (such as major triads) with three distracter tones; musically speaking, the underlying tone triplets were "ornamented". In each six-tone word, the first, the third and the sixth tone came from the underlying tone triplets, while the other tones were "ornaments". In Experiments 1 to 3, the ornaments were musically meaningful, because most of them just "connected" the tones of the underlying tone triplets. That is, the pitch of the ornament tones bridged that of the surrounding tones from the underlying triplet; as a result, the melodic lines in the ornamented tone words were "smoother" than in

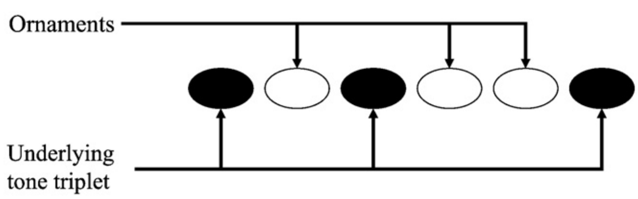

Fig. 2. Stimuli in Experiments 1 to 4. Participants were familiarized with a continuous concatenation of three six-tone words, each repeated 70 times. Each tone word (represented by the six ellipses) was constructed from an underlying three-tone sequence interspersed with three other tones (the "ornaments"). The first, the third and the sixth tone came from the underlying tone triplet, while the other tones were the ornaments. Following this familiarization, participants were tested on the underlying tone triplets and test items derived from these triplets.

the underlying tone triplets. Such "passing tones" are extensively used both for improvising ornaments and for composing melodies (e.g., Linde, 1958).

In Experiments 1 to 3, participants were familiarized with a continuous concatenation of 70 repetitions of either a major version of the tone words, or of a minor version of the tone words. Following this familiarization, participants had to choose between two types of three-tone items. In Experiment 1, they had to choose between major and minor versions of the underlying tone triplets, respectively. If they can track non-adjacent tone relations, they should prefer the version corresponding to the familiarization melody they had heard. Experiment 2 controlled for the possibility that participants might just track specific pitch classes by transposing the test items relative to the familiarization melody. Experiment 3 ruled out that participants simply tracked the mode of the familiarization melody (i.e., major or minor) by asking them to choose between test items that comprised the same tones but in different orders.

In Experiment 4, participants were familiarized with a concatenation of tone words composed of the same underlying tone triplets as in Experiments 1 to 3, but using different, arbitrary ornaments; if participants relied on the tonal structure of the melodies in Experiments 1 to 3 to track non-adjacent melodies, they should experience more difficulties in Experiment 4 than in the other experiments.

There is an important caveat relating to what participants learn in such experiments. Experiments such as Creel et al.'s (2004) are usually described in terms of TPs among tones. However, there is no evidence that such TPs can actually be learned (except in infants; see Saffran & Griepentrog, 2001). In fact, computing TPs among tones requires identifying the tones, and thus absolute pitch. Human adults, however, generally process melodies in terms of intervals (that is, frequency ratios) or even contours (that is, sequences of ups and downs), but not in terms of their absolute identity (although even non-musicians seem to have some form of perfect pitch for highly familiar melodies; see Halpern, 1989; Schellenberg & Trehub, 2003). In line with these results, experiments that have explicitly tested the contributions of relative vs. absolute pitch for statistical processing have found that non-musicians failed to show any sensitivity to TPs among tones when presented with random melodies, while they have a moderate sensitivity to absolute pitch cues when tested with simplified tonal melodies (Saffran & Griepentrog, 2001; Saffran, 2003).

Moreover, it is well known that melodies are perceived in a hierarchically organized fashion (e.g., Jackendoff & Lerdahl, 2006; Lerdahl & Jackendoff, 1983). However, as has been pointed out more than half a century ago (Lashley, 1951), TP-like mechanisms that encode sequences by forming associations among neighboring sequence elements are unlikely to account for hierarchical structure, raising the possibility that TPs might not be used in melodic perception either. As a result, the aforementioned "TPs among tones" might be better described in terms of interval sequences. For a better comparison with the previous literature, however, I will present the results in terms of TPs among tones, but one should keep in mind that in the experiments presented below (and the previous literature) participants probably learned adjacent or non-adjacent interval sequences.

1. Experiment 1

1.1. Materials and method

1.1.1. Participants

Twenty native speakers of Italian, recruited at SISSA, Trieste (Italy), took part in this experiment (11 females, 9 males; mean age 24.5, range 18–33). They were randomly assigned to one of two mode conditions (see below). Participants were recruited irrespectively of their musical training.

To make sure that the results were not carried just by a few "musicians", I reanalyzed the data after excluding participants with more than three years of musical training (2, 4 and 2 participants in Experiments 1, 2 and 4, respectively; there were no "musicians" in Experiment 3). The results of these analyses were similar to those reported below; importantly, all significant results remain significant (and all non-significant results remain non-significant) after removal of the musicians. I will thus report just the overall analyses including the entire sample of participants.

1.1.2. Materials and apparatus

The stimuli were created using text files in the ABC music notation language that were converted to midi files using abc2midi (http://abc.sourceforge.net/abcMIDI/); the latter were converted to wav files using timidity++ (http://timidity.sourceforge.net) and finally to aiff files using sox (http://sox.sourceforge.net/). To avoid direct cues to tone word onsets, streams were ramped in and out for 10 s using audacity (http://audacity.sourceforge.net). Tone duration was 333 ms (180 bpm). The experiments were run using PsyScope X (http://psy.ck.sissa.it). Stimuli were presented over headphones. Responses were collected on pre-marked keys on the keyboard.

1.1.3. Pre-training

Before starting the experiment, participants completed a pre-training phase to get familiarized with the response keys. The pre-training consisted of 10 trials. In each trial, participants heard two syllables, one of which was 'so.' Their task was to indicate whether 'so' was the first or the second syllable. 'So' was the first syllable in half of the trials, and the second one in the other half.

1.1.4. Familiarization

After the pre-training phase, participants were informed that they would listen to a melody, and were instructed to attend to it. Following this, they were randomly assigned to one of two melodies, one in major and the other one in minor; below, we will refer to these melodies as the "mode conditions". Both melodies consisted of concatenations of three six-tone "words." These tone words were constructed from three underlying three-tone sequences by interspersing the tones of the underlying tone triplets with three additional tones (the "ornaments"; see Fig. 2). The underlying and the ornamented versions of the tone words are shown in Table 1. In the major mode condition, for example, the ornamented version of the underlying tone triplets $C_4E_4G_4$ was constructed by inserting one tone between the first and the second tone, and two tones between the second and the third tone, yielding the ornamented tone word $C_4D_4E_4F_4F^{\#}_4G_4$. As mentioned above, the ornaments were mostly

Table 1
Stimuli used in Experiment 1. In the tone columns, letters and numbers stand for pitch classes and octave numbers, respectively (the middle C being C_4). Intervals starting with a capital and a lower-case M represent major and minor intervals, respectively. Tone words were constructed from three underlying three-tone sequences. Participants were familiarized with "ornamented" versions of these tone triplets comprising six tones, and then tested on the underlying tone triplets. Half of the participants were familiarized with the major version, and the remaining participants with the minor version.

Ornamented		Underlying	
Tones	Intervals	Tones	Intervals
Major			
$C_4D_4E_4F_4F^{\#}_4G_4$	M2↑M2↑m2↑m2↑m2↑	$C_4E_4G_4$	M3↑m3↑
$F_4G_4A_4F_4F^{\#}_4G_4$	M2↑M2↑M3↓m2↑m2↑	$F_4A_4G_4$	M3↑M2↓
$D_5A_4B_4A_4F^{\#}_4G_4$	M4↓M2↑M2↓m3↓m2↑	$D_5B_4G_4$	m3↓M3↓
Minor			
$C_4D_4E^{b}_4F_4F^{\#}_4G_4$	M2↑m2↑M2↑m2↑m2↑	$C_4E^{b}_4G_4$	m3↑M3↑
$F_4G_4A^{b}_4F_4F^{\#}_4G_4$	M2↑m2↑m3↓m2↑m2↑	$F_4A^{b}_4G_4$	m3↑m2↓
$D_5A^{b}_4B^{b}_4A^{b}_4F^{\#}_4G_4$	m5↓M2↑M2↓M2↓m2↑	$D_5B^{b}_4G_4$	M3↓m3↓

"passing tones" that "smoothed" the melodies. To create the familiarization melody, the resulting ornamented tone words were concatenated in random order with no silences between them, and with the constraint that the same tone word could not occur twice in a row. Each tone word was repeated 70 times, yielding a familiarization melody of 7 min.

As mentioned above, participants were exposed to one of two "mode conditions", corresponding to a major and a minor version of the familiarization melody, respectively. As shown in Table 1, the crucial difference between the major and the minor version was that the third tone in the minor tone words was a semitone lower than the third tone in the corresponding major tone words. In the tone word constructed from the underlying tone triplet $D_5B_4G_4$, the second and the fourth tone were also shifted by a semitone (see Table 1).

The TPs across tones are shown in Table 2. On average, first order TPs within tone words were 0.688, and first order TPs across tone words were 0.250. Second order TPs averaged to 0.659 within tone words and 0.292 across tone words, respectively. As a result, most first or second order TPs were higher within tone words than across tone words.

1.1.5. Test

After listening to the familiarization melody, participants were presented with pairs of the underlying tone triplets. In each trial, they were instructed to choose the one that they considered more familiar. All pairs comprised one major and one minor tone triplet. As a result, legal items for one group (e.g., the major group) were foils for the other group (e.g., the minor group) and vice versa. The nine possible pairings of major and minor underlying tone triplets were presented twice in random order, once with the major item first and once with the minor item first. Tones were played at half the rate of the familiarization tones (that is, at 90 bpm, or with a tone duration of 666 ms), equating the inter-stimulus interval between two tones from the underlying triplet in the familiarization melody and in the test items.

Table 2
First and second order transitional probabilities among tones in the major condition of Experiment 1. Letters and numbers stand for pitch classes and octave numbers, respectively (the middle C being C_4). Rows give the first tone in a transition and columns give the second tone. The transitional probabilities for the minor condition were very similar (maximal absolute difference from the probabilities given here: 0.005). First order transitional probabilities averaged to 0.688 within tone words and 0.250 across tone words, respectively. Second order transitional probabilities averaged to 0.659 within tone words and 0.292 across tone words, respectively.

From	To								
	C_4	D_4	E_4	F_4	$F^{\#}_4$	G_4	A_4	B_4	d_5
First order									
C_4	0	1.0	0	0	0	0	0	0	0
D_4	0	0	1.0	0	0	0	0	0	0
E_4	0	0	0	1.0	0	0	0	0	0
F_4	0	0	0	0	0.667	0.333	0	0	0
$F^{\#}_4$	0	0	0	0	0	1.0	0	0	0
G_4	0.251	0	0	0.251	0	0	0.251	0	0.247
A_4	0	0	0	0.333	0.333	0	0	0.333	0
B_4	0	0	0	0	0	0	1.0	0	0
D_5	0	0	0	0	0	0	1.0	0	0
Second order									
C_4	0	0	1.0	0	0	0	0	0	0
D_4	0	0	0	1.0	0	0	0	0	0
E_4	0	0	0	0	1.0	0	0	0	0
F_4	0	0	0	0	0	0.667	0.333	0	0
$F^{\#}_4$	0.335	0	0	0.335	0	0	0	0	0.330
G_4	0	0.251	0	0.251	0	0.251	0.247	0	0
A_4	0	0	0	0	0.333	0.333	0.333	0	0
B_4	0	0	0	0	1.0	0	0	0	0
D_5	0	0	0	0	0	0	0	1.0	0

1.1.6. Analysis

T-tests were calculated after transformation with the empirical logistic transform (e.g., Jaeger, 2008) using a bias of 0.5 and tested against a chance level of 0 (corresponding to a proportion of correct responses of 50%). The significance of the resulting *t* values was assessed both using the *t* distribution and a one-sample permutation test associated with the *t* statistic.

I used logistic mixed effects models to compare conditions and/or experiments (Pinheiro & Bates, 2000; see Baayen, Davidson, & Bates, 2008, for an informal introduction), using the lme4 package (Bates & Maechler, 2008) for the R software package (version 2.6.1, http://www.R-project.org). Neither random intercepts for the test pairs, nor for whether or not the first test item of a pair was the correct choice added significantly to the likelihood of the models (as assessed by a likelihood ratio test). As a consequence, I kept only a random intercept for participants.

All statistical tests are two-tailed with a chance level of 0.05.

1.2. Results

As shown in Fig. 3, participants preferred legal items to foils with the opposite mode ($M=59.4\%$, $SD=11.6\%$), $t\ (19)=3.6$, $p=0.002$, Cohen's $d=0.81$, $p_{permut}=0.003$. (As mentioned above, t-test are computed on data transformed using the empirical logistic function.) A mixed effects model with mode group as fixed effect predictor and a random intercept for participants revealed no difference between the mode conditions, $Z=1.5$, $p=0.135$, ns.

1.3. Discussion

Before accepting the conclusion that participants were sensitive to the relations among non-adjacent tones, one needs to rule out a possible confound. In the test pairs in Experiment 1, the middle tone in the foils had never occurred in the familiarization melody; if participants kept track of all the tones they had heard, they might prefer legal items simply for this reason.

A considerable amount of research makes this possibility unlikely. Indeed, recognizing that a particular tone has occurred during familiarization would require "perfect pitch," that is, the ability to

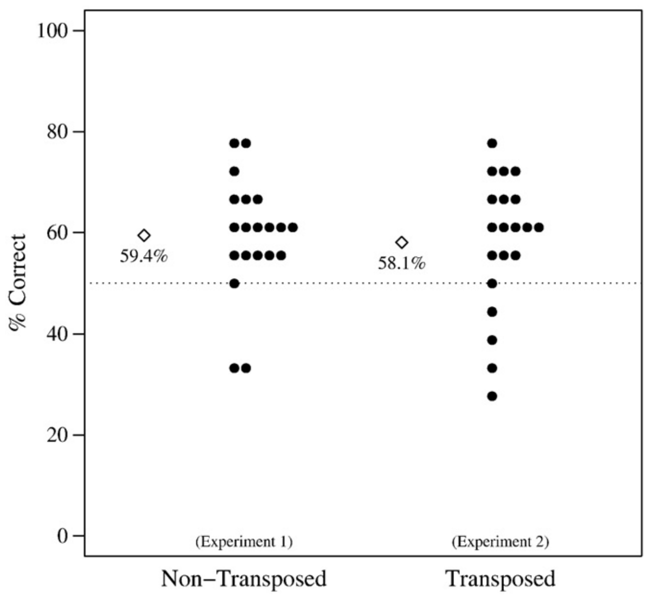

Fig. 3. Results of Experiments 1 and 2. Dots represent averages of individual participants, diamonds represent population averages, and the dotted line represents the chance level of 50%. Participants successfully tracked non-adjacent tone relations both when the test items were played on the original pitch levels (Experiment 1) and when they were transposed (Experiment 2).

identify tones by their fundamental frequency. However, at least human adults encode melodies in relative terms (e.g., in statistical learning experiments, Saffran & Griepentrog, 2001; Saffran, Reeck, Niebuhr, & Wilson, 2005), while the strategies used by infants are less clear. In other words, rather than encoding melodies as sequences of fundamental frequencies, they encode them as sequences of frequency *ratios*. This ability allows listeners to recognize a song when sung by a soprano or a basso, as the frequency ratios remain the same.

Still, even non-musicians appear to have perfect pitch for highly familiar songs (e.g., Halpern, 1989; Schellenberg & Trehub, 2003); in order to rule out any such confound, Experiment 2 controls for this possibility by transposing test items by a semitone upward in the major condition, and downward in the minor condition. As a result, the middle tones in the *foils* occurred during the familiarization melody, but not the middle tones in the legal items. Hence, if the participants' choices were based on the recognition of particular pitches, they should choose the foils rather than legal items.

2. Experiment 2

2.1. Materials and method

2.1.1. Participants

Twenty new native speakers of Italian, recruited at SISSA, Trieste (Italy), took part in this experiment (10 females, 10 males; mean age 22.4, range 19–34). They were randomly assigned to the two mode conditions.

2.1.2. Familiarization

The familiarization phase of Experiment 2 was identical to that in Experiment 1.

2.1.3. Test

The test items used in Experiment 2 were similar to those used in Experiment 1 except that they were transposed by a semitone. Specifically, in the major condition, the test items were shifted *upward* by a semitone. As a result, the middle tones in (minor) foils had occurred in the (major) familiarization melody, while the middle tones in the legal items had *not* occurred during the familiarization melody. In the minor condition, in contrast all test items were transposed downward by a semitone; as in the major condition, only the middle tones of foils but not those of legal items had occurred during the familiarization melody.

2.2. Results

Participants preferred legal items to foils ($M=58.1\%$, $SD=13.4\%$), $t\ (19)=2.7$, $p=0.015$, Cohen's $d=0.6$, $p_{permut}=0.019$. A mixed effects model with mode group as fixed effect predictor and a random intercept for participants revealed no difference between the mode conditions, $Z=0.7$, $p=0.498$, ns. A mixed effects model with mode group and experiment (i.e., Experiment 1 vs. Experiment 2) as fixed effect predictors and a random intercept for participants revealed no main effect or interaction.

2.3. Discussion

The results of Experiments 1 and 2 suggest that participants can keep track of relations among non-adjacent tones when tested in a musically meaningful way. While these results are not particularly strong, they are comparable to those reported for moderate streaming cues by Creel et al. (2004), and reliable since the results of Experiment 2 replicate those of Experiment 1.

A prior preference for major or minor cannot account for these results, as legal items for the major group were foils for the minor group and vice versa. Likewise, participants did not simply base their choices on the recognition of particular tones; if they had followed such a strategy, they should have chosen foils rather than legal items in Experiment 2.

Table 3
Counts of the intervals used in Experiment 1. Intervals starting with capital and lowercase M represent major and minor intervals, respectively.

	Major		Minor	
	Ornamented	Underlying	Ornamented	Underlying
M2↓	1	1	2	0
M2↑	5	0	4	0
m2↓	0	0	0	1
m2↑	6	0	7	0
M3↓	1	1	0	1
M3↑	0	2	0	1
m3↓	1	1	1	1
m3↑	0	1	0	2
M4↓	1	0	0	0
m5↓	0	0	1	0

Neither could participants have succeeded by tracking particular intervals (that is, frequency ratios). Indeed, Table 3 shows that all intervals are equally represented in major and minor test items, respectively, with two exceptions: the major test items contain one major upward third more than the minor test items, while the minor test items contain one minor upward third more. However, neither interval occurred during the familiarization melodies. Likewise, the major test items contain a major downward second, while the minor test items contain a minor downward second; in the familiarization items, in contrast, the *minor* tone words contained one additional major downward second compared to major items, while neither major nor minor items contained any minor downward seconds. Hence, if participants simply tracked specific intervals, they should have favored foils in the minor condition, and be at chance in the major condition. In contrast to this prediction, they preferred legal items in both conditions. Hence, a plausible conclusion is that participants preferred legal items due to a sensitivity to non-adjacent tone relations.

(As mentioned before, non-musicians are unlikely to track TPs among tones, because this would require perfect pitch; moreover, sequence encoding mechanisms similar to TPs are unlikely to allow for hierarchical structure (e.g., Lashley, 1951), raising the question of how TPs can be reconciled with the hierarchical processes involved in melody perception (e.g., Jackendoff & Lerdahl, 2006; Lerdahl & Jackendoff, 1983). The results presented here should, therefore, not be taken as evidence that participants can track TPs among non-adjacent tones; they might have tracked intervals among non-adjacent tones rather than TPs, but the important point is that they successfully tracked melodic properties that were not limited to adjacent tones.)

While the combined results of Experiments 1 and 2 seem to suggest that listeners can track relations among non-adjacent tones, there are several alternative accounts according to which participants might not have learned any tone relations at all. First, upon hearing the familiarization melodies, participants might simply have recognized the underlying, interleaved melodies, because they were likely to have heard them before (e.g., a major triad). While this account still requires participants to be sensitive to non-adjacent tone relations (they could not perceive the interleaved melodies otherwise), previous research makes it unlikely. In Dowling's (1973) experiments, for example, where participants had to track interleaved melodies, they perceived (non-adjacent) target melodies only when these were pre-specified; once they knew which melody to watch out for, they seemed to selectively attend to the relevant parts of the overall, adjacent melody (see also Bregman (1990), chapter 4, for more examples where explicit instructions or extensive training led participants to attend predominantly to parts of the overall melody).

In the current experiments, in contrast, participants were not informed about the structure of the melodies; as a result, they had no reason to listen specifically to the underlying tone triplets while ignoring the ornaments. This conclusion is reinforced by the fact that the adjacent, ornamented melodies were at least as well-formed as the underlying non-adjacent tone triplets, as the "smoother" intervals used in the ornamented tone words are at least as frequent in music as the larger intervals used in the underlying tone triplets (Youngblood, 1958). It thus seems unlikely that participants just attended to the underlying tone triplets, and ignored the ornaments. Rather, they seem to have attended to the overall, ornamented melodies. Despite perceiving an overall melody among adjacent tones, participants were thus sensitive to non-adjacent tone relations in the absence of any streaming cues.

However, there is a second alternative explanation of the results of Experiments 1 and 2. In the test phases of both experiments, participants always had to choose between major and minor items. As a result, if they just tracked the mode of the familiarization melodies, they might simply have matched the mode of the test items to that of the familiarization melody, possibly without any sensitivity to non-adjacent tone relations at all. This possibility is addressed in Experiment 3.

3. Experiment 3

While the results of Experiments 1 and 2 seem to suggest that participants successfully tracked relations among non-adjacent tones, there is an alternative explanation that might not require such an ability. Possibly, participants simply remembered whether the familiarization sequences were in major or in minor, and selected those test items that matched the mode they had been familiarized with.

Experiment 3 addresses this possibility. After the same familiarization as in Experiment 1, participants had to choose between items that were either both in major, or both in minor. Specifically, participants in the major conditions had to choose between the major test items from Experiment 1, and foils that comprised the same tones as the major test items but in different orders; that is, the foils were permutations of the major test items. Likewise, in the minor condition, participants had to choose between the minor test items from Experiment 1 and permutations of these test items.

If participants just tracked the mode of the familiarization melodies in Experiments 1 and 2 without being sensitive to non-adjacent tone relations, they should fail in Experiment 3, since they have to choose between items in the same mode. In contrast, if they tracked non-adjacent tone relations, they should successfully discriminate correct items from foils.

3.1. Materials and method

3.1.1. Participants

Twenty new native speakers of American English took part in this experiment (10 females, 10 males; mean age 20.1, range 18–29). They were recruited through the Harvard study pool and were randomly assigned to the two mode conditions.

3.1.2. Familiarization

The familiarization phase was identical to that in Experiments 1 and 2.

3.1.3. Test

3.1.3.1. Major condition. In the major condition, participants had to choose between the major test items from Experiment 1, and items comprising the same tones but in different orders; these items were thus permutations of the "correct" test items. The test items are shown in Table 4. The nine test pairs constructed from three "correct" items and three foils were presented twice, once with the correct items first, and once with the foil presented first. Test items were presented in random order with the constraint of having at most three pairs starting or ending with a correct item in a row.

Table 4
Test items used in Experiment 3. In the tone columns, letters and numbers stand for pitch classes and octave numbers, respectively (the middle C being C_4). Intervals starting with a capital and a lower-case M represent major and minor intervals, respectively. In the major condition, participants had to choose between the major test items from Experiment 1 and permutations of these test items. In the minor condition, participants had to choose between the minor test items from Experiment 1 and permutations of these test items.

Original		Permuted	
Tones	Intervals	Tones	Intervals
Major			
$C_4E_4G_4$	M3↑m3↑	$E_4C_4G_4$	M3↓M5↑
$F_4A_4G_4$	M3↑M2↓	$G_4A_4F_4$	M2↓M2↓
$D_5B_4G_4$	m3↓M3↓	$B_4D_5G_4$	m2↑M5↓
Minor			
$C_4E^{b}_4G_4$	m3↑M3↑	$E^{b}_4C_4G_4$	m3↓M5↑
$F_4A^{b}_4G_4$	m3↑m2↓	$A^{b}_4G_4F_4$	m2↓M2↓
$D_5B^{b}_4G_4$	M3↓m3↓	$B^{b}_4D_5G_4$	M3↑M5↓

3.1.3.2. Minor condition. The test items are shown in Table 4. Similarly to the major condition, participants had to choose between the minor test items from Experiment 1 and permutations of these test items. Test trials were constructed with the same constraints as in the major condition.

3.2. Results

As shown in Fig. 4, participants preferred legal items to foils ($M = 61.4\%$, $SD = 13.9\%$), t $(19) = 3.6$, $p < 0.002$, Cohen's $d = 0.81$, $p_{permut} = 0.002$ A mixed effects model with mode group as fixed effect predictor and a random intercept for participants revealed no difference between the mode conditions, $Z = 0.8$, $p = 0.403$, ns.

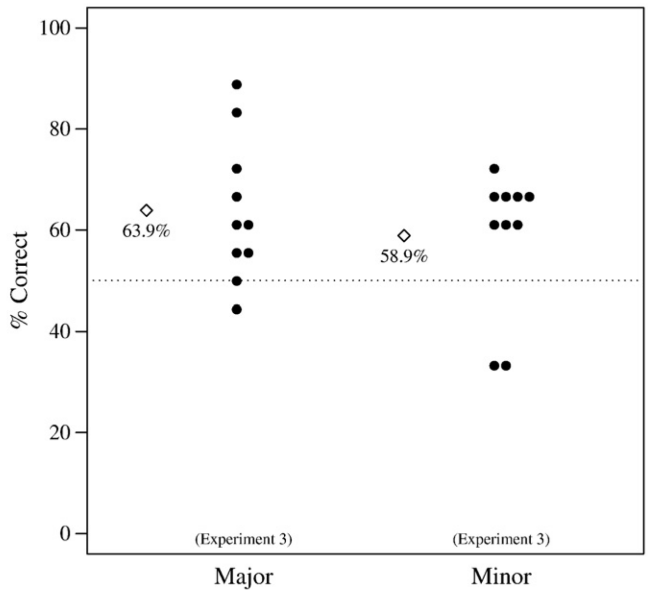

Fig. 4. Results of Experiment 3. Dots represent averages of individual participants, diamonds represent population averages, and the dotted line represents the chance level of 50%. Participants preferred the underlying tone triplets from the familiarization melodies over permutations of these triplets, even though both test items were in the same mode.

3.3. Discussion

In Experiment 3, participants successfully chose between the underlying tone triplets from the familiarization melodies, and permutations of these triplets. Importantly, both items in each test pair had the same mode; that is, major tone triplets were pitted against major foils, and minor triplets against minor foils. Hence, if, in Experiments 1 and 2, participants simply tracked the mode of the familiarization melodies, they should have failed in Experiment 3. As they successfully chose underlying tone triplets over foils, they must have tracked non-adjacent tone relations.

In contrast to Creel et al.'s (2004) experiments, participants in Experiments 1 to 3 successfully tracked non-adjacent tone relations. As mentioned in the introduction, a major difference between the experiments presented here and Creel et al.'s (2004) experiments is that the latter authors used random melodies, while the current experiments used tonal melodies. Experiment 4 asks whether this difference can account for the discrepancy between the current and Creel et al.'s (2004) experiments. In this experiment, participants were familiarized with the same underlying tone triplets as in Experiment 1, but with more random "ornaments". If tonal structure is crucial to tracking non-adjacent tone relations, participants might experience difficulties tracking such relations in Experiment 4.

4. Experiment 4

Experiments 1 to 3 suggest that people can track non-adjacent tone relations when presented with musically meaningful tonal stimuli. Given that Creel et al. (2004) did not find evidence for such an ability, it is important to find out whether listeners can track non-adjacent tone relations with arbitrary melodies, or just with tonal melodies such as those used during Experiments 1 to 3. Experiment 4 addresses this question. In this experiment, participants were familiarized with tone words made from the same underlying tone triplets as in Experiments 1 to 3; the ornaments, in contrast, were chosen to sound random, and were, therefore, no longer "passing tones".

4.1. Materials and method

4.1.1. Participants

Twenty new native speakers of Italian took part in this experiment (14 females, 6 males; mean age 21.9, range 19–30). They were recruited at SISSA, Trieste (Italy) and randomly assigned to the two mode conditions.

Table 5
Stimuli used in Experiment 4. In the tone columns, letters and numbers stand for pitch classes and octave numbers, respectively (the middle C being C_4). Intervals starting with capital and lower-case M represent major and minor intervals, respectively. Tone words were constructed from three underlying three-tone sequences. Participants were familiarized with "ornamented" versions of these tone triplets comprising six tones, and then tested on the underlying tone triplets. Half of the participants were familiarized with the major version, and the remaining participants with the minor version.

Ornamented		Underlying	
Tones	Intervals	Tones	Intervals
Major			
$C_4B^b_3E_4F^\#_4C^\#_4G_4$	M2↓m5↑M2↑M4↓m5↑	$C_4E_4G_4$	M3↑m3↑
$F_4F^\#_4A_4D_4C^\#_4G_4$	m2↑m3↑M5↓m2↓m5↑	$F_4A_4G_4$	M3↑M2↓
$D_5C_5B_4C_4C^\#_4G_4$	M2↓m2↓M7↓m2↑m5↑	$D_5B_4G_4$	m3↓M3↓
Minor			
$C_4B^b_3E^b_4F^\#_4C^\#_4G_4$	M2↓M4↑m3↑M4↓m5↑	$C_4E^b_4G_4$	m3↑M3↑
$F_4F^\#_4A^b_4D_4C^\#_4G_4$	m2↑M2↑m5↓m2↓m5↑	$F_4A^b_4G_4$	m3↑m2↓
$D_5C_5B^b_4C_4C^\#_4G_4$	M2↓M2↓m7↓m2↑m5↑	$D_5B^b_4G_4$	M3↓m3↓

4.1.2. Familiarization

As in the previous experiment, participants listened to a melody composed of a concatenation of ornamented tone words. The underlying three-tone sequences were the same as in Experiments 1 and 2. The new ornaments were rather arbitrary tones; the new ornamented versions are shown in Table 5.

Within tone words, there were nine transitions among adjacent tones with a TP of 1.0, and four with a TP of 0.5; second order TPs were either 1.0 (8 transitions) or 0.5 (4 transitions). Across tone words, first and second order TPs were 0.33, and thus lower than within tone words.

4.1.3. Test

The test phase was identical to that of Experiment 1.

4.2. Results

As shown in Fig. 5, participants preferred legal items to foils ($M=58.1\%$, $SD=15.5\%$), $t\ (19)=2.3$, $p=0.031$, Cohen's $d=0.52$, $p_{permut}=0.032$. However, in contrast to the other experiments, a mixed effects model with mode group as fixed effect predictor and a random intercept for participants revealed that participants performed significantly better in the minor condition compared to the major condition, $Z=2.3$, $p=0.021$. Analyzing the performance in the two groups separately, participants preferred legal items only in the minor group ($M=65.0\%$, $SD=13.9\%$), $t\ (9)=3.4$, $p=0.008$, Cohen's $d=1.1$, $p_{permut}=0.013$, but not in the major group ($M=51.1\%$, $SD=14.3\%$), $t\ (9)=0.25$, $p=0.808$, $p_{permut}=0.769$, ns.

4.3. Discussion

Experiment 4 asked whether participants could track non-adjacent dependencies also in less well-behaved melodies than in Experiments 1 to 3 by using somewhat arbitrary ornaments. Participants in the minor condition (but not in the major condition) succeeded in tracking these relations. These results suggest that the tracking of non-adjacent tone relations is possible, but perhaps less robust than that of adjacent relations; indeed, while adjacent relations are tracked in rather arbitrary

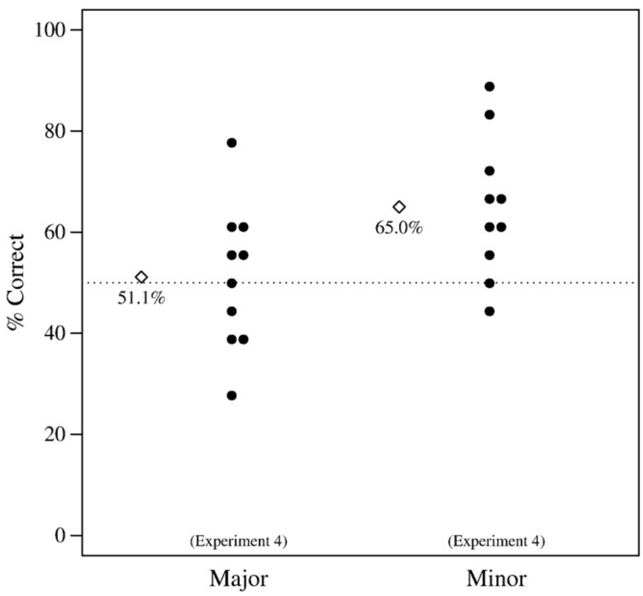

Fig. 5. Results of Experiment 4. Dots represent averages of individual participants, diamonds represent population averages, and the dotted line represents the chance level of 50%. When the ornaments were arbitrary tones, participants could track non-adjacent tone relations only for the minor mode group but not for the major mode group.

melodies (e.g., Saffran et al., 1999), participants succeeded only in the minor condition of Experiment 4.

Why did participants perform better in the minor mode group than for the major mode group? Possibly, they may have matched the global configuration of the melodies to the test items, and the minor ones may still have sounded sufficiently "minor-like." To quantify this impression, I analyzed the melodies with Krumhansl's (1990) key finding algorithm. According to this algorithm, the cumulative tone duration for each pitch class in a melody is computed (e.g., how long the pitch class C is sounded in total during the melody). Then, the algorithm evaluates how well these cumulative durations fit empirically determined tone profiles for a tonality (e.g., C major). (These tone profiles are empirically obtained using the probe-tone technique (Krumhansl & Kessler, 1982), which measures the relatedness of each pitch class to a tonality.) For each of the 24 major and minor keys, these profiles are correlated with the cumulative tone durations of the 12 pitch classes in the melody, and the best fit is chosen as the key of the melody.

The results of these analyses suggest that the key finding algorithm cannot account for the results of Experiment 4. In all experiments presented here, the familiarization melody in the major condition correlated most strongly with a major mode, and that in the minor condition with a minor mode, respectively. (The test items contained the same tones in all experiments, and correlated strongly with the appropriate modes, $p<0.001$). Importantly, the correlation coefficients for the winning modes were comparable in all experiments (Major: 0.54 (Experiments 1 to 3) and 0.537 (Experiment 4); minor: 0.443 (Experiments 1 to 3) and 0.454 (Experiment 4)). As a result, one would expect participants to perform similarly in Experiment 1 and in Experiment 4. It thus remains unclear why participants performed better in the minor condition of Experiment 4 than in the major condition.

That being said, the results of Experiment 4 allow for two important conclusions. First, at least in the major condition, participants failed to learn the non-adjacent melodies although, according to a prominent key finding algorithm, it should be as easy to find the mode of the familiarization streams in Experiment 1 and in Experiment 4, a result that is unexpected if participants simply tracked the mode of the familiarization melodies. Together with Experiment 3, these results suggest that participants did not simply track the mode of the familiarization melodies, but that they were sensitive to non-adjacent melodies, at least in Experiments 1 to 3.

Second, the results of Experiment 4 show that, under some conditions, participants experience difficulties tracking non-adjacent melodies. While they readily tracked them when musically well-behaved melodies were used, participants failed in the major condition of Experiment 4 where the familiarization melody was more "random", suggesting that the tonal structure of the familiarization melodies of Experiments 1 to 3 facilitated the tracking of non-adjacent melodies.

5. General discussion

Can listeners track probabilistic relations among non-adjacent elements? In the domain of language, results have been conflicting. While some researchers failed to find evidence for a sensitivity to relations among non-adjacent elements Newport & Aslin, 2004, others argued for the opposite conclusion (e.g., Ebbinghaus, 1885/1913; Endress & Bonatti, 2007; Endress & Mehler, 2009; Gómez, 2002; Peña et al., 2002). A plausible conclusion is that listeners can track, to some extent, relations among non-adjacent elements, but that, all things being equal, tracking relations among adjacent elements is easier. Moreover, listeners might have biases that make some non-adjacent relations easier to track than others, because some non-adjacent relations are more relevant than others for certain learning problems (Bonatti, Peña, Nespor, & Mehler, 2005; Mehler, Peña, Nespor, & Bonatti, 2006; but see Keidel, Jenison, Kluender, & Seidenberg, 2007).

Here, I start investigating these issues in the domain of music. In previous research, listeners seemed to successfully track non-adjacent tone relations only when the non-adjacent tones were perceptually adjacent due to streaming cues (Creel et al., 2004). (As mentioned in the introduction, the previous literature assumed that listeners can track transitional probabilities over tones, an assumption that is unlikely to be correct as it would require perfect pitch even for random melodies. Here, I make no such assumption. Listeners might have tracked transitional probabilities among tones, or, more likely, adjacent or non-adjacent *intervals* between tones.) However, these authors used random melodies, raising the question of whether participants would perform better if musically more meaningful material is used.

When presented with musically meaningful three-tone sequences that were interspersed with three other tones (used as "ornaments"), participants successfully tracked the non-adjacent relations among the tones in the underlying tone triplets. That is, after familiarization with such ornamented tone triplets, participants chose major versions of the underlying tone triplets over minor versions if they had been familiarized with a major melody, and minor versions over major versions when they had been familiarized with a minor melody. Further experiments showed that participants did not simply track specific tones or intervals in the familiarization melodies, or just their mode (i.e., major or minor); instead, participants seem to have noticed the non-adjacent tone relations among the tones in the underlying tone triplets. Finally, when the underlying tone triplets were ornamented with out of key tones, participants found the non-adjacent tone relations more difficult to track.

Together, these results allow for two conclusions. First, non-adjacent tone relations seem to be more difficult to compute than relations among adjacent tones; while adjacent tone relations are computed for rather arbitrary melodies (e.g., Saffran et al., 1999), the experiments presented here and those by Creel et al. (2004) show that non-adjacent tone relations are computed preferentially with tonal stimuli. Second, participants seem to have a genuine sensitivity to non-adjacent tone relations, even though this sensitivity can be observed only under appropriate testing conditions; that is, this sensitivity seems to be much stronger when participants are exposed to tonal, musically meaningful tone sequences as opposed to random tone sequences as in Experiment 4 and Creel et al.'s (2004) experiments. As in the speech case (Bonatti et al., 2005), listeners thus seem to bring specific biases to the learning task that make non-adjacent tone relations easier to learn in tonal as opposed to random melodies.

Acknowledgments

I am grateful to Ágnes Kovács, Jacques Mehler, Mohinish Shukla, and Juan M. Toro for their helpful comments and discussions. The study is supported in part by the HFSP grant RGP 68/2002, the Regione Friuli-Venezia-Giulia (L.R. 3/98) and the European Commission Special Targeted Project CALACEI (contract no. 12778 NEST) to Jacques Mehler, and by the Mind, Brain, and Behavior Interfaculty Initiative at Harvard University.

References

Aslin, R. N., Saffran, J., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321−324.

Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390−412.

Bates, D. M., & Maechler, M. (2008). *lme4: Linear mixed-effects models using S4 classes, R package version 0.999375-9*. http://www.cran.r-project.org/web/packages/lme4/index.html

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science*, *16*(8).

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: The MIT Press.

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(5), 1119−1130.

Croonen, W. L. (1994). Effects of length, tonal structure, and contour in the recognition of tone series. *Perception and Psychophysics*, *55*(6), 623−632.

Dowling, W. J. (1973). The perception of interleaved melodies. *Cognitive Psychology*, *5*(3), 322−337.

Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception and Psychophysics*, *50*(4), 305−313.

Dowling, W. J., Kwak, S., & Andrews, M. W. (1995). The time course of recognition of novel melodies. *Perception and Psychophysics*, *57*(2), 136−149.

Dowling, W. J., Lung, K. M., & Herrbold, S. (1987). Aiming attention in pitch and time in the perception of interleaved melodies. *Perception and Psychophysics*, *41*(6), 642−656.

Ebbinghaus, H. (1885/1913). *Memory: A contribution to experimental psychology.*New York: Teachers College, Columbia University (http://www.psychclassics.yorku.ca/Ebbinghaus/).

Endress, A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, *105*(2), 247−299.

Endress, A. D., & Mehler, J. (2009). Primitive computations in speech processing. *The Quarterly Journal of Experimental Psychology*, *62*(11), 2187−2209.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*(6), 499−504.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(24), 15822−15826.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431−436.

Halpern, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory and Cognition*, *17*(5), 572−581.

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, *78*(3), B53−B64.

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, *100*(1), 33−72.

Jaeger, T. F. (2008). Categorical data analysis: Away from anovas (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434−446.

Keidel, J. L., Jenison, R. L., Kluender, K. R., & Seidenberg, M. S. (2007). Does grammar constrain statistical learning? Commentary on Bonatti, Peña, Nespor, and Mehler (2005). *Psychological Science*, *18*(10), 922−923.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* New York: Oxford University Press.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334−368.

Lashley, K. (1951). The problem of serial order in behavior. In L. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112−136). New York: Wiley.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Linde, H. -M. (1958). *Kleine Anleitung zum Verzieren alter Musik.* Mainz: Schott.

Mehler, J., Peña, M., Nespor, M., & Bonatti, L. (2006). The "soul" of language does not use statistics: Reflections on vowels and consonants. *Cortex*, *42*(6), 846−854.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127−162.

Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in speech processing. *Journal of Memory and Language*, *53*(2), 225−237.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298*(5593), 604−607.

Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-PLUS.* Berlin, Germany: Springer.

Saffran, J. R. (2003). Absolute pitch in infancy and adulthood: The role of tonal structure. *Developmental Science*, *6*(1), 35−43.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926−1928.

Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology*, *37*(1), 74−85.

Saffran, J. R., Johnson, E., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27−52.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606−621.

Saffran, J. R., Reeck, K., Niebuhr, A., & Wilson, D. (2005). Changing the tune: The structure of the input affects infants' use of absolute and relative pitch. *Developmental Science*, *8*(1), 1−7.

Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, *69*(2), 105−134.

Schellenberg, E. G., & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science*, *14*(3), 262−266.

Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception and Psychophysics*, *67*(5), 867−875.

Youngblood, J. (1958). Style as information. *Journal of Music Theory*, *2*, 24−35 (as cited by Krumhansl, 1990).