Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009) Jyväskylä, Finland
Jukka Louhivuori, Tuomas Eerola, Suvi Saarikallio, Tommi Himberg, Päivi-Sisko Eerola (Editors)

Adaptive memory representations of musical tempo and pitch

Sabine Strauß & Oliver Vitouch

Dept. of Psychology, University of Klagenfurt, Austria
sabine.strauss@uni-klu.ac.at, oliver.vitouch@uni-klu.ac.at

ABSTRACT

Recent studies showed that the perception of extremely shifted versions of familiar pieces of music affects judgements about the original tempo and pitch. This paper aims to give an overview on musical adaptation effects (MAEs) for musical tempo and pitch. The focus is set on the questions if the recently found MAEs are generally comparable between the tempo and the pitch domain, and if they can be found to the same extent. In a series of eight empirical studies (N = 288, mean age = 26, SD = 8.6, 78 % female participants) the influence of extremely shifted versions (tempo-accelerated or pitch-shifted) of audio signals on memory representations was tested. Within this paper the focus will be set on a study on musical pitch (N = 30, mean age = 23, SD = 3, 77 % female participants). We used six TV themes, pre-selected for familiarity, as stimuli. The influence of different presentation conditions on the persistency and comparability of MAEs in the time and pitch domain were tested. The results of repeated-measures ANOVAs revealed treatment x probe interactions. Overall, participants had difficulties to distinguish between a shifted and an original probe version after hearing an extreme version in the treatment. MAEs vary between the domains of interest. The adaptation effects for musical tempo are stronger and more systematic than the effects found for musical pitch. The results give further insight into the memory representations of musical pitch in comparison with the results from the time domain.

I. INTRODUCTION

In general musical adaptation effects (MAEs) are a facet of the musical long-term memory. Flexible long-term memory representations play a crucial role for helping humans to orientate in a constantly changing environment in general, and especially in the musical environment. Long-term memory processes provide standards for the comparison of novel information, but also allow for adapting these standards (e.g., Bower, Thompson, Schill, & Tulving, 1994).

Human faces are typically assumed to elicit stable long-term memory representations. Carbon and Leder (2005, 2006) have shown that the presentation of extremely distorted familiar faces (faces of celebrities) affects the ability to distinguish between an original and a shifted version of the face. This adaptation effects were not only short-termed but last for minutes and even for 24 hours (Carbon, Strobach, Lanton, Harsànyi, Leder, & Kovàcs, 2007).

Studies have shown that memory representations for musical tempo and pitch can be remarkably stable across time, even for very long time spans (e.g., Levitin, 1996; Levitin & Cook, 1996; Schellenberg & Trehub, 2003). This project deals with memory representations for musical tempo and musical pitch. According to the results of the initial tempo study, the focus of the study reported here was set on the musical domain pitch with a comparable experimental design as in the initial tempo study (Strauß et al., 2006). The stimuli used in this study were digitally altered on the pitch level.

In the context of musical long-term memory processes theoretical input on absolute pitch has to be taken into account. Absolute pitch can be described as a rare ability of auditory perception and memory. Absolute pitch is characterized by the ability to identify or produce the pitch level of a musical piece without contextual cues or reference pitches. Implicit memory for pitch relations could have an important impact on MAEs in the domain for pitch. Schellenberg and Trehub (2003) showed that absolute pitch memory is widespread among adults with no musical training. They used instrumental soundtracks of familiar television programs as stimuli. Participants heard five seconds excerpts either in the original or shifted (upward or downward) by one or two semitones. Participants could successfully identify the original pitch levels. The reported results confirm the accuracy of implicit memory for pitch relations. Schellenberg and Trehub (2003) also refer to research on knowledge which is acquired without knowledge (see, e.g., Goshen-Gottstein, Moscovitch, & Melo, 2000; Reber & Allen, 2000; Tillman, Bharucha, & Bigand, 2000), which could be applied to pitch levels of familiar pieces of music.

Predominantly we were not interested in the number of absolute pitch possessors within this study but in the question if the adaptation effects can also be found for the domain of pitching. Within this project, MAEs for pitch were investigated and compared with the recently found effects in the time domain.

II. AIMS

The aim of the study reported here was to investigate adaptation effects in the domain of pitching in comparison to the recently reported results from the tempo domain. The overall research question was: Does the perception of extremely distorted versions of familiar pieces of music (TV themes) affect judgements about the original pitch level of these pieces?

The results of this study should shed light on the question, if adaptation effects for the two domains tempo and pitch are comparable, and therefore domain-general in the auditory realm. Additionally, the influence of musical training on adaptation processes was tested.

III. METHOD

A. Subjects

Overall, a number of 288 people (mean age = 26, *SD* = 8.6, 78 % females) participated in a series of eight empirical studies on MAEs. Within this paper the focus will be set on a study in which 30 people participated (mean age = 23 years, *Min* = 19, *Max* = 29, *SD* = 3.0, 77 % females). All participants were students of Klagenfurt University. They were recruited as part of study requirements and were naive to the aim of the study.

B. Stimuli

We used the themes of six US television series from the 1960s, 1970s, 1980s, and 1990s, pre-selected for familiarity, as stimuli.

- The Waltons (USA 1972–1981)
- I Dream of Jeannie (USA 1965–1970)
- Knight Rider (USA 1982–1986)
- Dallas (USA 1978–1991)
- The X-Files (USA 1993–2002)
- Roseanne (USA 1988–1997)

As we could already show in past studies (Strauß & Vitouch, 2005, Strauß et al., 2006) TV themes are appropriate stimuli for the investigation of musical long-term memory. They are in general very well-known and are rapidly recognized by certain age cohorts, who absorbed them at a high rate (Strauß & Vitouch, 2005).

The stimuli for the pitch condition were altered, using the online freeware audacity (www.audacity.com), in only one direction – shifted higher on the pitch level. The pitch shifts had no influence on the tempo of the musical piece. The digital acceleration also had no impact on the quality of the musical piece. This was tested systematically in a pre-study.

Each TV theme was used in three different versions:

1. the original version,
2. an extreme version (+ 7 semitones), and
3. a shifted version (+ 2 semitones).

C. Procedure

The test was conducted fully computer-assisted in the X-Lab of Klagenfurt University, using the experimental software E-prime. A maximum of four participants could listen to the test stimuli via headphones in a preset volume level at one time-point. Participants were individually instructed via monitor and asked to enter their responses via keyboard. The stimuli were presented in a random order. This was implemented to reduce interactions between the participants during the parallel test session. In the first part of the procedure participants heard a TV-theme-set (treatment-version of the TV theme for 15s, silence 2s, pink noise 3s, silence 2s, probe-version for a max. of 30s) and should judge if the last heard version (probe-theme) is an original or not by pushing the appropriate key (see figure 1).

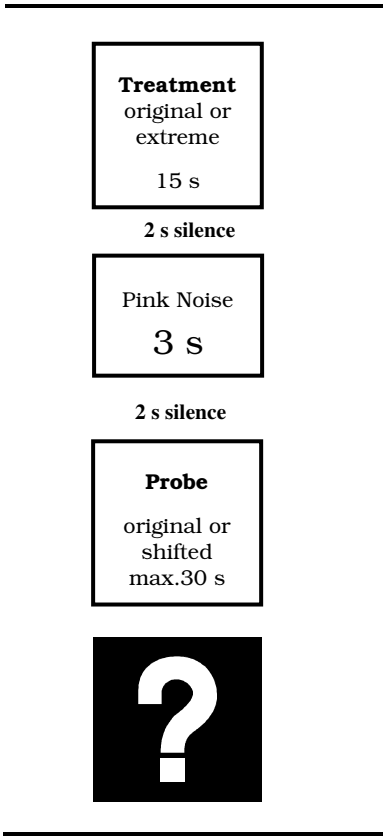

Figure 1: Scheme for one experimental trial. The trial starts at the top (treatment) and closes with the question if the version heard last was an original or not.

The treatment and probe versions were masked. Participants did not get any information about the TV theme versions. They did also not get any feedback if the theme is an original version or not. Neither did participants get feedback about the correctness of their decisions. They were asked to make their decisions on the validity of a probe theme based on their knowledge about the TV theme and not on the previous heard treatment versions. After the test procedure the recognition task and the questionnaire were presented. In the recognition task participants were asked to name the six TV themes. If participants were not able to remember the title of the TV theme, the TV theme was also counted correct if they could instead describe the TV series. The questionnaire included questions on demographic data, favorite TV series, frequency of TV series consumption, when participants had seen the series from the study last, and questions on the level of musical education.

D. Design

The design of the study was fully balanced. Every possible treatment (original vs. extreme) of the six TV themes was crossed with every possible probe version (original vs. shifted). This 2 x 6 x 2 design results in a total of 24 trials. The trials were randomly sampled by E-prime. The procedure including the questionnaire and recognition task took about 20 minutes for each participant.

IV. RESULTS

A. Recognition task

The recognition rates were to some extent lower than in the tempo study (Strauß et al., 2006). The best recognized theme was Knight Rider with 92 % of the participants recognizing it. The recognition rates of each theme are displayed in table 1. The recognition values of the tempo study are displayed in brackets behind the current data. Within this study only the TV themes from the late 1980s and 1990s (Knight Rider, The X-Files, and Roseanne) were well recognized. The "older" TV themes from the 1970s and early 1980s had lower recognition rates as expected. The theme of the TV series Emergency Room had the lowest recognition frequencies in the tempo study. For this reason it was replaced by the theme of Roseanne.

The recognition rates were still satisfying, but were lower than in the tempo study. The reasons for this could lie in the age distribution of participants (mean age_{tempo}: 26.3 years, mean age_{pitch}: 23 years). The mean age of participants in the tempo study was 26 years, and the mean age of participants in the pitch study was 23 years. In a direct comparison the age distribution of participants was quite similar to the age distribution of the tempo study, but a mean age difference of three years could be accountable for the lower recognition values. It could have been that some of the youngest participants had never seen the series from the 1960s, 1970s and early 1980s (I dream of Jeannie, Dallas, and The Waltons), even though that there are many reruns on the current TV programme.

Another reason could be the E-prime test format. Participants could skip (accidentally) the question by pressing any key. The input format was changed for future studies.

Table 1: Recognition rates for the TV themes in %

TV themes	freq (%)	freq (%)$^{+}$
Knight Rider	92	90
X-Files	77	87
Roseanne	53	*
Jeannie	38	63
Dallas	23	57
Waltons	15	67

Note: $^{+}$ = Comparative values of the tempo study are displayed in column 2;* = not used in the tempo study.

B. Adaptation processes

A three-way repeated-measures ANOVA showed no treatment x probe interaction (p = .715). The factor probe was not significant (p = .946). The factor treatment on the other hand *was significant (p < .001).* This means that the average yes rate was lower in trials with extreme versions, independent from the probe versions. The details on the three-way repeated-measures ANOVA are displayed in table 2.

Table 2: Overview of the results of three-way repeated-measures ANOVA

Factor	Pitch
Probe	F (1, 25) = 0.005; p = .946
Treatment	F (1, 25) = 20.271; p < .001
TV theme	F (5, 125) = 3.127; p = .011
Probe x Treatment	F (1, 25) = 0.137; p = .715
Probe x Treatment x TV theme	F (1, 25) = 0.756; p = .583

Note: Only those interactions which were relevant for testing the hypothesis are displayed.

Participants were not better in distinguishing between an original and a shifted probe after hearing an original treatment (without explicitly knowing that it was an original version). Participants seem to be irritated by the treatment to some extent because the average rejection of a shifted probe fell from about 67 % in the control condition down to 48 % in the extreme treatment condition. The main effect treatment is displayed in figure 2. There is an adaptation effect for musical pitch but the effect is not straightforward comparable with the adaptation effects for musical tempo. Participants' decisions are just around the chance level.

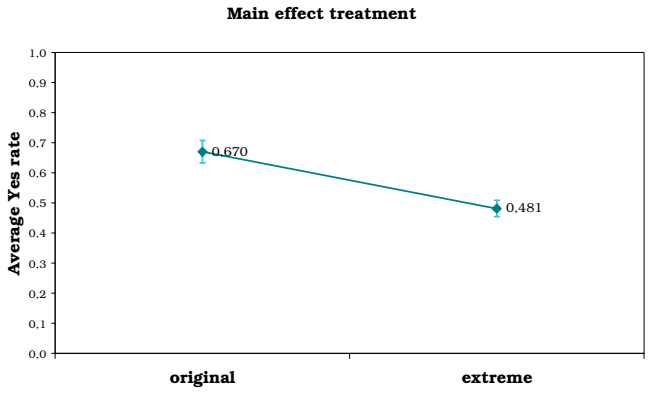

Figure 2: Main effect treatment. Note: Error bars are SEs.

The means of the main interaction treatment x probe are displayed in table 3. The means for the yes rates of participants' decisions show that participants who heard an extreme version in the treatment had difficulties in identifying an original probe version.

Table 3: Mean yes-rates for the treatment x probe interaction

	probe original		probe shifted	
	treatment		treatment	
	original	extreme	original	extreme
mean	.660	.494	.679	.468
SE	.048	.057	.038	.055

An adaptation effect for musical pitch was found but the effect is not directly comparable with the effect found for musical tempo. The reasons therefore have to be considered.

In the realm of absolute pitch it is assumed that absolute pitch possessors often make a mistake in pitch labelling within one semitone (see, e.g., Lockhard & Byrd, 1981; Miyazaki, 1988). Schellenberg and Trehub (2003) argue that ordinary adults' memory (adults without any musical training) for music is similar to the memory of absolute pitch possessors. In line with this argumentation a question has to be considered. Are shifted versions (higher about two semitones) falsely identified as original versions because of an adaptation effect, or because the implicit memory for the pitch level which is distorted? Schellenberg and Trehub (2003) found that participants could identify the original pitch levels out of one or two semitones shifted upward or downward excerpts of familiar pieces of music. These findings and the theoretical assumption that memory for absolute pitch is widespread would strongly support that the results of the present study can be explained by an adaptation effect. The findings of Dowling (1994) about precise memory representations of music as well as relational components would also support adaptation and not a pitch bias.

To summarize, the results of the three-way repeated-measures ANOVA revealed no significant treatment x probe interaction, although the factor treatment was significant. The impact of the factor treatment could be interpreted as adaptation effect. There are reports in the literature on preferred pitch levels (see, e.g., Helmholtz, 1863) and preferred intervals (octaves, fifths, fourths and thirds). The pitch levels of the six TV themes are all common. They are all prototypically for the genre of TV themes. The results cannot simply be explained by the level of pitch. The treatment appears to have a strong influence on the decision if a probe theme is an original or not. The results of the tempo study show that participants seem to have a kind of intuitive knowledge about the correct tempo. The correct pitch level appears to be more difficult to identify, and it could be that it requires a higher level of musical expertise. This is assumed for the general memory for musical pitch and perception of pitch relations which are a function of musical expertise (see, e.g., Krumhansl, 1999), and increase with the level of musical training.

We assumed that absolute pitch possessors would be to some extent some kind of resistant to adaptation processes. Absolute pitch possessors would be able to name the exact pitch level, and would give more attention to the pitch information than others would. We assume that they would not update the information on the distorted pitch level because their attention would be drawn to the absolute pitch level of the stimuli. The assumptions will be tested empirically in a future study with an extreme-group of absolute pitch possessors.

C. Domain comparison: Pitch and tempo

The adaptation effect for pitch was found, but was to some extent weaker than the adaptation effect for musical tempo. Only the determining factor treatment was significant in the pitch study. In table 4 the results of the repeated-measures ANOVAs for both domains are contrasted.

Table 4: Comparison pitch and tempo

Factor	Tempo	Pitch
Probe	F (1, 29) = 11.752, p = .002	F (1, 25) = 0.005, p = .946
Treatment	F (1, 29) = 21.847, p < .001	F (1, 25) = 20.271, p < .001
TV theme	F (5 ,125) = 4.425, p = .001	F (5, 125) = 3.127, p = .011
Treatment x TV theme	F (1, 29) = 21.181, p < .001	F (1, 25) = 0.756, p = .583

D. Results from the questionnaire data

1) Impact of the level of musical education

In the questionnaire participants were asked for their level of musical education. They should answer one question on the level of musical education, and one question on the instrument they are playing. The frequencies of participants with a musical education are displayed in figure 3. The group of non-musicians (58 %) was larger than the group of musicians. The types of instruments participants reported to play are displayed in figure 4. The largest percentage of participants (46 %) stated to play no instrument. People who had some musical education were merged to a subgroup called musicians. Additionally the impact of musicians defined as between subject factor in a mixed repeated-measure ANOVA was estimated. The factor musicians was not significant (F (1, 24) < 1, p = .992). The variable musician was also tested as covariate to the three-way repeated-measures ANOVA. The influence of the covariate was also not significant (F (1, 24) < 1, p = .992).

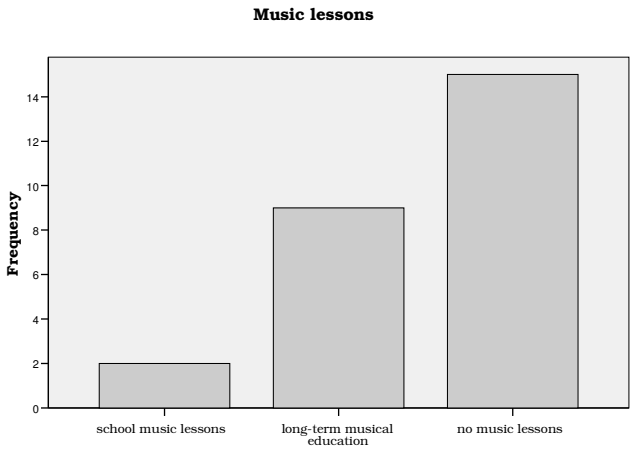

Figure 3: Music lessons. Note: The bar school music lessons contains all answers of participants who had special music lessons at school for example in special school types. School music lessons do not include the regular and compulsory music lessons at school.

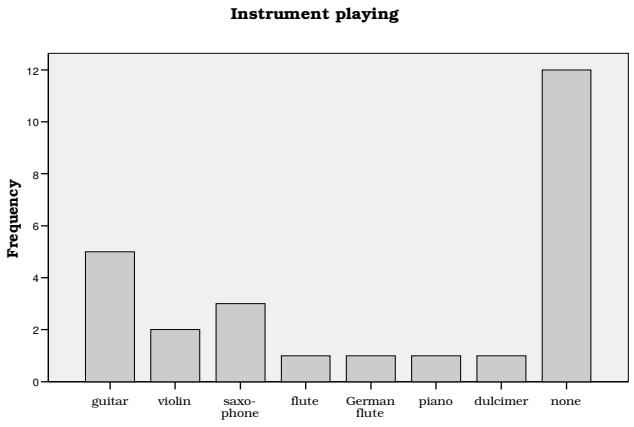

Figure 4: Frequency of instruments participants are playing.

2) Frequency of last seen values for the TV themes

Table 5 shows the frequencies of time points when participants reported to have seen the TV series last. As expected the largest percentage of participants saw the TV series in their childhood or youth. The TV series Knight Rider, The X-Files, and Roseanne were the only series which were seen, from some participants within the last year. The participants were asked in an open question format for the remembered time point when they had seen the TV series last. The open question format could also be partially explaining the high missing values.

Table 5: Frequency of the last-seen values for the TV themes in %

TV theme	last seen				
	1	2	3	4	5
Jeannie	57.7	7.7	3.8	x	14.3 (19.2)
Waltons	23.1	7.7	3.8	x	50.0 (15.4)
Dallas	38.5	3.8	15.4	x	30.8 (11.5)
Knight Rider	38.5	15.4	15.4	15.4	11.5 (3.8)
X-Files	11.5	30.8	42.5	7.7	3.8 (3.8)
Roseanne	30.8	11.5	19.5	15.4	7.6 (15.4)

Note: 1 = childhood, 2 = youth, 3 = last 5 years, 4 = within the last year, 5 = never (missing). Missing values are displayed in brackets.

V. CONCLUSIONS

MAEs were found for both domains of interest: musical tempo and musical pitch. The adaptation effect for musical tempo is quite strong and distinct. Adaptation effects for pitch could be found too, but were to some extent weaker than the effect for musical tempo. Only the factor treatment (original vs. extreme) was significant in the pitch study. This factor seems to be the determining factor for the adaptation effect in both domains.

As in the tempo study, the presentation of extremely pitch-altered versions of TV themes reduces the ability to distinguish between original and shifted versions. Participants seem to rapidly integrate new information into a flexible mental representation frame. The results show that the factor treatment is the determining factor for the adaptation effect.

The results are also in line with Dudai's (2004) neurophysiological theory of memory formation. He states that new information gets quickly integrated into a flexible memory representational frame. Like in the domain of face processing it seems to be evolutionary plausible and important to integrate new pitch information into memory representations.

Overall, the pitch condition seems to be in general more difficult, and not as self-explanatory as the tempo condition. The results of the tempo study show that people seem to have a kind of intuitive knowledge about the correct tempo. The correct pitch level appears to be much more difficult to identify. The identification of a pitch level seems to require a higher level of musical expertise.

The level of musical education had no influence on MAEs in the present study. People with higher levels of musical education did not perform better on the task.

Based on the results of the tempo study we are interested how long MAEs last. In the visual domain, adaptation effects were found after durations of five minutes up to 24 hours between the treatment and the test phase (Carbon & Leder, 2005; Carbon et al., 2007). Within the project of eight empirical studies on MAEs the aspect of duration will be considered.

REFERENCES

Bower, G. H., Thompson-Schill, S., & Tulving, E. (1994). Reducing retroactive interference: An interference analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20,* 51-66.

Carbon, C.-C., & Leder, H. (2005). Face adaptation: Changing stable representations of familiar faces within minutes? *Advances in Cognitive Psychology, 1,* 1-7.

Carbon, C.-C., & Leder, H. (2006). The Mona Lisa effect: Is "our" Lisa fame or fake? *Perception, 35,* 411-414.

Carbon, C.-C., Strobach, T., Langton, S. R. H., Harsànyi, G., Leder, H., & Kovàcs, G. (2007). Adaptation effects of highly familiar faces: Immediate and long lasting. *Memory & Cognition, 35,* 1966-1976.

Dowling, W. J. (1994). Melodic contour in hearing and remembering melodies. In R. Aiello & J. Sloboda (Eds.), *Musical perceptions* (pp. 173-191). Oxford: Oxford University Press.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual Review of Psychology, 55,* 51-86.

Goshen-Gottstein, Y., Moscovitch, M., & Belo, B. (2000). Intact implicit memory for newly formed verbal associations in amnesic patients following single study trials. *Neuropsychology, 14,* 570-578.

Helmholtz, H. v. (1866). *Handbuch der physiologischen Optik.* Hamburg: Voss.

Krumhansl, C. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126,* 159-179.

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, *56*, 414-423.

Levitin, D. J., & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics, 58,* 927-935.

Lockhead, G. R., & Byrd, R. (1981). Practically perfect pitch. *Journal of the Acoustical Society of America, 70,* 387-389.

Miyazaki, K. (1988). Musical pitch identification by absolute pitch possessors. *Perception & Psychophysics, 44*, 501-512.

Reber, A. S., & Allen, R. (2000). Individual differences in implicit learning: Implication for the evolution of consciousness. In R. G. Kunzendorf & B. Wallace (Eds.), *Advances in consciousness research: Vol. 20. Individual differences in conscious experience* (pp. 227-247). Amsterdam: John Benjamins.

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

Strauß, S., & Vitouch, O. (2005). Jeannie, Kirk, & John Boy: Title thems of Tv series as triggers of old and biographical memories. In I. Manenica (Ed*.), 7th Alps-Adria Conference in Psychology – Abstracts* (pp.221-222). Zadar, Croatia: Dept. of Psychology, University of Zadar.

Strauß, S., Vitouch, O., Ladinig O., Augustin, D., Carbon, C.-C., & Leder, H. (2006). Memory representations of musical tempo: Stable or adaptive? In M. Baroni, A. R. Addessi, R. Caterina & M. Costa (Eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition & 6th Triennial Conference of the European Society for the Cognitive Sciences of Music* (pp. 90-94). Bologna: Bononia University Press.

Tillman, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review, 107,* 885-913.