Musicæ Scientiæ
Fall 2009, Vol XIII, n° 2, 231-272

Expectancy in Sami Yoiks revisited: The role of data-driven and schema-driven knowledge in the formation of melodic expectations

TUOMAS EEROLA*, JUKKA LOUHIVUORI*
AND EDWARD LEBAKA**
* Department of Music, University of Jyväskylä, Finland
** Department of Music, Pretoria University, South Africa

• **Abstract**

This study extends a previous study concerning melodic expectations in North Sami yoiks (Krumhansl *et al*., 2000) in which a comparison between expert and non-expert listeners demonstrated the existence of a core set of principles governing melodic expectancies. The previous findings are reconsidered using non-Western listeners (traditional healers from South Africa) in a modeling investigation. Comparison of different models made it possible to separate the role of data-driven and schema-driven models in melodic expectancies and to reveal any possible Western bias in previous studies. The results of the experiment, in which African listeners rated the fitness of probe-tones as continuations of North Sami yoik excerpts, indicated that data-driven models are adequate in explaining the expectancies, regardless of the cultural background of the listeners. The frequency-based models exerted more influence on listeners unfamiliar with the yoik style, the Western schematic model had most impact on Western listeners, and the style-specific models carried most predictive power for those listeners who possessed knowledge about yoiks.

Keywords: music cognition, melodic expectancy, auditory model, cross-cultural, statistical learning, expectation.

Introduction

Expectancy has a central role in all forms of human behaviour. It enters into perception, planned motor movements, speech production and comprehension, all of which continuously shape our reactions to events in the real world. In music, too, expectancy has long stimulated the minds of music theorists, psychologists and musicians, because of the tangible connection between expectancies and tension, surprise and other musical devices in music. Recent work on melodic expectancies (Krumhansl, Louhivuori, Toiviainen, Järvinen & Eerola, 1999; Krumhansl, Toivanen, Eerola, Toiviainen, Järvinen & Louhivuori, 2000; Krumhansl, 2000; Roh & Yi, 2000) has shown how music draws on common psychological principles of expectation, even if musical cultures have a distinct effect on the application of these principles. These principles have been captured in a cognitively oriented musico-theoretical model of melodic expectations (Narmour, 1990). This model, which is data-driven, has been intensively investigated by music psychologists (Aarden, 2002; Larson, 2006; Margulis, 2005; Krumhansl, 1995a, b; Schellenberg, 1997; Schellenberg, Adachi, Purdy & McKinnon, 2002; Thompson & Stainton, 1998; von Hippel, 2000). Another data-driven explanation of melodic expectancies relies on the listener's sensitivity to statistical properties in the melodies, a factor which has also been shown to influence the formation of the melodic expectations (Krumhansl *et al.*, 1999; 2000; Krumhansl, 2000; Oram & Cuddy, 1995). Both of these models are presumed to be universal, and the claim of universality has been upheld in cross-cultural contexts using various styles of music.

The studies mentioned above subscribe to the notion of *universalism* in cross-cultural psychology, which assumes that there are world-wide similarities, or universalities in underlying psychological processes, and especially in human cognitive processes (Berry, Poortinga, Segall, & Dasen, 1992). Substantial similarities in responses to music-related tasks have been observed among listeners from different cultures, in studies where listeners have evaluated simple auditory stimuli (Poortinga, 1972; Bragg & Crozier, 1974), and similarities have also been observed in more specific investigations of the temporal organization (Toiviainen & Eerola, 2003) and tonal organization (Castellano, Bharucha, & Krumhansl, 1984; Kessler, Hansen & Shepard, 1984) of music. However, the agreement among people from various cultures starts to diminish once the focus of the studies turns towards the culturally learned properties of music. Studies on emotional responses to music have yielded similarities between listeners from separate cultures, yet differences too have emerged (Balkwill & Thompson, 1999; Gregory & Varney, 1996; Meyer, Palmer & Mazo, 1998). Overall, there are reasons to believe that melodic expectancy may be positioned somewhere between universal and culturally determined processes, since the relevant studies, reviewed below, do appear to demonstrate both similarities and differences between listeners from different cultures.

In experiments on melodic expectancies, listeners from separate musical cultures rate the fitness of different possible continuations of melodies. The first such cross-cultural study involved a comparison of sung responses to intervals, using as respondents German, American and Hungarian musicians (Carlsen, 1981; Unyk & Carlsen, 1987). Later, a probe-tone methodology was used to compare the fitness ratings given by Indian and Western listeners to North Indian ragas (Castellano, Bharucha, & Krumhansl, 1984) and fitness ratings given by Western and Balinese listeners to

Balinese music (Kessler, Hansen & Shepard, 1984). In a similar vein, native Chinese and American listeners' responses to Chinese and British folk songs have been studied (Krumhansl, 1995b). More recently, probe tones after excerpts from Finnish spiritual folk hymns (Krumhansl *et al.*, 1999) and North Sami yoiks (Krumhansl *et al.*, 2000) were rated by listeners who were active in the musical culture in question, and by Western musicians unfamiliar with the style. The results of these studies have generally supported the claim of universality for certain bottom-up principles in Narmour's model. In addition, the importance of statistical information in the formation of expectancies has been demonstrated. In spite of these results, the relative contribution of the principles in Narmour's model has varied across styles, regardless of the musical background of the listeners. Thus, the findings support the notion that expectations are governed by common psychological principles and stylistic knowledge has a distinct, additive effect on these principles. It should also be noted that all the groups of listeners in previous cross-cultural studies have been significantly exposed to Western music. Indeed, one could say that the listeners have been living in areas permeated with Western musical culture, and hence the research into the universality of Narmour's principles has not been very firmly based.

The present study aims to distinguish the role of innate expectancies, Western schematic knowledge and acoustic information in the formation of expectancies. The study takes as its point of departure a previously reported yoik study (Krumhansl *et al.*, 2000). It makes use of the data from that study, proceeding now to compare the responses of Western, Finnish and Sami listeners with those of a group of listeners who are not strongly influenced by the Western musical tradition, namely Pedi traditional healers from South Africa. This will allow us to reconsider the universality of Narmour's principles. It will also allow us to ascertain how far Western schematic knowledge — which has previously been documented as a predictor — is truly related to typical Western melodic expectancies.

The original study by Krumhansl *et al.* (2000) attempted to differentiate between data-driven and schema-driven models. Yet some uncertainties may still remain in terms of the roles of the models in the formation of expectancies: the point is that the responses of both non-experts and experts could have reflected generic Western melodic expectancies instead of truly data-driven, bottom-up, non-stylistically influenced expectations. Another aspect of our investigation derives from the fact that listeners may well adapt to the regularities of any music they hear, and will form their expectations based on frequency information provided by the musical context. This issue will be addressed using statistical models in the prediction of the expectancies. In addition, our analysis will be supplemented with a short-term memory model based on auditory modelling. This will extend the data-driven explanations to cover automatic processing of auditory signals by the auditory system, thus providing an additional explanatory level that goes beyond pitch-based models.

Data-Driven and Schema-Driven Expectancy Models

The issues pertaining to melodic expectancies can be divided into data-driven and schema-driven processes. In data-driven processes, expectancies arise from the musical material itself, without the application of specific knowledge or higher level cognitive processing[1]. A great deal of pre-processing is assumed to occur in pitch-based models that do not normally consider the formation of pitch sensation from the perspective of auditory perception. In the present study, three data-driven processes come under consideration: (a) the auditory memory model (Leman, 2000), (b) implication-realization principles of melodic expectancies (Narmour, 1990), and (c) statistical information within the melodies themselves. The schema-driven processes are addressed by considering (a) the ratings of two-tone contexts by Western listeners (Krumhansl, 1995a) and (b) various yoik-specific models, with a variety of predictors, including predictors derived from statistical style analyses of yoiks and predictors based on tones assumed to be correct in yoiks (Krumhansl *et al.*, 2000). Before going into more detailed hypotheses concerning the various explanatory models, the sections below will describe the original study and offer an initial presentation of a new model.

A Cross-Cultural Study of Melodic Expectations in North Sami Yoiks

The original study was a collaboration between members of a large research group consisting of music psychologists, ethnomusicologists and modellers of music cognition[2]. The aim was to compare three different approaches — behavioural experiments, statistical analyses and self-organizing neural network simulations — in studying the role of learning in melodic expectancies. In this study, excerpts were chosen from a corpus of North Sami yoiks. This musical style originates from Lapland, Northern Scandinavia, and has several distinct musical features that make it suitable for a study of expectancies. Yoik is a vocal style in which a solo singer 'yoiks' about people, animals and nature. Yoik is highly improvisatory and often syllables without linguistic meaning are used. The musical structure consists of short, cyclic motifs that often belong to the major pentatonic scale; large interval leaps (especially perfect fourths and fifths) are not uncommon.

Altogether, eight yoik excerpts were chosen for the study. The excerpts were selected on the basis that they were followed by strong, possibly unexpected continuations

(1) The *data-driven processing* is instantaneous and consists of basic principles of perceptual organization. For example, grouping of auditory events is achieved by pitch proximity and temporal contiguity. *Schema-driven processing* of sound is learned behaviour, which involves the detection of familiar acoustic patterns, such as tonal functions of tones in a key or phrasing structure, which then guide the organization of relevant auditory details (Bregman, 1990).

(2) Carol Krumhansl from Cornell University, Jukka Louhivuori, Pekka Toivanen, Petri Toiviainen, Topi Järvinen and the first author from the University of Jyväskylä, and Annukka Hirvasvuopio from the University of Tampere, Finland.

that would perhaps be known to the expert listeners but not to non-expert listeners. All the excerpts were transposed to have a common tonic note of C. The behavioural experiment was similar to earlier research on musical expectancy (Krumhansl, 1995b; Schmuckler, 1989). The procedure is that each context melody is presented to the listeners many times, followed on each occasion by a single probe-tone. Using a rating scale, the listeners assess the degree to which the probe-tone fits their expectations about how the melody might continue. In our experiment, the range of these probe-tones was from G_3 to E_5, resulting in a total of 22 probe-tones for each excerpt.

Three groups of listeners participated in the study on North Sami yoiks. The first group consisted of Sami listeners, *i.e.* seven native Sami who had extensive experience in yoiks. The second group was made up of Finnish music students from the University of Jyväskylä. Although the members of this group are predominantly trained in Western music, they are also relatively well versed in the yoik style due to an ethnomusicological emphasis at the Department of Music. The third group consisted of Western musicians unfamiliar with the yoik style. The exact procedure, materials, models and results are detailed in the original article (Krumhansl *et al.*, 2000). Note, however, that the previous study does not include auditory modelling, which will be dealt with later.

Extending the Previous Study: Listeners with less influence of the Western Musical Tradition

In extending the previous study, two immediate considerations present themselves. Firstly, in the modern world it is difficult to find listeners who are not exposed to Western music. Secondly — assuming that one can find such listeners — if the aim is to compare models relating to melodic expectation, one has to be sure that the listeners are actually familiar with the concept of melodies. In actual fact, division of the frequency spectrum of sound into discrete pitches, allowing the replication of these pitches in each octave (octave equivalence), and combining these pitches in various ways does seem to be common to almost all music cultures even if the fine details of this organization may vary (*e.g.* having different tuning systems). Admittedly, the didjiridu of aboriginal Australia uses timbral contour patterns as principles of musical organization (Kendall & Carterette, 1991). In this music culture, questions of expectancies in terms of discrete pitches would not be relevant, at least to this instrument. Nevertheless, most musical cultures do organize music in terms of discrete pitch categories to create melodic sequences. Although most of these musical cultures do not use the Western equal tempered tonal system, a common foundation is a logarithmic system, with an octave being noted as a universal property (Dowling & Harwood, 1986; Carterette & Kendall, 1999). Hence, there is no lack of musical cultures that would be suitable for cross-cultural comparison of melodic expectations.

The rationale behind selecting South African participants for the current study is as follows. In 1999, two South African musicians visited the Department of Music in Jyväskylä, Finland, and heard performances of North Sami yoiks in a concert. The musicians asked for more details about this musical style, having found the style foreign but still approachable. They experienced certain elements in the yoiks — particularly the pentatonic scale structure, the singing style, and use of a small number of anchor tones — as bearing some resemblance to the musical styles they were familiar with. Hence there was basic agreement on what parameters were expected in music and a sense that the scale systems were not impossibly remote. This led us to think that South African participants might be willing to participate in the kind of study we envisaged, even if they were not very familiar with Western musical styles. We then contacted the South African ethnomusicologist Edward Lebaka from the University of Pretoria, who has studied *Pedi* traditional healers and their musical traditions (2001). The Pedi are one of communities of people living in the Limpopo Province of South Africa, in an area bordered by the Limpopo, Vaal and Komati Rivers. Other linguistic communities (*Xitsonga, Amashangana, TshiVenda*) live also in the Limpopo area. *The Pedi* have their own language, Sepedi, also known as Northern *Sotho*, and possess cultural and musical traditions that will be briefly described below.

All in all, the traditional Pedi healers meet the goals of this project perfectly as they are all musicians (performing their traditional music regularly, almost every weekend) and they live far from the major cities in a relatively remote, rural area in the northern part of South Africa (see the map in Figure 1). In that region the number of radios and television sets per village is low and thus the local people, including the traditional healers, have relatively little exposure to Western music, as the consequence of the lack of electricity[3]. Also the reception for radios, televisions and mobile phones is deplorable, because it is a mountainous area. The reasons for the relative lack of contact with Europeans during the late nineteenth and early twentieth century included the mountainous environment and political disputes that dampened colonial ambitions (Blacking, 1962). Even today access to the area is difficult and the Pedi are able to sustain their traditional ways in many areas of life, including music and traditional healing customs. Due to high unemployment rate in this area, many parents and some of the youth are living a hand to mouth existence. As such it is difficult for them to pay a visit from their villages to the cities, where they could easily be influenced by the Western musical tradition. Moreover, the traditional healers are forbidden to listen to the radio as "Healers don't listen to radio, they are not allowed, because it is not important! They need to be 'pure', focus on the other/spiritual world, and they don't have time or the inclination" (Masoga, personal communication, October 22, 2006, Jyväskylä, Finland).

Traditional healers have an esteemed position in southern African societies.

(3) The houses in the Sekhukhune area, where traditional healers live were only electrified in 2001 and the experiment was conducted in 2000. Very few families who could afford to buy radio were using the batteries. For the television they were using petrol generators.

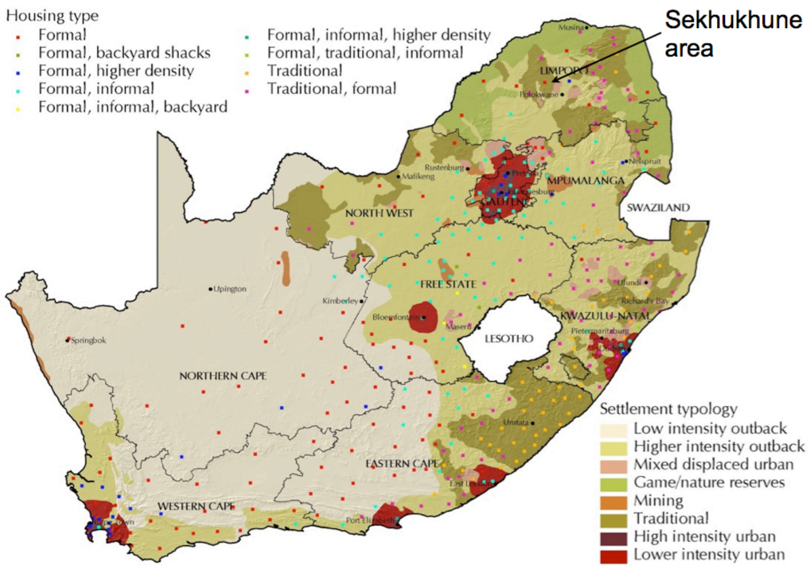

Figure 1.

Map of human settlements in South Africa showing Sekhukhune area in Limpopo province with a cross (SOE, 2008).

People seek their advice for problems ranging from social dilemmas to major medical illnesses. In their diagnostic practice, they drum, dance, and follow specific rituals (Hewson, 1998). These rituals involve ancestral spirits that are called up by dance, music and drumming (Janzen, 2000). In Pedi tradition, music and dance are integral parts of the work of the healers (Lebaka, 2001; Huskisson, 1958). The music of the Pedi people consists of call and response singing and drumming, including polyrhythms that arise both in the accompaniment and in its relation to the singing (Kaemmer, 1998). The national dance of the Pedi is the *Dinaka*[4] (Pedi reed pipes), which involves the rhythmic movement of the legs and drumming to the music of end-blown pipes with different pitches (pentatonic or heptatonic scale) played in *hocket* — *i.e.*, with the individual tones of a melody alternating between different players in close succession (Kaemmer, 1998). Sometimes the spectators support the performers by hand clapping. The musical scale used appears to be basically pentatonic (Barz, 2003) although Blacking (1970) has characterized the music of the neighbouring *Venda* people as strongly diatonic. However, this may be a result of Christian missions, which have had a strong influence on African music in general (Kaemmer, 1998, p. 718): the missionaries brought with them Christian hymns and

(4) Also known as *Kiba*, which encompasses a large variety of musical styles in the diverse cultural communities. For detailed information about the other musical genres, see Masoga, 2006).

military music, and they often translated the hymns into African languages while discouraging the use of African music (Kaemmer, 1998, p. 716).

In fact, the influence of the Christian missions is quite complex. The elements they brought blended with indigenous southern African music and formed the *makhwaya*, meaning choral singing that uses formalized western four-part harmonization and tempered intervals (Coplan, 1998; 2003). It is true that the Pedi traditional healers too have encountered the Western scale and four-part harmonies. However, they do not perform such music themselves.

Another advantage of working with the Pedi is that their language is not a tonal language, or at least not to the same extent as some languages in Western Africa. The speech tones in West African languages have a stronger effect on melodic structures (Kaemmer, 1998, p. 701) than in southern Africa, although tonal patterns of speech have been related to the beginnings of musical phrases in southern African music (Rycroft, 1983). All in all, while it is true that the cultures involved in the present study (South African, Sami, and Western) do share some commonalities in the melodic domain, for instance, reliance on pentatonic and diatonic pitch structures and sung sequences, many other musical features, such as rhythm, may display large, possibly incommensurable differences.

Function of singing in Pedi culture

Each Sekhukhune traditional healer receives "a personal song" from ancestors either before or while undergoing training as a traditional healer. The personal song has exceptional ritualistic value for the particular healer. For instance, should the traditional healer "faint" or "pass out into a trance", it is on hearing his / her personal song as performed by colleagues that recovery is induced and finally gained (Lebaka, 2001). Every song has a different function in the ritual. For example the song "Salane" in its original context was intended to be sung as an epilogue. "Salane" is also recited when problems are encountered by the traditional healers with respect to the divination bones. For example, when the traditional healer is unable to interpret divination bones, the song is recited. In most cases amongst the Pedi traditional healers singing, dancing, hand-clapping and instrumental playing (such as drums and whistles) are synchronized together (Lebaka, 2001).

Musical and other parameters (dance, lyrics *etc.*) in singing in Pedi culture has a different role compared to what we as Europeans are familiar with. The focus is not in the melody, but more in the rhythm and how singing goes together with dance. Also the role of improvisation is great. For a person not familiar with the singing style the melodies sounds monotonous and repetitive in nature. Only after careful listening and learning it comes easier to make a difference between different songs. During the interview of Simon Sete, a traditional healer, one of the authors (JL) asked about how they are able to distinguish melodies from each other. When a specific song was played to him from a video clip, it took him a while before he was able to recognize the song. The decision was made more on the basis of the lyrics and the rhythmic aspects of the song than on melodic structures.

Another occasion where the focus on other than melodic aspects of the music, was clear when we asked Mogomme Masoga to sing simultaneously with the videotape recorded during the ritual in Sekhukhune in 2002. This singing was recorded for further analysis, and parts are displayed in Figure 2. The panel A of the figure displays the fundamental frequency of the typical melodic pattern repeated five times and the panel B contains the notation of this pattern. Panel C depicts the intonation of the singing and panel D is the intonation pattern from the Dinaka reed pipes. This information contained within this figure will be explained later in detail. In addition to singing, he frequently changed singing into imitation of rhythmic patterns. We interpreted this as a sign of a shift of focus from repetitive melodic phrases into probably musically and functionally more important rhythmic structures and drumming. There was a clear tendency to imitate rhythmic patterns instead of singing melodic variants of the songs heard in the tape.

The songs are in pentatonic scale, with no interval smaller than a tone. The melodic outline has to a large extent, evolved to conform with the speech patterns of the language. There is no key change in the course of the tune, but a distinct feeling of the tune hovering round a central note. The Pedi originally had only one word, "segalo" which applied to all three elements, "accent", "pitch" and "tone", and another "tuma nnoši" to the "vowel sound" (Lebaka, 2001). The melodic structure and pentatonic character the songs can be seen in the transcription of one typical melody line sung by Mogome Masoga (Figure 2). It is noteworthy that at the same time participants of the ritual sing many different variants. Thus the version Mogomme Masoga uses is only one example of many possible variants.

Scale and pitch usage in Pedi singing

To investigate the scale, intonation and hierarchy of pitches of Pedi traditional music more carefully, we analyzed two samples of field recordings of the same group of Traditional healers that will act as the informants in the experimental part. In the first sample, the video recordings with high-quality audio track were played back to Dr Mogomme Masoga, a Traditional healer himself, who sang with the other healers in the video and his singing was recorded. This allowed noise free extraction of fundamental frequencies of his singing. The second sample was Reed Pipe Kiba Group performing 7-8 common melodic patterns. To obtain information about the intonation (tuning), F_0 estimation algorithm was applied to the audio excerpts using MIR Toolbox (Lartillot & Toiviainen, 2007). Frequency histogram is shown in Figure 2 (panel D). This analysis largely confirms the music-theoretical notions about their music (the pentatonic scale structure). A pitch-class distribution was also obtained from these analyses by multiplying the frequency distribution with Gaussians distributions assigned to tempered pitch-classes, resulting in pitch-class profile representing the musical style.

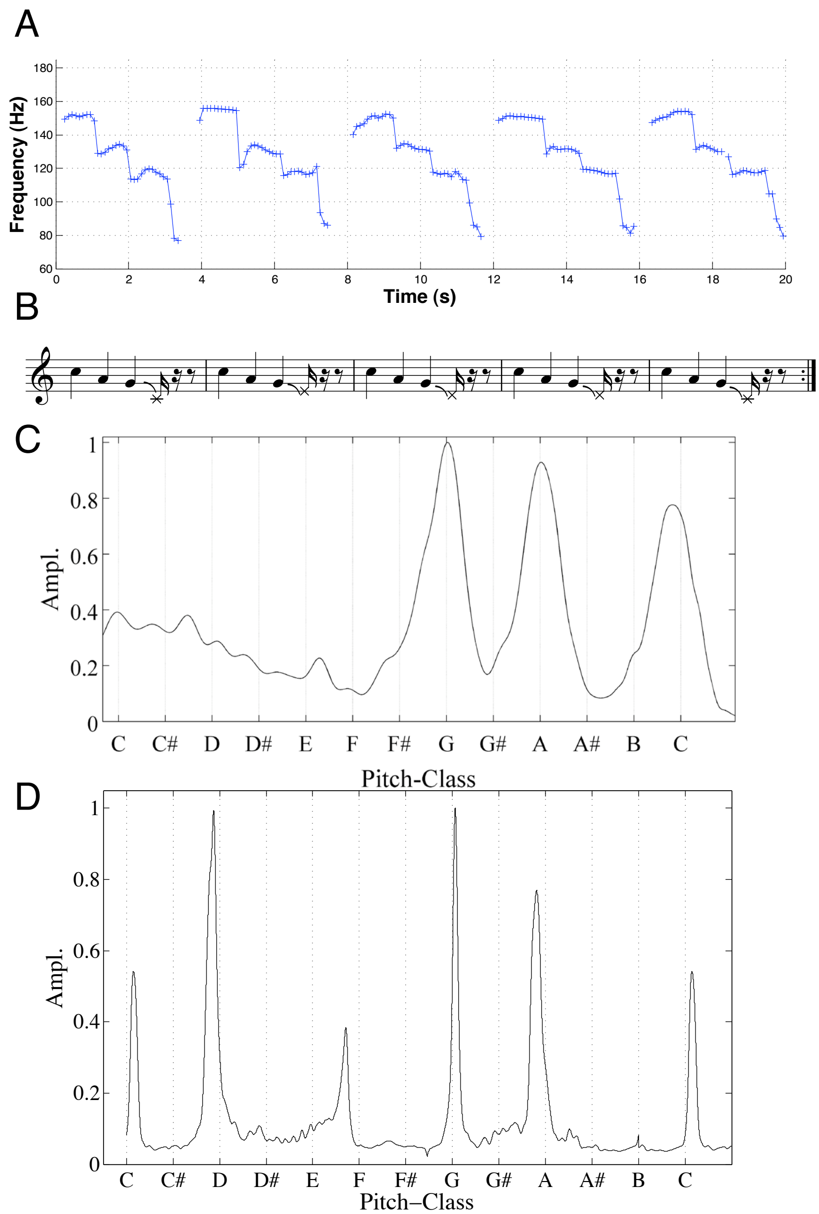

Figure 2.

An excerpt of a Pedi traditional healer song performed by Mogomme Masoga simultaneously with an authentic performance recorded in Sekukhune area in 2002. A) Melodic contour (median filtered F_0), B) Notation of the excerpt, C) Intonation of the pitches for the same excerpt (transposed to G as the central tone), D) octave collapsed F_0s of a recorded sample of five Dinaka Reed Pipes (also recorded in a field trip to Sekukhune area in 2002) that form the pentatonic scale.

The Auditory Memory Model

Recently, Leman (2000) has presented an auditory modelling approach to account for the fitness ratings of the original probe-tone studies by Krumhansl & Kessler (1982). The model extends the work by Huron and Parncutt (1993) by providing a framework of expectations based on acoustic memory. The auditory memory model relies entirely on data-driven processing.

The model postulates several stages of processing. First of all, the auditory periphery is simulated using a model created by Van Immerseel and Martens (1992). In this, the outer and the inner ear are simulated using filters; the basilar membrane is simulated using 40 equidistantly spaced channels (on a critical band scale). The signals in these channels are converted into firing probabilities for the hair cells. The resulting signal is called the *auditory nerve image*. Next, periodicity information is obtained from the primary image using an autocorrelation technique. This results in a *periodicity pitch image*. Finally, the auditory pitch image is processed using two types of short-term memories. The first memory is the lingering *sensory memory*, extremely brief in duration and episodic in nature. The second memory is called the *echoic memory*: this includes the global auditory pitch images over the whole sequence (integrated over time). The duration of both memories (echoes) is indicated by the half-decay of time of the memory (T, 0.1 s for the sensory memory and 1.5 s for the echoic memory). These two memories are compared at the offset time of the probe tone by means of a correlation coefficient to obtain a degree of fit of the probe-tone to the context (Method II in Leman's article). The articles by Leman (2000); and by Leman, Lesaffre & Tanghe (2000) should be consulted for a complete explanation of this model. In order to obtain predictions based on the model, the IPEM Toolbox (a collection of functions for Matlab, see Leman *et al.*, 2000) was used.

In Figure 3, rows a-c illustrate (a) the waveform of a yoik (*Anden Inga*) and the probe-tone, (b) the converted auditory nerve image and (c) the periodicity pitch image. Two versions of the periodicity pitch image were created, each having different memory decay (sensory and echoic memory images, rows d and e in Figure 3). The fit of the probe to the context can be explained as a fit between these two memory images at the time of the probe-tone, achieved by means of a correlation coefficient at the offset time of the probe-tone (in Figure 3, column f shows the correlation between the sensory and echoic memory images, resulting in a low correlation [.35] at the time the probe-tone $A_{\sharp 4}$ is heard). The half-decay values for these memories, given above, were the same values that were used in Leman's reanalysis of the probe-tone study (2000). This operation was performed for each yoik and each probe-tone.

The application of this model extends the scope of the original yoik study and is especially interesting because the model has high explanatory power in Leman's reanalysis of the first set of probe-tone studies. In other words, the most straightforward explanation of the fitness ratings lies in the way sensory memory and echoic memory match at the time of the probe tone.

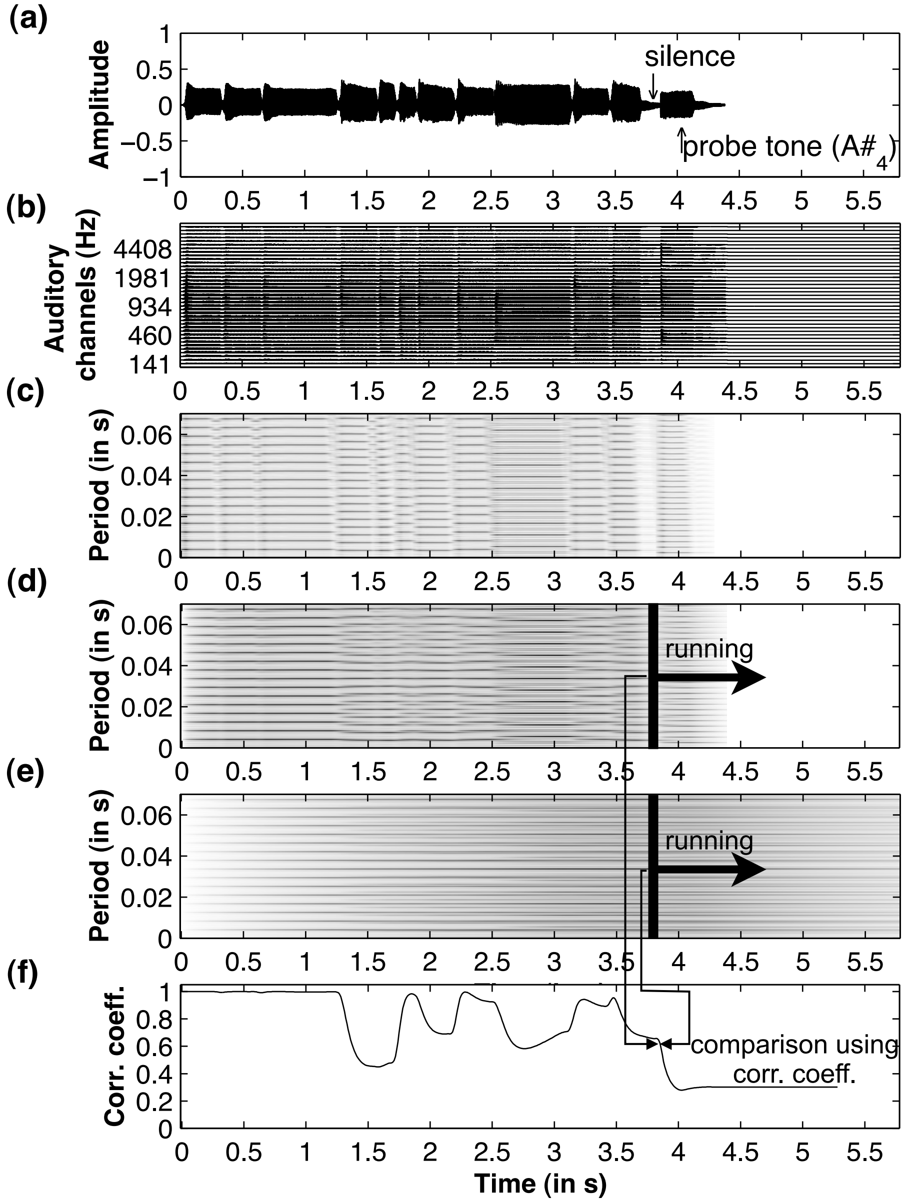

Figure 3.

*Pitch images for the yoik (*Anden Inga*) followed by a probe tone ($A_{\#4}$). From top to bottom: (a) audio signal, (b) auditory nerve image, (c) the periodicity pitch image, (d) the sensory memory image, (e) the short-term memory image, and (f) the comparison between sensory and echoic memory images.*

Hypotheses

The next section summarizes the hypotheses of our addition to the original study, extending from data-driven to schema-driven processes. The hypotheses concern mainly non-Western listeners, but the hypotheses concerning the auditory memory model will be specified also for those listeners who were detailed in the original article.

The data-driven models come in three varieties: (a) the auditory memory model, (b) the event-frequency model, and (c) the implication-realization model (Narmour, 1990). According to our hypothesis, data-driven information will be important for all groups of listeners, but its contribution will be more prominent for those groups that do not possess stylistically appropriate knowledge. This is because those unfamiliar with the musical style in question will have to base their fitness judgements more directly on the surface details of the contexts. By contrast, experts who are familiar with the music can let their stylistic knowledge determine their expectations; consequently the explanatory power of the data-driven models should be lower for these individuals. This was in fact the case in the original study where both the implication-realization model and the event-frequency models correlated inversely with the level of stylistic knowledge (highest correlations for the Western listeners, then the Finnish listeners; lowest correlations for the Sami listeners). By this reasoning, the African ratings should be at the same end of the continuum as the Western listeners. In addition, the implication-realization model should be able to explain the ratings of the African listeners, at least to the same degree as was observed in previous studies, since it has been claimed to be automatic and innate (Narmour, 1990), and has been found to apply across different cultural backgrounds (Krumhansl, 1995b; Krumhansl *et al.*, 1999, 2000; Roh & Yi, 2000).

In the original study, the components of the implication-realization model correlated most strongly with the responses given by the Western listeners, except for the principle of consonance (Krumhansl *et al.*, 2000, p. 38). The Sami participants' expectancy for consonant tones was explained by the combination of pentatonic scale structure and the predominance of large consonant intervals — perfect fourths and fifths — in the yoik style; these features were reflected in the fitness ratings given by the experts, on the basis of their stylistic knowledge.

Concerning the remaining models, the auditory memory model is hypothesized to operate in a similar way to the other two data-driven models. The final data-driven hypothesis concerns the event-frequency models (involving tone and interval distributions) that have been strong predictors of expectancies in previous studies (Oram & Cuddy, 1995) and have also been noted to transcend cultural boundaries (Castellano *et al.*, 1984; Krumhansl *et al.*, 2000; Krumhansl, 2000). In other words, listeners who are unfamiliar with the musical style in question can form an abstraction of the musical style based on frequency information. Thus we have the hypothesis that listeners across musical cultures should be sensitive to the statistical properties of melodies; or more precisely, the effect of statistical information should be more pronounced for the non-expert listeners, as they do not possess other forms of knowledge for that style.

Regarding schema-driven models, there are two varieties: (a) the western schematic model and (b) yoik models, that is, models derived from the yoik style. The first category is represented by results from a study by Krumhansl (1995a) in which intervals were followed by a third tone, which listeners rated in terms of how well it matched their expectations for how the interval would continue. The experiment produced consistent ratings, demonstrating that these minimal contexts were sufficient to invoke a full set of expectancies for Western listeners. Studies using production methodologies demonstrate similar patterns (Carlsen, 1981; Schmuckler, 1990; Thompson, Cuddy & Plaus, 1997) and thus the rating data obtained by Krumhansl (1995a) will be used here as representative of the general, Western schematic model. In the original study, the magnitude of the correlations between the Western schematic model and the ratings given by three groups of listeners brought out a clear continuum (Western listeners .68, Finnish .58, and Sami .47). On the basis of this, one would expect the African responses to be on the low end of the continuum (the Sami end) as they do not possess information on schematic Western continuations. The other category, yoik models, is represented by the common tone patterns (single-tone, two-tone, three-tone distributions) in a sample of yoiks and also the correct subsequent tones in the particular yoiks used in the experiment. The typical statistical properties of yoiks are hypothesized to pick out those who possess stylistic knowledge, although it has been shown that these contingent probabilities can be grasped rapidly by listeners unfamiliar with the style (Castellano *et al.*, 1984; Oram & Cuddy, 1995). Finally, the African listeners' ratings are hypothesized not to correlate with the correct tones any more than Western listeners' responses, the latter group also being completely unfamiliar with the yoiks.

Behavioural experiment

Method

• **Participants.** Due to the complex organization involved in getting the South African participants together at the same place and at the same time, the experiments were conducted on two separate occasions (December 2000 and October 2002). The South African sample included 15 women and 16 men. All the South African participants, hereafter referred to as the *African* group, were born and grew up in the Pedi culture in Limpopo Province, which is situated in the rural, northern region of South Africa. The age range was 15-69 years (mean 42.7 years, *SD* = 14). None of the participants had undergone formal music training but all of them had been involved in musical activities as part of their work as religious healers.

In present-day South Africa, which is a multiracial nation, there are eleven official languages. The majority of the participants informed us that their native tongue was Northern Sotho (Pedi), a language spoken by 3.8 million people in South Africa (Katzner, 1986). None of the participants was familiar with yoik style. Their participation in the experiment was voluntary, and the groups were paid compensation for their efforts on a collective basis.

• **Apparatus and stimulus materials.** The stimuli were the same as those reported in the Krumhansl *et al.* (2000) study, except that in order to shorten the duration of the experiment, only five out of the eight yoiks in the original study were used on this occasion. The selection of these five yoiks was made by looking at the intersubject correlations of the ratings given by the groups in the original study for each yoik, discarding those with the lowest agreement. The five remaining yoiks were *Anden Inga, Elle Sunna, Čappa Magdalena*, *Haldi*, and *Bierra Bierra* (shown in Figure 4). Otherwise, the stimulus material was identical to the original study (using performed excerpts that were synthesized using an English horn sample, with continuation tones consisting of the 22 chromatic scale tones in the range G_3 to E_5, where the numeral specifies the octave; C_4 = middle C) apart from the presentation technique, in which the original audio tracks of the experiment trials were recorded on a CD, together with a spoken voice indicating the number of the trial. Two new random orders were used (one for each group session).

• **Procedure.** The procedure followed the one reported earlier (Krumhansl *et al.*, 2000) but minor alterations were necessary. The participants were told that they would hear excerpts from melodies, starting at the beginning of a phrase and stopping in the middle of a phrase. Their task was to evaluate how well the continuation tone matched their expectations about what might follow in the melody, using a scale from 1 (“extremely bad continuation”) to 7 (“extremely good continuation”). These ratings will be referred to below as *fitness judgments*. Graphical symbols familiar to the participants were added to the response sheets (thumbs up or down sign at both ends of the visual rating scale) to ensure that the scale was used in a consistent manner. The whole experiment was carefully explained to the native expert, Edward Lebaka, who took care of the translation during the experiment. The actual response sheets were not translated, as all the participants could speak and read English. The initial spoken introduction to the experiment was given in English but the exact task and the procedure was also explained to the participants by Mr Lebaka in their mother tongue to ensure that all participants understood the task correctly. After the experiment, we interviewed several participants and also gave them more information about the study and its aims. During the interviews, we verified that the concept of melodic expectancies had been properly understood and discovered that the probe-tone task itself had been envisaged as a sort of musical game. The participants also expressed their keen interest in the yoik music and in the research project in general.

Figure 4.

The five excerpts used in the experiment (from Krumhansl et al.*, 2000). The probe positions are indicated by (PP).*

Results

• Similarities Within and Between the Groups

The initial pool of participants numbered 31. Despite careful instructions and practice trials, the responses of six participants had to be discarded from the data analysis. Two participants documented the number of each trial instead of writing down their fitness ratings for each excerpt. One participant used the value of 1 (extremely bad continuation) for 95% of his responses. In addition, three participants were removed from the analysis because of the clear randomness of their responses (the correlations with other participants' ratings were < ±.1 in all cases). The rest of the responses (25 participants) were consistent between the participants as measured by the reliability analysis (*Cronbach's* α = .79, *df* = 109, 2616, $p < .001$). Nunnaly (1978) has suggested that 0.7 is an acceptable Cronbach's alpha, though some other sources give 0.8 as the criterion value (Carmines & Zeller, 1979; McGraw & Wong, 1996). The consistency index obtained in the present study borders on the higher threshold values found in the literature, though the consistency is lower than that obtained in the original study (average intersubject correlations between .51 and .72 in the present study, as opposed to α between .88 and .92 in Krumhansl *et al.*, 2000, p. 33).

As the focus of attention is on the overall trend of the responses, the mean ratings of the group will be used in further analyses of the data. Let us first consider the correlations between the different groups of listeners (Table 1).

Table 1
Correlations Between the Fitness Judgements Given by the Four Participant Groups

	Sami	**Finnish**	**Western**	**African**
Sami	—	.87	.81	.71
Finnish		—	.92	.82
Western			—	.84
African				—

Note. $p < .01$ in cases, $df = 108$.

The correlation matrix demonstrates the high overall similarity of the responses given by the four groups. The Sami listeners received the lowest correlations to the other groups, probably as they expected certain stylistic conventions to occur in the yoiks. The non-Western group of listeners stretches the previously noted continuum of responses (from Western schematic to expert responses). The greatest dissimilarity of the African ratings was with the ratings given by the Sami listeners (.71), their ratings being closer to those of the other non-experts (a correlation of .84 with the Western listeners). The differences between the correlations of Sami and Finnish (.87) and Sami and Western listeners (.81) were significantly different using Steiger's (1980) test for two dependent correlations, $Z_1 = 3.38$, $p < .01$. Similarly, correlations between Sami and Western (.81) and Sami and African (.71) listeners were statistically different ($Z_1 = 2.92$, $p < .01$). However, the correlation between Finnish and African (.82) when compared to the correlations between Western and African listeners (.84) were not significantly different ($Z_1 = -1.02$, $p = .23$, N.S). Figure 5 below shows the group means and standard errors for all the probe-tones for one yoik, exemplifying the characteristic differences and similarities between the groups.

In Figure 5 we can see the fitness judgements for the yoik *Bierra Bierra*. The ratings given by the African group are most similar to those given by the Western group (high values for G_4, A_4, $A\sharp_4$, B_4, C_5). The African listeners, like the Western listeners, expected the continuation tone to be very close to the last tone occurring in the context (hence $A\sharp_4$, B_4 or even $C\sharp_5$), although the Western listeners did not think $C\sharp_5$ would fit the musical expectancies set by the context. The leading tone (B), which rarely occurs in yoiks, was more favourably rated by the Western and African listeners than the Sami listeners. For their part, the experts expected the genuine

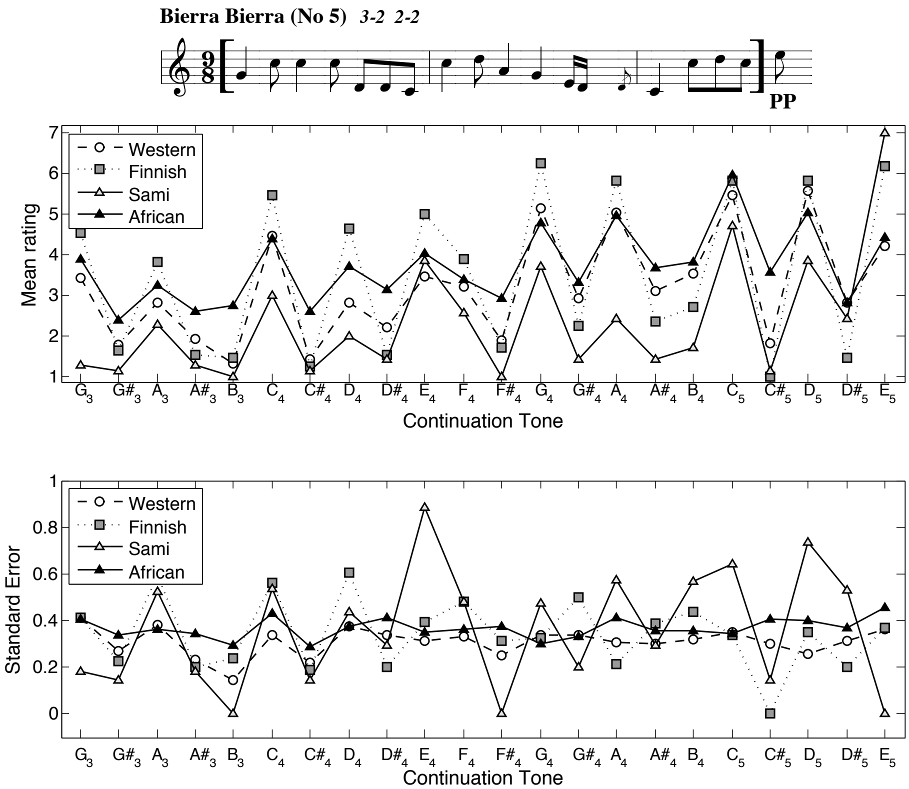

Figure 5.

Yoik Bierra Bierra *and the mean fitness judgements (upper panel) and the standard errors (lower panel) by all groups of listeners for all probe tones.*

continuation tone (E_5) and gave comparatively low ratings to the other tones. It is worth noting that the influence of pentatonic scale (C, D, E, G, and A) may be seen in the all the non-expert groups' ratings. It is worth of note that the standard errors of the African participants are similar in range to those of Western and Finnish participants thus eliminating the possibility that the patterns simply reflect wider range of responses. Below, we shall examine the ratings further, matching them with the various data-driven models and schema-driven models.

- **Data-Driven Expectancies: the Implication-Realization Model**

We hypothesized that the data-driven models would be better at explaining the non-expert listeners' (African and Western) ratings than the experts' ratings. In addition, we expected that the implication-realization model would be able to predict the ratings of the African listeners to much the same extent as it predicted the ratings of the non-expert listeners in previous studies; this is because the model has been claimed to be innate (Narmour, 1990) and to work across different cultural backgrounds (Krumhansl, 1995b; Krumhansl *et al.*, 1999, 2000). To evaluate a simpler version of the model, Schellenberg's two factor model (1997) will also be used to predict the ratings of the four groups. A regression analysis was performed to assess how much of the variance in the fitness ratings of all groups could be explained by two versions of the implication-realization model (as coded by Krumhansl, 1995a and Schellenberg, 1997). The results relating to the implication-realization model are summarized in Table 2.

Table 2
Comparison of Implication-Realization Model Principles (Original and Revised) Using Regression and Showing Squared Semipartial Correlations (sr^2) for Individual Principles

	Sami	**Finnish**	**Western**	**African**
Original (*df* = 105,4)	R^2 = .22	R^2 = .15	R^2 = .28	R^2 = .29
Registral Direction	.04*	.01	.01	.01
Intervallic difference	.09*	.06*	.07*	.06*
Proximity	.06*	.03	.01	.00
Unison	.07*	.05*	.03*	.05*
Registral return	.00	.00	.01	.01
Revised (*df* = 108,1)	R^2 =.12	R^2 =.09	R^2 =.26	R^2 =.20
Proximity (revised)	.03*	.05*	.16*	.12*
Pitch reversal	.03	.00	.02	.00

Note. * = significant predictor at $p < 0.05$ level.

In Table 2, we can observe that all the principles together could explain 10-30% of the variance in the fitness ratings of the four groups. The overall difference between the two models (original and revised) are fairly small (magnitude of 7% across the groups), suggesting that the simpler, revised model offers a more parsimonious explanation of the data. Moreover, the regression model results for the African group are similar to that of the Western group for both versions of the implication-realization model. This suggests that at least some of the principles in Narmour's model did apply across cultural boundaries. There were, however, two principles in original model which did not correlate significantly with the ratings of any of the groups (Registral Return and Registral Direction). These have also failed to predict ratings in earlier studies (Schellenberg, 1996; Krumhansl *et al.*, 2000; Schellenberg *et al.*, 2002). In fact, von Hippel (2000) has provided a simpler explanation regarding Registral Direction and Proximity, a point which will be addressed later. Schellenberg's two factor model contains one principle (Proximity, revised) that accounts almost as much variance as the original principles together if one looks at the ratings of the two groups that do not have appropriate stylistic knowledge (Western and African participants). Therefore this variable will be used in further analysis of the data to represent data-driven expectations. Moreover, the prediction rates were consistently higher for the two non-expert groups than for the expert groups, suggesting that individuals who do not possess appropriate stylistic knowledge base their fitness ratings more on local pitch structures than do experts.

- **Data-Driven Expectancies: the Auditory Memory Model and Event-Frequency Models**

According to the hypotheses, the auditory memory model should predict the responses of non-expert groups better than the expert groups, since the former do not possess appropriate stylistic knowledge of the yoik style and may therefore form their fitness evaluations on the basis of the acoustic similarity between the probe and the context. The auditory memory model has been found to be robust in explaining probe-tone ratings for chord contexts (Leman, 2000) but it is uncertain how far the model can account for the ratings given to melodically rich musical contexts like the ones in this study.

Table 3 displays the correlations between the auditory memory model, the event-frequency models (*i.e.* including the frequency of tones and intervals in the contexts), the schema-driven model and the ratings by the four groups of listeners. If we compare the correlations between the auditory memory model and the ratings for the other data-driven models, we notice that the auditory memory model was not particularly successful in explaining the probe-tone fitness ratings given by any group. Admittedly, when we look at the expert and non-expert groups there is a difference that would fit the hypotheses, in terms of high correlations for the Western and Finnish listeners. This could imply that these groups rely more on acoustic information than the Sami listeners. However, doubt is cast on this explanation by the fact that the correlation of the auditory memory model with the African listeners' ratings is weaker than with the other non-expert groups' ratings. One way of examining this phenomenon could be to control for the variables in Narmour's model by partialing them out in the correlation analysis. Yet this manoeuvre did not remove the unexpected relationship between the Auditory memory model and the African fitness ratings (Sami $r = .10$, $p > .05$, NS; Finnish .24, $p < .05$; Western .30, $p < .01$; and African .13, $p > .05$, NS; $df = 108$ in all cases).

The most plausible reason for the modest success of the auditory memory model is that these yoik melodies evoke a sense of musical context. This is a feature which distinguishes the current experiment from the initial probe-tone studies, in which

Table 3
Simple Correlations Between The Groups and Predictors

	Sami	Finnish	Western	African
Data-driven: auditory				
Auditory model	.27**	.34***	.34***	.28**
Data-driven: event-frq.				
PCs in contexts	.50***	.55***	.64***	.59***
IVs in contexts	.40***	.41***	.48***	.46***
Data-driven: impl. -real.				
Proximity	.30**	.29**	.51***	.45***
Pitch Reversal	.30**	.18	.31***	.28**
Schema-driven: event-frq.				
PCs in yoiks	.70***	.77***	.67***	.61***
2-tones in yoiks	.60***	.63***	.71***	.72***
3-tones in yoiks	.64***	.55***	.60***	.61***
Western schem.	.44***	.58***	.66***	.53***
PCs in Pedi	.57***	.57***	.45***	.41***
Idiosyncratic				
Correct tones	.33***	.23*	.14	.17

Note. * = significant correlation at $p < 0.05$ level, ** at $p < 0.01$ level, *** at $p < 0.001$

the context was kept stylistically neutral by using Shepard-tone timbres, and by establishing a sense of key through the presentation of typical cadences or complete major or minor scales. The auditory memory model may be better suited to predicting the fitness ratings with these kinds of stylistically neutral key-defining contexts in which the immediate context of chords may explain the ratings. In the present case, the yoik melodies with their recurring musical patterns probably induce expectancies that are influenced by a longer musical context (*i.e.* the specific tones and intervals used in the whole excerpt) than just the match between the different auditory short-term memories. An example of this was presented in Figure 5, where the African participants rated the probe-tone g favourably, presumably because it had occurred often in the context, even though it did not immediately precede the probe position.

The tone and interval distributions in the yoik contexts demonstrated patterns of correlations (also shown in Table 3) that were more in accordance with the hypothesis that the auditory memory model should match the non-expert groups better than the expert groups. First of all, we can see that the ratings of the Western and African groups correlated slightly more highly with the tones that had occurred within the yoik contexts than did the ratings given by the Sami or Finnish groups. The same holds true for the intervals. This difference, though minor in magnitude, indicates that the non-experts formed their melodic expectancies more directly from the information provided by the musical context than did those subjects who had knowledge of the yoik style. In other words, listeners seem to pay attention to event frequencies in the melodies, and frequently occurring events act as cognitive reference points that are established by repetition. Similar results were observed in a cross-cultural study by Castellano *et al.* (1984). In their study, Western listeners adapted to the frequency information provided by North Indian musical contexts. More direct frequency effects were found by Oram and Cuddy (1995) in their direct manipulation of frequencies within musical contexts. In other words, listeners who are unfamiliar with the musical style in question do adapt to the structural regularities of contexts, as proposed by the hypothesis, and in agreement with conclusions of previous studies (Krumhansl *et al.*, 1999; Castellano *et al.*, 1984).

The correlations between the principles of the two-factor implication-realization model and the participants ratings can be found on Table 3. A proposal by von Hippel (2000; also von Hippel & Huron, 2000) explains the principles of pitch proximity and pitch reversal (Registral Direction) using an ecological model; one which is rooted in the physiological constraints involved in music making. Von Hippel uses the terms *tessitura* and *mobility*, the first of these relating to the fact that the median range of the melody tends to be favoured and thus more expected. Mobility for its part explains why melodies change direction after large leaps by simply observing that they would otherwise run out of the comfortable melodic range. The predictive power of these two interrelated formulations was tested in the current material. Tessitura was quantified as the distance (in semitones) between the probe-tone and the median pitch height of the melody. Mobility uses a lag-one autocorrelation between successive pitches to predict the pitch height of the probe-tone relative to the last tone of the context. Hence, tessitura predicts whether listeners expect probe-tones close to median pitch height, whereas mobility takes into account the extent to which the pitch height is predictable from the previous pitch heights. The predictions of tessitura correlated statistically significantly with the Western listeners ($r = .32$, $p < .01$, $df = 108$), and with the Finnish and African listeners (.24, and .23, $p < .05$); however, the correlation between tessitura and the Sami listeners was not statistically significant (.16, $p > .05$, NS.). For its part, mobility did not correlate with the ratings given by any of the groups. In this case, the predictions from both models were lower than those obtained from Narmour's implication-realization model.

- **Schema-Driven Expectancies: the Western Schematic Model and Yoik-Related Models**

In order to determine (a) whether the Western schematic model truly contributes to Western melodic expectancies, and (b) whether statistical information on the yoik style in general relates to the ratings given by experts, the predictions of these models were correlated with the actual probe-tone ratings given by each of the four groups of listeners. The Western schematic model was represented by the results of a probe-tone study by Krumhansl (1995a), in which Western listeners rated the continuation tones of two-tone contexts. The yoik-related models consisted of a number of yoik-related statistical variables that were derived from an analysis of a small corpus of yoiks (see Krumhansl *et al.*, 2000, p. 21-25); this corpus was believed to contain a reasonably representative selection from a more exhaustive sample of yoiks (Kantola, 1984). The variables used were single-tone, two-tone and three-tone distributions in the yoik corpus. In addition, a major pentatonic tone profile applied to each yoik in the experiment was used as a yoik-related model. Note that the actual subsequent tone of each yoik was taken to represent expert knowledge of particular melodic continuations within the yoiks.

The Western schematic model revealed a fairly clear continuum, moving from the Western listeners (r = .66) through the Finnish (.58) listeners and then to the Sami (.44) and African listeners (.53; all $p < .001$, $df = 84$) — the last-named group consisting of persons without Western schematic knowledge. The fact that the Sami listeners actually correlated less well with the Western schematic model than did the African listeners may be due to such intervening factor as expectation of the actual subsequent tone in each yoik. Indeed, the correlation of the African listeners with the Western schematic model is surprisingly large. Thus, one has the impression that the Western schematic model encapsulates not only Western expectations, but also proximity and other melodic predictors.

If the principles in Narmour's model are removed from the analysis of Western schematic expectations (by partialing out the principles in Narmour's model in the correlation analysis), the Western schematic model shows a statistically significant correlation only with the Western and Finnish listeners (r = .32 and r = .30, respectively, both $p < .01$ and $df = 108$); the correlations with the Sami and the African listeners are then greatly reduced (r = .08 and r = .13, both $p > .05$, NS., respectively). Consequently, this model supports the notion of Western expectancies. Note that although the Sami listeners were familiar with Western music, they had specific expectancies in mind and thus their ratings were not governed by generic principles to the same extent as those of the Finnish and Western groups.

Next we shall focus on the frequency-based variables derived from a statistical style analysis of a yoik corpus. These models should display higher correlations for those listeners who are familiar with the style (the Sami group, followed by the Finnish group). The first-order statistics (involving single-tone distributions in the yoiks) were in accordance with the hypothesis (r = .70, .77, .67, .61, with Sami, Finnish, Western, and African listeners, respectively, all $p < .001$, $df = 108$). Once again, this points towards a large overlap between the variables. The distributions of tones in the yoik style are highly collinear with the tone distributions in the yoik excerpts used in the experiment ($r = .89$, $p < .001$, $df = 108$). If we control for this overlap by partialing out the tone distributions in the yoik excerpts in the correlation analysis, the correlations between the first-order tone distributions in the yoik style and the ratings given by the Sami, Finnish, Western and African groups still display the same trend, albeit a weak one, in line with the hypothesis ($r = .47$, .51, .44, and .45, all $p < .001$, respectively). These trends were more obscure in higher-order statistical patterns, as evidenced by the correlations between the two- and three-tone distributions and the Sami ($r = .60$, .64.), Finnish (.63, .55), Western (.71, .60), and African (.72, .61, all $p < .001$, $df = 108$) listeners' ratings. The three-tone distributions show a higher correlation with the Sami group that possessed the appropriate stylistic knowledge of the yoik style, but the two-tone distributions in the yoiks actually correlate better with the groups who did not possess any stylistic knowledge of yoiks. We assume that this anomaly is due to a high overlap between the tone transitions in the yoik style and the tone distributions and transitions in the yoik excerpts used in the experiment ($r = .86$ and .63, $p < .001$, $df = 108$).

All in all, teasing out the role of stylistic knowledge — as manifested by the higher-order statistical patterns — in the formation of expectancies in yoiks is difficult, since all the groups appear to be sensitive to the distributions of tones and the pattern of tones in the yoiks. Krumhansl has argued (2000) that the first-order tone distributions do not provide enough information to distinguish styles from each other, whereas the second- and third-order tone distributions are more revealing in this sense. It is also worth pointing out that these stylistically appropriate variables explain the fitness ratings of the probe-tones better than the auditory model does; in addition, they provide a better explanation than would be obtained purely from the tone and interval distributions of the contexts. In a regression analysis, 69% of the variance in the probe-tone ratings of the African listeners was explained by these schema-driven variables ($F = 57.32$, $p < .001$, $df = 3,80$); this is 15% more than in the case of the data-driven variables, suggesting a rather high overall impact of the schema-driven variables. A more detailed comparison of the different levels and models of expectancies will be presented at the end of the results section.

In addition to the statistical patterns in yoiks, a pentatonic tone profile might represent yoik-related knowledge, since yoiks often consist of a major (anhemitonic) pentatonic scale. A statistical study of yoiks by Kantola in 1984 found out that the pentatonic scale represented 71% of the repertoire of 154 yoiks from the Northern Sami area. To explore the relevance of this, tones C, D, E, G and A were coded as 1 and other tones as zero; this variable was correlated with the ratings of all the groups. The coding resulted in high correlations between the variable and the ratings of the Sami and Finnish listeners ($r = .77$ and .80, respectively; both $p < .001$ and $df = 108$), with somewhat lower correlations between this variable and the Western and African listeners (r = .66 and .59, respectively, both p < .001). These correlations are in line with the hypothesis, emphasizing the role of stylistic knowledge and schema-driven knowledge. Moreover, the note distributions in pedi traditional singing extracted from the audio excerpts was also correlated with the responses, yielding moderate correlations for Sami and Finnish listeners (r = .57 and .57, both p < 0.001) and somewhat lower correlations for the Western and African listeners (r = .45, .41, both p < 0.001). This result seems at first counter-intuitive as we expected to find higher correlations for African listeners from a predictor that reflects their own musical culture. However, this finding resembles observations Castellano and her colleagues (1984; also Kessler *et al.*, 1984), and made in similar cross-cultural comparison where the listeners unfamiliar with the style in question were actually using simple heuristics derived from the contexts themselves (note frequency data from Western and Indian music) rather than relying on their own musical culture and its specific predictions.

Finally, knowledge of the actual melodic continuations in these yoiks is represented by a criterion-matching expectation, *i.e.* the one that has been coded for the analysis as the correct tone occurring at the probe-position in the context. The correlations of this variable with the ratings display a clear relationship with levels of stylistic knowledge; the ratings of the Sami (r = .33, p > .01) and the Finnish (r = .23, p > .05) listeners correlate with the variable whereas the African (r = .17, p > .05, NS) and Western (r = .14, p > .05, NS) listeners' ratings do not, as they were completely unfamiliar with the yoik style. Hence the hypothesis concerning the contribution of an expert model was supported.

- **Model Summary**

To assess the individual contribution of each predictor to the ratings of the four groups, a regression approach was taken. However, two problematic issues — multicollinearity and overfitting — for such an analysis needs to be first dealt with. We can assume that many of the 11 predictors outlined earlier them correlate as they are variant representations of the same data. Indeed, the correlations between all predictors, shown in A1 (Appendix), imply multicollinearity between certain types of predictors, especially the statistical descriptions of the yoiks and the contexts. These high correlations are problematic when the goal is to assess the individual contribution of these predictors to the ratings due to predictor redundancy leading to suppression, magnitude reduction of parameter estimates and inflation of standard errors (Morrow-Howell, 1994). To diagnose multicollinearity, variance inflation factors (VIF) were calculated for each predictor (shown in the last column of Table A1). Using a conservative criterion by Fox (1991), $\sqrt{VIF}$ over 2 is considered as an indication of harmful collinearity. Thus PCs in context, 2- and 3-tone transitions in yoiks were excluded from the regression analysis, leaving 8 predictors for the analysis.

Eight predictors, however, leaves the door open for overfitting, as the ratio of predictors to cases is lower than the usual recommendation (20:1, see Hair *et al*, 2006). To bring down the number of variables to five in the present study (110 cases), the three predictors were eliminated based on theoretical arguments. First, in comparing implication-realization model, we observed that only Proximity (revised) contributed to the ratings so we eliminate Pitch reversal. At this point, all data-driven predictor categories (auditory, event-frequency based models, implication-realization) have one predictor whilst the schema-driven models are represented by three predictors (PCs in yoiks, Western schematic and PCs in Pedi). From these three we want to keep PCs in yoiks as it is the most clear-cut example of the stylistically appropriate predictors for yoiks, leaving 5 (correct tones in yoiks is also included) predictors for the final regression analysis, where all the predictors are entered into the regression simultaneously. A summary of this analysis is shown in Table 4, the explanatory power of the models is shown in R^2s, the individual contribution of the predictors is shown using *squared semipartial correlations* (sr^2). For a less strict approach to the regression, we alternatively used all 11 predictors to explain the responses for all groups, shown at the bottom row of the Table 4.

It is evident from Table 4 that the overall prediction rate for the four different groups is fairly similar in magnitude and the actual differences can be observed in the contribution of individual predictors (indicated by sr^2 values). Broadly speaking, the groups who have not been exposed to yoik style (Western & African) are making their responses in terms of data-driven models, especially in terms of the simple pitch proximity heuristic. Interestingly, the auditory model does not contribute to the regression equation for any of the groups. *Intervals in the contexts* seem to make a marginal contribution. Perplexingly, the responses of all groups are best predicted by the representation of the stylistic probabilities (*PC in yoiks*), which represents schema-driven expectations in form of typical pitch-classes in yoik style. We must interpret this more broadly as it is difficult to argue that groups not exposed to the yoiks could be using stylistically appropriate schematic information as such when forming their impressions on the fitness of the continuation tones. This predictor correlates with the pitch-classes in the contexts (see Table 3), which might be part of the explanation. Another plausible explanation is that the contexts evoke a sense of the stylistically appropriate patterns, which are best represented by *PCs in yoiks* predictor and not by the *PCs or IVs in contexts* predictors as they might be too sparsely populated (*e.g.*, typically only 4 PCs out of 12 possible PCs in each yoik). Unfortunately, the current data cannot further disentangle these associations without making a leap of faith in terms of reliability and interpretability of the regression model by reporting the analysis using all the predictors. Finally, the correct tones in these particular yoiks, can be seen to contribute to fitness ratings of continuations of the Sami group and not to the other groups, as expected. The correlations in Table 3 tell a similar overall story, showing low match between the auditory model and the ratings. The data-driven principles (event-frequency) correlate fairly well with the ratings and no large differences are evident in these between the groups. The Western

Table 4

Summary of Regression Results Showing Squared Semipartial Correlations (sr^2) for Individual Principles. Predictors in Italics Denote the 5-Predictor Regression Analysis

	Sami	Finnish	Western	African
	sr^2	sr^2	sr^2	sr^2
Data-driven: auditory				
Auditory model	.01	.00	.00	.00
Data-driven: event-frq.				
PCs in contexts	-	-	-	-
IVs in contexts	.02*	.02*	.02*	.03*
Data-driven: impl. -real.				
Proximity	.05*	.05*	.17*	.12*
Pitch Reversal	-	-	-	-
Schema-driven: event-frq.				
PCs in yoiks	.33*	.41*	.31*	.25*
2-tones in yoiks	-	-	-	-
3-tones in yoiks	-	-	-	-
Western schem.	-	-	-	-
PCs in Pedi	-	-	-	-
Idiosyncratic				
Correct tones	.03*	.00	.00	.00
Summary				
Chosen predictors (5)	R^2 =.64	R^2 =.71	R^2 =.74	R^2 =.61
All predictors (11)	R^2 =.77	R^2 =.80	R^2 =.81	R^2 =.73

Note. * = significant predictor at $p < 0.05$ level.

schematic model correlates most highly with the Western listeners. Although all the schema-driven predictors have all high correlations, a reliable interpretation of their individual role in the expectation might be haphazard due to the remaining correlations between the data-driven event-frequency models.

The explanatory power of the data-driven predictors (proximity and event-frequency models representing the contexts) was twice as high for the Western and African listeners who did not have stylistic knowledge at their disposal than it is for the listeners who knew the yoik style. Although in regression the implication-realization model was reduced into single predictor, this principle of pitch proximity is able to contribute significantly to the expectancies across the listener groups, echoing the findings of earlier studies (Schellenberg, 1997; Krumhansl, 1995a). A more permissive regression model with all the 11 predictors yielded, not surprisingly, slightly higher prediction rate (an increment of about 10% of the variance on average). What is noteworthy in this riskier analysis strategy that the relative magnitudes of the prediction rates between the groups remains virtually the same as in the more stringent regression with 5 predictors.

In sum, a considerable portion of the variance in the ratings of the groups could be explained by a small set of the predictors in the regression analysis. The impact of the variable categories depended mostly on the different levels of knowledge, as predicted. The frequency-based models exerted more influence on listeners unfamiliar with the yoik style, the Western schematic model had most impact on Western listeners, and the yoik models carried most predictive power for those listeners who possessed knowledge about yoiks.

Discussion

In the previous study by Krumhansl *et al.* (2000), the musical background of the listeners (Sami, Finnish and Western listeners) in all cases included familiarity with the Western musical style; thus the separation of expectancies into various innate and schematic components may have manifested generic Western melodic expectancies and not have allowed sufficient room for data-driven, non-stylistically influenced expectations. In the replication, in which traditional healers from South Africa rated the fitness of probe-tones to North Sami yoiks, the overall pattern of results was generally similar to that observed with the Western listeners — who, it should be remembered, were also unfamiliar with the yoik musical style.

The data-driven models were largely independent of stylistic knowledge and were moderately successful in explaining the melodic continuations. A revised version of Narmour's implication-realization model predicted the fitness ratings of African listeners to the same extent as it predicted the ratings of the Western listeners, supporting the claim of the innateness of at least the pitch proximity principle of the model. The auditory memory model for its part was insufficient in explaining the probe-tone ratings in the yoiks, even if the listeners were totally unfamiliar with the style and idiom. The low success of the auditory model was assumed to be a result of the sequential patterns rather than the key or chord-related continuations evoked by the contexts. Support for this explanation came from observations on how the patterns in the contexts (tone and interval distributions) were picked up by listeners unfamiliar with the style. In other words, the statistical information provided by the context is of real importance. This is in line with the findings of Castellano *et al.* (1984), Kessler *et al.* (1984) and the more recent studies of melodic expectations (Oram & Cuddy, 1995; Krumhansl *et al.*, 1999, 2000). It is also consistent with studies on language processing by Aslin, Saffran, & Newport (1998; also Saffran, Johnson, Aslin, & Newport, 1999), in which infants as young as eight months have been shown to learn the transitional probabilities of syllables in an artificial language and tones in artificial tone sequences, during a brief exposure to them.

The Western schematic model was shown to correlate most highly with Western and Finnish listeners. However, both this model was found to overlap with the data-driven models. For example, a closer inspection of the Western schematic model revealed a large overlap with pitch proximity (revised), which renders the inferences from these correlations to be cautious. The typical tone distributions present in yoiks led to a similar dilemma, as this emerged as the strongest predictor even for the non-expert groups — *i.e.* for those who did not know the yoik style and thus could not in fact be expected to rate their expectations on the typical distributions. However, this schema-driven variable overlapped considerably with a data-driven variable, namely, the tone distributions in the contexts themselves. These overlaps in the models may be difficult to avoid in stimulus material derived from — and taken to be representative of — a particular musical style, as there will always be a high correlation between the statistical properties of the chosen stimulus materials and the style in general. This brings out the finding that all groups of listeners appear to adapt relatively easily to the frequency-based information provided by contexts.

The results concerning both the data-driven and the schema-driven models are along the lines of previous cross-cultural investigations of melodic expectancies (Castellano *et al.*, 1984; Kessler *et al.*, 1984; Krumhansl, 1995b; Schellenberg, 1996, 1997; Krumhansl *et al.*, 1999, 2000). However, our study was only the third in which this particular paradigm was applied outside a Western context (the first two were in the comparison of Western and Indian listeners by Castellano *et al.*, 1984, and Balinese and Western listeners by Kessler *et al.*, 1984). In this respect, the overall consensus between culturally dissimilar groups of listeners in perceiving how melodies may be continued is an important indication of commonalities in musical processing.

The extension of the explanatory variables into auditory modelling is an approach that has rarely been applied to the study of expectations (the exceptions are two reanalyses of the early probe-tone studies by Huron and Parncutt, 1993, and Leman, 2000). Although the explanatory power of the auditory model was inferior to pitch-based models in the yoik contexts, the role of acoustic information might have been more important if real sung excerpts had been used. Vocal timbre, prosody and vowels may give important cues to possible continuations. In addition, the role of the auditory memory model may be more important in other musical styles than it is in yoiks.

One possible explanation for the traditional healers' performance at the same level of Finnish and Western subjects in predicting the correct next tone in yoiks may be related to better developed learning strategies in aural memory tasks. This hypothesis is based on observations one of the authors (JL) made during several field trips since 1998 in South Africa, Botswana and Kenya, and during the visits of South African music and dance groups in Finland. In these occasions, several professional Finnish music educators and researchers have had a possibility to teach music for school children, music school students and choristers in rural and urban contexts (*e.g.*, Louhivuori, Salminen, & Lebaka, 2005). An overall impression of these learning events has been an exceptionally fast learning of both melody and text of songs despite the participants' unfamiliarity with the style of music or the language (Finnish). Moreover, remarkably good pronunciation of Finnish language has been observed. A possible reason for this latter observation may be related to widespread multilingualism in these African countries and to existence of strong living oral tradition, which emphasizes of aural skills. For example, the number of official languages in South Africa is eleven (9 indigenous and 2 former colonial languages) although altogether the population belongs to 28 different language groups. Most people in South Africa are at least bilingual as English, although home language of only 9% of the population, has a high status and is the language of political and economical leadership (Mesthrie, 2006).

In addition to the influence of multilingualism, it is hypothesized that many African cultures still bear features typical of oral cultures. This may influence the development of aural cognitive skills and increase the ability to pick up statistical and other characteristic features in music more quickly and precisely than people whose aural cognitive skills are not as developed due to emphasis on visual information processing (such as written text and television in Western cultures). Similar reasoning was already suggested by Castellano and her colleagues (1984), in a study where Western listeners responses to unfamiliar musical style was observed to reflect the probabilities of the new musical style rather than their own fixed images of the own prior experience. Therefore the primary explanation that the underlying similarities of the musical structures (pentatonic scales, large interval sizes) may not be as persuasive explanation for the better performance of traditional healers over Finnish and Europeans in predicting the correct next tone in yoik melodies. The answer may be looked from other directions such as special aural memorization skills developed by people who live in a culture that emphasizes oral transmission. Cultural differences in aural memorization skills and cognitive strategies should be studied more carefully using paradigms such statistical learning experiments (*e.g.*, Oram & Cuddy, 1995) in order to understand and explain the findings in this paper.

A critical assessment of the current study would identify two areas for improvement. The first is in the methodology. Probe-tone methodology is very much a retrospective method, and one which may result in distortion of the expectancies due to the repetition of the context. A recent variant of the method, called the concurrent probe-tone method, overcomes this by asking the listeners provide their fitness ratings whilst the music is actually playing (Krumhansl & Toiviainen, 2000; Toiviainen & Krumhansl, 2003; Eerola, Toiviainen, & Krumhansl, 2002). Beyond this, in studying questions of musical expectancies one could perhaps use methods that may be more natural in a musical sense, such as production methods (*e.g.*, Unyk & Carlsen, 1987; Auhagen, 1994; Larson, 1997). However, production tasks face the problem of production skills, and this could lead to differences between what is actually expected and what can be produced (by singing or by other means). Perhaps the most promising direction is towards methods using reaction times and brain imaging; these have the advantage of displaying implicit and unprocessed knowledge that is difficult to obtain otherwise (see Koelsch, 2005; Koelsch, & Friederici, 2003; also Huron 2006 for other methods). Cross-cultural comparisons of musical expectancies using these methods have not yet been reported, though a few studies have used brain imaging techniques to examine the overall differences between listeners from different cultures in processing music (Cohen, Granot, Pratt, & Barnea, 1993; Demorest, Morrison, Aylward, Cramer, & Maravilla, 2002). More direct investigations of expectancies have been conducted in studies on language processing (*e.g.* Schlosser, Aoyagi, Fulbright, Gore, & McCarthy, 1998; Milberg, Blumstein, Katz, Gershberg & Brown, 1995).

The second area of improvement lies in the premises surrounding the musical cultures studied, including the choice and presentation of the musical stimuli and in relation to whether the exact tasks and the models are appropriate to the musical culture in question. For example, in the current study, the role of the tone distributions in the contexts used could not be distinguished from the tone distributions in yoiks in general, as they were almost identical. Moreover, in our study synthesized excerpts were used in order to reduce the number of variables, but the scope of the variables would ideally be enlarged to encompass the more refined aspects of the music (timbre, accents, intonation).

CONCLUSIONS

To sum up, we conclude that all listeners use data-driven heuristics consisting of (a) dynamic strategies (frequency of events) and (b) innate or at least culturally-transcending heuristics (such as the pitch proximity principle). In addition, depending on their stylistic knowledge, they also use schema-driven strategies to form their expectations about melodic continuations. The data-driven principles are important even if the listeners are familiar with the style. For expert listeners, schema-driven expectancies have a predominant role. Yet schema-driven expectancies often make use of innate heuristics that are available also to non-expert listeners. Although the different sources of expectancies may create conflicting predictions, we did not observe substantial differences due to these sources. However, the assumption of the additivity of different components of expectancies, as observed by Schmuckler (1989), needs to be addressed in future studies.

The pervasiveness of expectancies has generated a certain amount of theoretical discussion concerning the origins of expectancies in music: whether they are all learned from one's own musical culture or whether there are some universal heuristics that transcend musical cultures. Addressing any notion of universality underlying these mechanisms would take one into the domain of evolutionary psychology (Barkow, Cosmides, & Tooby, 1992). This is obvious if we follow the rationale behind the two levels of expectancies presented previously. All expectancies are acquired through exposure to the surroundings, though different kinds of memory structures are employed. Data-driven expectancies employ various short-term memories (echoic memory, short-term or working-memory). If these data-driven memories are beneficial for an organism and if the events occur frequently enough, the knowledge they represent may be transferred into long-term memory and may be used on future occasions (schema-driven expectancies). If a certain context is particularly stable over an evolutionary time period, an organism may develop a *predisposition* towards learning the particular regularity, also called an *innate* mechanism. Hence, the assumption about the innateness of Narmour's model or simple quantification of pitch proximity relies on the supposition that it has been useful in predicting the behaviour of acoustic signals for tens of thousands of years. One principle, pitch proximity, may tap into such a regularity, as described in detail in Bregman's theory of auditory stream segregation (1990). According to this theory, successive pitches are more likely to be perceived as belonging to a same stream if they are close in pitch and temporally proximate (*cf.* van Noorden, 1975). The recent revisions regarding principles of Proximity and Registral Direction in Narmour's model by von Hippel (2000), forming the concept of Tessitura, may offer an even more ecologically valid formulation although these were not implicated in the present study. To put it another way: humans have structured the auditory scene by finding suitable rules of thumb for detecting sound sources reliably and for tracking the vocal utterances of fellow communicators. These heuristics state that successive pitches from a single communicator tend to be close in proximity, and that the pitch range also has a tendency to be fairly limited. Both assumptions can be observed in music (Huron, 2001; von Hippel, 2000), and are exploited in Narmour's (1990) and von Hippel's (2000) models of melodic expectancies — which of course does not mean that these parsing mechanisms were developed for predicting the continuations of musical events.

The present study used cross-cultural comparisons to estimate the contribution of data-driven and schema-driven processes in expectancy formation. Other ways of teasing out the culture-specific and culture-transcending qualities of musical processing occur in the field of infant studies. This area has already produced findings on several human predispositions for processing music and has illustrated how culture-specific, schematic processes in music develop (for a review, see Trehub & Trainor, 1993).

Predispositions have also been studied in emotion research related to music, in which similar acoustic cues have been found to work across languages and cultures (Scherer, Banse & Wallbott, 2001; also Gabrielsson & Juslin, 2003 and Balkwill & Thompson, 1999) despite the many culture-specific associations that are involved in music.

Further research on the mechanisms underlying the formation of expectations across cultures is essential in music cognition, despite the methodological and epistemological difficulties detailed earlier. One should, however, note the argument of many music anthropologists, musicologists and ethnomusicologists: that the essential features of musical culture as a phenomenon do not lie in the similarities between cultures but rather in the unique features of a culture that can only be studied from within that culture. As Davidson and Torff (1992, p. 136) have put it, "it is fruitless to attempt to isolate a cognitive core in the individual independent of cultural and local forces." Indeed, the cultural context is important, and in order to reduce ethnocentrism in the explanatory models, the research questions should also be derived from the unique properties of the musical cultures themselves. This would lead to more cultural sensitivity, especially if the questions and hypotheses could be drawn from the cultures studied, particularly those non-Western cultures that are currently used to examine the validity of Western-oriented models and hypotheses. Moreover, any conclusions about features that may transcend cultures will require a comparison of several diverse cultures (Berry, Poortinga, Segall, & Dasen, 1992).

However, a strong challenge to the rejection of cross-cultural approaches is offered by the psycholinguist Steven Pinker, who takes the view that "[u]niversal mental mechanisms can underlie superficial variation across cultures" (2002, p.37). Here he echoes the ideas presented earlier by Chomsky (1975, 2000) in the field of linguistics, and by Lomax (1968) and Harwood (1976) in the field of music. Indeed, linguistics is taking great strides in applying evolutionary psychology and neurobiology to explanations of how language is acquired and used (Hauser Barner, & O'Donnell, 2007). Although the search for common mental mechanisms underlying music may raise objections from ethnomusicologists and culturally oriented music analysts, the field of music cognition will undoubtedly benefit from taking a multidisciplinary path; one that has already been trodden by linguists and by other cognitive scientists.

Acknowledgements

We humbly thank Simon Sete, a traditional healer from Sekhukhune area, without whom we could not have carried out the study. We are also grateful to Dr Mogomme Masoga and Professor Caroline Niekerk for providing critical reading of the manuscript and suggesting numerous improvements to our discussion of the African music and culture.

Address for correspondence:
Tuomas Eerola
Finnish Centre of Excellence in Interdisciplinary Music Research
University of Jyväskylä, Finland
e-mail: tuomas.eerola@jyu.fi

Appendix

Table A1
Correlations Between All Predictors and Variance Inflation Factors (VIF)

	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	$\sqrt{VIF}$
1. Auditory model	.38***	.19*	.02	.14	.44***	.38***	.26**	.22*	.09	.33***	1.16
2. PC in contexts		.41***	.33***	.49***	.52***	.86***	.66***	.19	.03	.18	2.19
3. IV in contexts			.40***	.34***	.21*	.41***	.43***	.30**	.07	.12	1.18
4. Proximity				.55***	-.03	.38***	.35***	.51***	.02	.06	1.34
5. Pitch reversal					.12	.43***	.52***	.06	.10	-.01	1.40
6. PC in yoiks						.61***	.51***	.38**	.22*	.59***	1.63
7. 2-tones in yoiks							.80***	.32**	.09	.36***	2.83
8. 3-tones in yoiks								.21	.31**	.22*	2.03
9. Western schem.									.01	.30*	1.49
1. Correct tones										.22*	1.14
11. PC in Pedi											1.36

Note. * $p < .05$, ** $p < .01$, *** $p < .001$, $df = 108$, except for Western schem. $df = 84$.

• References

Aarden, B. (2002). Expectancy vs. retrospective perception: Reconsidering the effects of schema and continuation judgments on measures of melodic expectancy. In C. Stevens, D. Burnham, G. McPherson, E. Schubert & J. Renwick (eds), *Proceedings of the 7th international conference on music perception and cognition* (pp. 469-72). Adelaide: Causal Productions.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by human infants. *Psychological Science*, *9*, 321-24.

Auhagen, W. (1994). *Experimentelle untersuchungen zur auditiven tonalitätsbestimmung in melodien.* Regensburg: Gustav Bosse Verlag.

Balkwill, L. L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception*, *17*, 43-64.

Barkow, J. H., Cosmides, L., & Tooby, J. (1992). *The adapted mind: Evolutionary psychology and the generation of culture.* Oxford University Press.

Barz, G. F. (2003). Sotho/Tswana music. In L. Macy (ed), *The new Grove Dictionary of music online.* Retrieved from http://www.grovemusic.com

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (eds). (1992). *Cross-cultural psychology: Research and applications.* Cambridge, UK: Cambridge University Press.

Blacking, J. (1962). Musical expeditions of the Venda. *African Music*, *3*(1), 54-72.

Blacking, J. (1970). Tonal organization in the music of two Venda initiation schools. *Ethomusicology*, *14*, 1-56.

Bragg, B. W. E., & Crozier, J. B. (1974). The development with age of verbal exploratory responses to sound sequences varying in uncertainty level. In D. Berlyne (ed), *Studies in the new experimental aesthetics* (pp. 91-108). Washington, DC: Hemisphere Publishing Corporation.

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

Carlsen, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology*, *1*, 12-29.

Carmines, E.G., & Zeller, R. A. (1979). *Reliability and validity assessment.* Beverly Hills, CA: Sage Publication.

Carterette, E. C., & Kendall, R. A. (1999). Comparative music perception and cognition. In D. Deutsch (ed), *The psychology of music* (pp. 725-91). San Diego, CA: Academic Press.

Carterette, E. C., Kendall, R. A., & DeVale, S. C. (1993). Comparative acoustical and psychoacoustical analyses of gamelan instrument tones. *Journal of the Acoustical Society of Japan (E)*, *14*(6), 383-96.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General*, *113*, 394-412.

Chomsky, N. (1975). *Reflections on language.* New York: Pantheon.

Chomsky, N. (2000). *New horizons in the study of language and mind.* New York: Cambridge University Press.

Cohen, D., Granot, R., Pratt, H., & Barnea, A. (1993). Cognitive meanings of musical elements as disclosed by event-related potential (ERP) and verbal experiments. *Music Perception*, *11*(2), 153-84.

Coplan, D. (1998). Popular music in South Africa. In R. Stone (ed), *The Garland Encyclopedia of World Music. Volume 1: Africa* (pp. 759-80). New York: Garland Publishing.

Coplan, D. (2003). Popular styles and cultural fusion. In L. Macy (ed), *The new Grove Dictionary of music online*. Retrieved from http://www.grovemusic.com

Davidson, L., & Torff, B. (1992). Situated cognition in music. *World of Music*, *34*(3), 120-39.

Demorest, S. M., Morrison, S. J., Aylward, E. H., Cramer, S. C., & Maravilla, K. R. (2002). An fMRI study of cross-cultural music comprehension. In C. Stevens, D. Burnham, G. G. McPherson, E. E. Schubert & J. J. Renwick (eds), *Proceedings of the 7th international conference on music perception and cognition* (pp. 141). Adelaide, Australia: Causal Productions.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. New York: Academic Press.

Eerola, T., Toiviainen, P., & Krumhansl, C. L. (2002). Real-time prediction of melodies: Continuous predictability judgments and dynamic models. In C. Stevens, D. Burnham, G. G. McPherson, E. E. Schubert & J. J. Renwick (eds), *Proceedings of the 7th international conference on music perception and cognition* (pp. 473-76). Adelaide, Australia: Causal Productions.

Fox, J. (1991). *Regression diagnostics*. Beverly Hills, CA: sage Publications.

Gabrielsson, A., & Juslin, P. N. (2003). Emotional expression in music. In H. H. Goldsmith, R. J. Davidson & K. R. Scherer (eds), *Handbook of affective sciences* (pp. 503-34). New York: Oxford University Press.

Gregory, A. H., & Varney, N. (1996). Cross-cultural comparisons in the affective response to music. *Psychology of Music*, *24*, 47-52.

Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2006). *Multivariate Data Analysis*. Prentice-Hall: Englewood Cliffs, NJ.

Hauser, M., Barner, D., & O'Donnell, T. (2007). Evolutionary Linguistics: A New Look at an Old Landscape. *Language Learning and Development*, *3*(2), 101-32.

Harwood, D. L. (1976). Universals in music: a perspective from cognitive psychology. *Ethnomusicology*, *20*, 521-33.

Hewson, M. G. (1998). Traditional healers in southern Africa. *Annals of Internal Medicine*, *128*, 1029-34.

von Hippel, P. (2000). Redefining pitch proximity: Tessitura and mobility as constraints on melodic interval size. *Music Perception*, *17*(3), 315-27.

von Hippel, P., & Huron, D. (2000). Why do skips precede reversals? The effect of tessitura on melodic structure. *Music Perception*, *18*(1), 59-85.

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, *19*(1), 1-64.

Huron, D. (2006). *Sweet anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press.

Huron, D., & Parncutt, R. (1993). An improved model of tonality perception incorporating pitch salience and echoic memory. *Psychomusicology*, *12*, 152-69.

Huskisson, Y. (1958). *The social and ceremonial music of the Pedi*. Unpublished doctoral dissertation, University of Witwatersrand, Johannesburg.

Janzen, J. M. (2000). Theories of music in African ngoma healing. In P. Gouk (ed), *Musical Healing in Cultural Contexts* (pp. 46-66). UK: Ashgate.

Kaemmer, J. E. (1998). Southern Africa: An introduction. In R. M. Stone (ed), *The Garland Encyclopedia of World Music. Volume 1: Africa* (pp. 700-21). New York: Garland Publishing.

Kantola, T. (1984). Talvadaksen joikuperinne (The yoik tradition of Talvadas). In *Folkloristiikan tutkimuksia (Studies in folklore)*. Turku: University of Turku.

Katzner, K. (1986). *Languages of the world* (Rev. ed.). London: Routledge.

Kendall, R. A., & Carterette, E. C. (1991). Perceptual scaling of simultaneous wind instrument timbres. *Music Perception*, *8*(4), 369-404.

Kessler, E. J., Hansen, C., & Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and the West. *Music Perception*, *2*, 131-65.

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, *15*(2), 207-12.

Koelsch, S. & Friederici, A. (2003). Comparative Results of Different Neurophysiological Investigation Methods. *Annals of the New York Academy of Sciences*, *999*, 15-28.

Krumhansl, C. L. (1995a). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum*, *17*, 53-90.

Krumhansl, C. L. (1995b). Effects of musical context on similarity and expectancy. *Systematische musikwissenschaft*, *3*, 211-50.

Krumhansl, C. L. (2000). Tonality induction: A statistical approach applied cross-culturally. *Music Perception*, *17*(4), 461-79.

Krumhansl, C. L., Louhivuori, J., Toiviainen, P., Järvinen, T., & Eerola, T. (1999). Melodic expectation in Finnish Spiritual Folk Hymns: Convergence of statistical, behavioral, and computational approaches. *Music Perception*, *17*, 151-96.

Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition*, *76*, 13-58.

Krumhansl, C. L., & Toiviainen, P. (2000). Dynamics of tonality induction: a new method and a new model. In C. Woods, G. B. Luck, R. Brochard, S. A. O'Neill & J. A. Sloboda (eds), *Proceedings of the Sixth International Conference on Music Perception and Cognition* (pp. 1504-13). Keele, Staffordshire, UK: Keele University.

Larson, S. (1997). Continuations as completions: Studying melodic expectation in the creative microdomain Seek Well. In M. Leman (ed), *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology* (pp. 321-34). Berlin: Springer Verlag.

Lebaka, M. E. K. (2001). *The ritual use of music in indigenous African religion: a Pedi perspective.* Unpublished master's thesis, University of Pretoria, Pretoria, South Africa.

Leman, M. (2000). An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception*, *17*(4), 481-509.

Leman, M., Lesaffre, M., & Tanghe, K. (2000). *The IPEM toolbox manual*. Univ. of Ghent, IPEM-Dept. of Musicology: IPEM.

Lomax, A. (1968). *Folk song style and culture*. Washington D. C.: American Association for the Advancement of Science.

Louhivuori, J., Salminen, V. & Lebaka, E. (2005). "Singing Together" — A Cross-cultural Approach to the Meaning of Choirs as a Community. In P. Campbell, J. Drummond,

P. Dunbar-Hall, K. Howard, H. Schippers & T. Wiggins (eds), *Cultural Diversity in Music Education Brisbane*: Australian Academic Press.

Masoga M.A. (2006). *Maila-go-fenywa, Rangwato Magoro and Mmino wa Koša: Some perspectives on theory and practice*. Unpublished master's dissertation, University of South Africa, Pretoria, South-Africa.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30-46.

Mesthrie, R. (2006). Language, transformation and development: a sociolinguistic appraisal of post-apartheid South African language policy and practice. *Southern African Linguistics and Applied Language Studies*, *24*(2), 151-63.

Meyer, R. K., Palmer, C., & Mazo, M. (1998). Affective and coherence responses to Russian laments. *Music Perception*, *16*(1), 135-50.

Milberg, W., Blumstein, S., Katz, D., Gershberg, F., & Brown, T. (1995). Semantic facilitation in aphasia: effects of time and expectancy. *Journal of Cognitive Neuroscience*, *7*, 33-50.

Morrow-Howell, N. (1994). The m word: Multicollinearity in multiple regression. *Social Work Research*, *18*(4), 247-51.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. University of Chicago Press.

Nunnaly, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

Oram, N., & Cuddy, L. L. (1995). Responsiveness of Western adults to pitch-distributional information in melodic sequences. *Psychological Research*, *57*, 103-18.

Pinker, S. (2002). *The blank slate. The modern denial of human nature*. Great Britain: Penguin.

Poortinga, Y. H. (1972). A comparison of African and European students in simple auditory and visual tasks. In L. J. Cronbach & P. J. Drenth (eds), *Mental tests and cultural adaptation* (pp. 349-54). The Hague: Mouton.

Roh, S. K., & Yi, S. W. (2000). The effects of musical styles and experience on melodic expectancy. In C. Woods, G. B. Luck, R. Brochard, S. A. O'Neill & J. A. Sloboda (eds), *Proceedings of the Sixth International Conference on Music Perception and Cognition* (pp. 1635-43). Keele, Staffordshire, UK: Keele University.

Rycroft, D. (1983). The relationships between speech-tone and melody in Southern African Music. In J. P. Malan (ed), *South African Music Encyclopedia* (pp. 301-14). Cape Town: Oxford University Press.

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

Schellenberg, *E.G.* (1996). Expectancy in melody: Tests of the implication-realization model. *Cognition*, *58*, 75-125.

Schellenberg, *E.G.* (1997). Simplifying the implication-realization model of melodic expectancy. *Music Perception*, *14*, 295-318.

Schellenberg, E., Adachi, M., Purdy, K., & McKinnon, M. (2002). Expectancy in Melody: Tests of Children and Adults. *Journal of Experimental Psychology General*, *31*(4), 511-37.

Scherer, K. R., Banse, R., & G., Walbott, H. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, *32*(1), 76-92.

Schlosser, M. J., Aoyagi, N., Fulbright, R. K., Gore, J. C., & McCarthy, G. (1998). Functional MRI studies of auditory comprehension. *Human Brain Mapping*, *6*(1), 1-13.

Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception*, *7*, 109-50.

SOE (2008). *State of the Environment Maps.* Department of Environmental Affairs and Tourism, Republic of South Africa. Retrieved October 15, 2008, from http://soer.deat.gov.za/themes.aspx?m=121

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*, 245-51.

Thompson, W. F., & Stainton, M. (1998). Expectancy in Bohemian folk song melodies: Evaluation of implicative principles for implicative and closural intervals. *Music Perception*, *15*(3), 231-52.

Toiviainen, P., & Eerola, T. (2003). Where is the Beat?: Comparison of Finnish and South African Listeners. In *Proceedings of the ESCOM 2003 Conference, Hanover, Germany*: ESCOM.

Toiviainen, P., & Krumhansl, C. L. (2003). Measuring and modeling real-time responses to music: the dynamics of tonality induction. *Perception*, *32*(6), 741-66.

Trehub, S. E., & Trainor, L. J. (1993). Listening strategies in infancy: The roots of music and language development. In S. McAdams & E. Bigand (eds), *Thinking in sound: The cognitive psychology of human audition* (pp. 278-327). Oxford University Press.

Unyk, A. M., & Carlsen, J. C. (1987). The influence of expectancy on melodic perception. *Psychomusicology*, *7*, 3-23.

Van Immerseel, L., & Martens, J. P. (1992). Pitch and voiced/unvoiced determination with an auditory model. *Journal of Acoustical Society of America*, *91*, 3511-26.

Van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences.* Unpublished doctoral dissertation, Eindhoven University of Technology, Netherlands.

• Expectativas en los Sami Yoiks revisitados: el papel de los datos y esquemas conductores de conocimiento en la formación de expectativas melódicas

Este estudio amplía un estudio anterior concerniente a las expectativas melódicas en los Sami Yoiks del norte (Krumhansl *et al*., 2000) en el cual una comparación entre oyentes expertos e inexpertos demostró la existencia de un juego nuclear de principios que gobiernan las expectativas melódicas. Las conclusiones anteriores son reconsideradas utilizando oyentes no occidentales (curanderos tradicionales de Sudáfrica) en una investigación modelada. La comparación de modelos diferentes hizo posible separar el papel de los modelos conducidos por datos y conducidos por esquemas en las expectativas melódicas y revelar cualquier posible prejuicio occidental en los estudios anteriores. Los resultados del experimento, en el que los oyentes africanos valoraron el valor de tonos de sonda como continuaciones de los extractos de los Sami Yoik del norte, indicaron que los modelos conducidos por datos son adecuados para explicar las expectativas, independientemente del entorno cultural de los oyentes. Los modelos basados en las frecuencia ejercieron más influencia sobre oyentes no familiarizados con el estilo Yoik; el modelo esquemático occidental tenía mayor impacto sobre oyentes occidentales, y los modelos estilísticos específicos generaron una mayor capacidad de predicción en aquellos oyentes que tenían conocimiento de los Yoiks.

• Previsione negli yoik sami rivisitati: il ruolo della conoscenza *data-driven* e *schema-driven* nella formazione delle aspettative melodiche

Questa ricerca approfondisce uno studio precedente sulle previsioni melodiche negli yoiks dei sami del Nord (Krumhansl *et al*., 2000) in cui attraverso il confronto tra ascoltatori esperti e non esperti è stata dimostrata l'esistenza di un nucleo centrale di principi che governano le previsioni melodiche. Le precedenti scoperte sono state riprese utilizzando ascoltatori non occidentali (guaritori tradizionali sudafricani) nell'ambito di un'investigazione di modelli. Il confronto tra modelli distinti ha reso possibile l'isolamento del ruolo dei modelli *data-driven* e *schema-driven* nelle aspettative melodiche e la rilevazione di ogni possibile distorsione occidentale negli studi precedenti. I risultati dell'esperimento in cui gli ascoltatori africani valutavano l'adeguatezza di alcuni *probe-tones* come continuazioni di estratti di yoik dei sami del Nord indicano che i modelli *data-driven* sono adeguati nella spiegazione delle aspettative al di là del background culturale degli ascoltatori. I modelli basati sulla frequenza esercitavano un'influenza maggiore sugli ascoltatori che non avevano familiarità con lo stile degli yoik; il modello schematico occidentale aveva un impatto maggiore sugli ascoltatori occidentali e i modelli stile-specifici esprimevano una capacità predittiva maggiore negli ascoltatori che conoscevano gli yoik.

• L'attente dans les Yoiks des Sami revisitée : le rôle de la connaissance fondée sur les faits et fondée sur les schémas dans la formation des attentes mélodiques

Cette étude complète une étude précédente sur les attentes mélodiques dans les yoiks chez les Sami du Nord (Krumhansl *et al.*, 2000), dans laquelle une comparaison entre des auditeurs experts et d'autres non experts démontrait l'existence d'un ensemble central de principes gouvernant les attentes mélodiques. Les résultats précédents sont reconsidérés, avec des auditeurs non occidentaux (des guérisseurs traditionnels d'Afrique du Sud), dans une enquête visant à la modélisation. La comparaison de différents modèles a permis de distinguer le rôle des modèles fondés sur les faits et fondés sur les schémas dans les attentes mélodiques. Elle a révélé un possible biais occidental dans les précédentes études. Les résultats de l'expérience, dans laquelle des auditeurs africains évaluaient la pertinence de sons-sondes pour continuer des extraits de yoiks des Sami du Nord, indiquaient que les modèles fondés sur les faits sont adéquats pour expliquer les attentes, quel que soit l'arrière-plan culturel des auditeurs. Les modèles basés sur la fréquence exerçaient plus d'influence sur les auditeurs non familiers du style yoik, le modèle schématique occidental avait plus d'impact sur les auditeurs occidentaux, et les modèles de style spécifique avaient plus de pouvoir prédictif pour ceux des auditeurs qui connaissaient déjà les yoiks.

• Eine Replikationsstudie zu Erwartungen in Sami Yoiks: Die Rolle von datenerzeugtem und schemaerzeugtem Wissen in der Formung melodischer Erwartungen

Diese Studie erweitert eine bestehende Studie zu melodischen Erwartungen bei Nord-Sami Yoiks aus (Krumhansl *et al.*, 2000), in denen Hörvergleiche zwischen Experten und Nichtexperten einige Kernregeln zu melodischen Erwartungen ergeben hatten. Die bisherigen Befunde werden in einer modellbildenden Studie an nicht-westlichen Hörern (traditionellen Heilern aus Südafrika) neu untersucht. Der Vergleich verschiedener Modelle ermöglichte die Herausarbeitung der Rolle von datenerzeugten und schemaerzeugten Modellen melodischer Erwartungen und konnte weiterhin mögliche westliche Vorurteile in bestehenden Untersuchungen aufzeigen. Im Experiment bewerteten afrikanische Hörer die Angemessenheit von Testtönen als Fortführung von Hörbeispielen nordsamischer Yoiks. Die Ergebnisse zeigen, dass datenerzeugte Modelle die Erwartungen unabhängig vom kulturellen Hintergrund der Hörer adäquat erklären können, Die frequenzbasierten Modelle übten mehr Einfluss auf solche Hörer aus, die nicht mit dem Stil der Yoiks vertraut waren. Während das westliche Schemamodell den stärksten Einfluss auf westliche Hörer ausübte, brachten die stilspezifischen Modelle die stärkste Vorhersageleistung für diejenigen Hörer, die Wissen über die Yoiks besaßen.