9th International Conference on Music Perception and Cognition

Alma Mater Studiorum University of Bologna, August 22-26 2006

The information dynamics of melodic boundary detection

Marcus T. Pearce **Geraint A. Wiggins**

Centre for Cognition, Computation and Culture,
Goldsmiths College, University of London,
New Cross, London SE14 6NB, United Kingdom
{m.pearce,g.wiggins}@gold.ac.uk

ABSTRACT

Many published models of perceived grouping structure in music are inspired by Gestalt psychology, associating grouping boundaries with discontinuities or changes in various dimensions of the musical surface. We examine a complementary approach based on information dynamics in melody perception according to which boundaries are perceived at points of expectancy violation and predictive uncertainty. We discuss empirical evidence for and against the two theories, consider how they relate to one another and suggest methods for further empirical investigation and development of the information dynamics approach.

Keywords

Melody perception, grouping structure, segmentation, boundary perception, Gestalt rules, statistical learning, information dynamics, expectation.

INTRODUCTION

The perception of grouping structure in music involves the identification of boundaries between contiguous groups of musical material. Just as speech is perceptually segmented into words which subsequently provide the building blocks for the perception of phrases and complete utterances (Brent, 1999b), motifs or phrases in a melody are identified by listeners, stored in memory and made available for inclusion in higher-level structural groups (Lerdahl & Jackendoff, 1983; Peretz, 1989; Tan, Aiello, & Bever, 1981). The low-level organisation of the musical surface into groups allows the use of these primitive perceptual units in more complex structural processing and may alleviate processing and memory demands.

Grouping structure is generally agreed to be logically independent of metrical structure (Lerdahl & Jackendoff, 1983) and some evidence for a separation between the psychological processing of the two kinds of structure has been found in cognitive neuropsychological (Liegeoise-Chauvel, Peretz, Babai, Laguitton, & Chauvel, 1998; Peretz, 1990) and neuroimaging research (Brochard, Dufour, Drake, & Scheiber, 2000). In practice, however, metrical and grouping structure are often intimately related and both are likely to serve as inputs to the processing of more complex musical structures (Lerdahl & Jackendoff, 1983).

AIMS

While the majority of existing models of perceived grouping structure in music are inspired by Gestalt psychology, we propose a complementary theory based on expectancy violation and predictive uncertainty. We examine how it may operate in music perception, as well as its implications for and relationship with Gestalt-based approaches. Subject to further empirical corroboration, the theory offers the possibility of relating two areas of research on music perception: expectation and grouping.

BACKGROUND

The Gestalt Approach

Theoretical Perspective

The perception of melodic groups has traditionally been modelled through the identification of local discontinuities or changes between events in terms of temporal proximity, pitch, duration and dynamics (Cambouropoulos, 2001; Lerdahl & Jackendoff, 1983; Temperley, 2001; Tenney & Polansky, 1980). Perhaps the best known examples are the Grouping Preference Rules (GPRs) of the Generative Theory of Tonal Music (GTTM) which constrain the segmentation of a musical surface into a hierarchically organised structure of recursively embedded groups through the identification of perceived local segment boundaries (Lerdahl & Jackendoff, 1983).

The most widely studied of these GPRs predict that phrase boundaries will be perceived between two melodic events whose temporal proximity is less than that of the immediately neighbouring events (due to a slur, a rest or a relatively long inter-onset interval or IOI) or when the transition between two events involves a greater change in register, dynamics, articulation or duration than the immediately neighbouring transitions.

Empirical Perspective

Several empirical studies have examined the extent to which the GPRs of GTTM accurately predict listeners' perceived grouping boundaries in melodic music. Two experimental paradigms are commonly used to obtain the behavioural data: first, asking participants to explicitly indicate perceived boundary locations while listening to a melody (Deliège, 1987; Frankland & Cohen, 2004; Peretz, 1989); and second, a probe recognition paradigm in which recognition memory judgements are obtained for melodic fragments (the probes) which cross or adjoin predicted phrase boundaries in a previously heard melody (Dowling, 1973; Frankland & Cohen, 2004; Peretz, 1989; Tan et al., 1981). In general, the results provide strong support for those GPRs related to temporal proximity but more equivocal support for those related to parametric change (with the possible exception of dynamics and timbre, Deliège, 1987).

The Expectancy Approach

Theoretical Perspective

Narmour (1990) proposes a different model according to which perceptual groups are associated with points of closure where the ongoing cognitive process of expectation is disrupted. Meyer (1957) discusses three ways in which expectations may be disrupted: first, an event expected to occur in a given context is delayed; second, the context fails to stimulate strong expectations for any particular continuation; and third, the continuation is unexpected.

Building on these approaches, we propose that boundaries are perceived before unexpected events or following points of predictive uncertainty and we quantify these two metrics in information-theoretic terms by reference to a model of unsupervised inductive learning of melodic structure. Briefly, the models we propose are n-gram models in which the conditional probability of a melodic event e given the context c of the preceding n-1 events is estimated on the basis of the frequency with which that symbol occurred in the same context in the prior experience of the model. The simplest way of estimating probabilities in this way is the *maximum likelihood* method:

$$p(e|c) = \frac{count(ce)}{count(c)}$$

where ce denotes the concatenation of c and e. Given a trained n-gram model, the degree to which an event appearing in a given context in a melody is unexpected can be defined as the *information content*, $h(e|c)$, of the event given the context:

$$h(e|c) = log_2 \frac{1}{p(e|c)}$$

Given an alphabet E of events which have appeared in the prior experience of the model, the uncertainty of the model's expectations in a given melodic context can be defined as the *entropy* or average information content of the events in E:

$$H(c) = \sum_{e \in E} p(e|c)\, h(e|c)$$

Extending the proposals of Meyer (1957), we suggest two related hypotheses: first, we would expect boundaries to be perceived before unexpected events (i.e., when p is low and h is high); and second, we would expect boundaries to be perceived after contexts associated with high predictive uncertainty (i.e., when H is low). In both cases, boundaries are predicted to occur when the context fails to inform the listener about forthcoming events leading to cognitive representations of a melody that maximise likelihood and simplicity (cf. Chater, 1996, 1999). The definitions of high and low in these contexts must be quantified, perhaps in relation to the values of p, h and H for the previous event or averaged over a window of previous events.

The two hypotheses are clearly related to one another to the extent that p is used in computing H although the manner in which they are related will depend on identities of c and e as well as the statistical structure of the training data. Evidently, we would like to know which statistics are computed by listeners and what influence they have on the perception of grouping boundaries. We regard these as empirical questions to be addressed by behavioural experiments with human listeners.

To our knowledge the second hypothesis has not been studied in experimental psychological research. Regarding the first hypothesis, however, it has been demonstrated that infants and adults reliably identify grouping boundaries in sequences of synthetic syllables (Saffran, Aslin, & Newport, 1996) and isochronous tone sequences (Saffran, Johnson, Aslin, & Newport, 1999) on the basis of higher digram (n=2) transition probabilities within than between groups. Brent (1999a) formalises this segmentation strategy in two models: the first is based on digram transition probabilities; the second is based on *pointwise mutual information*, $I(x,y)$, which measures how much the occurrence of one event reduces the model's uncertainty about the occurrence of another event (Manning & Schütze, 1999) and is defined as:

ISBN 88-7395-155-4

$$I(x,y) = log_2 \frac{p(xy)}{p(x)\,p(y)}$$

While digram probabilities are asymmetric with respect to the order of the two events, pointwise mutual information is a symmetric measure in this regard.[1] The models proposed by Brent predict a boundary before an event in a sequence when the statistic associated with that event (either digram probability or pointwise mutual information) is lower than that associated with its two immediate neighbours.

Brent (1999a) found that the pointwise mutual information model outperformed the transition probability model in predicting word boundaries in phonemic transcripts of infant-directed speech. Similar strategies for identifying word boundaries have been implemented using recurrent neural networks (e.g., Elman, 1990). These and other related approaches to segmentation and word discovery in natural language are reviewed by Brent (1999a, 1999b).

Empirical Perspective

Extending the work of Narmour (1990), Pearce and Wiggins (2004, 2006) have demonstrated that expectation in melodic pitch structure can be accurately modelled as a process of prediction based on the statistical induction of regularities in various dimensions of the melodic surface.

Furthermore, recent empirical research on implicit learning (Cleeremans, Destrebecqz, & Boyer, 1998; Seger, 1994) has begun to examine the manner in which adults and infants use induced statistical regularities to identify grouping boundaries in tone sequences (Creel, Newport, & Aslin, 2004; Saffran, 2003; Saffran & Griepentrog, 2001; Saffran et al., 1999; Saffran, Reeck, Niebuhr, & Wilson, 2005; Tillmann & McAdams, 2004). The experimental paradigm used typically involves training participants on isochronous tone sequences composed of three-tone groups where the only consistent cue to grouping boundaries is that digram transition probabilities are lower between than within groups. Since the goal is to examine learning, the tones themselves and the groups are carefully chosen to avoid creating familiar tonal contexts. In the test phase, participants undertake a two-alternative forced-choice task in which they select the most familiar of a three-tone group and a non-group (or part-group). Above chance performance is taken to indicate that the participants have induced the first-order statistical properties of the training sequences and used these in identifying grouping boundaries.

Research using this experimental approach has yielded evidence that infants and adults use the implicitly learnt statistical properties of pitch (Saffran et al., 1999) and pitch interval (Saffran, 2003; Saffran & Griepentrog, 2001; Saffran et al., 2005) sequences to identify segment boundaries before unexpected events. Tillmann and McAdams (2004) used the paradigm to investigate statistical learning of sequences of sounds which were constant in pitch but differed in timbre. Three conditions were created in which timbral similarity either supported, contradicted or was neutral with respect to the grouping of sounds on the basis of transition probabilities. The influence of training was assessed by comparing the responses of the trained participants with those of a control group with no prior training. The results indicated a significant influence of induced transition probabilities but also a general bias to segment on the basis of timbral similarity which did not interact with the inductive learning.

Finally, Creel et al. (2004) set out to examine whether the ability to learn statistical dependencies in tone sequences extends to dependencies between non-adjacent events. The results indicated that this was only the case when non-adjacent groups were distinguished in term of register or timbre. In the latter case, there appeared to be an interaction between timbral similarity and learning although, given the results reported by Tillmann and McAdams (2004) and summarised above, this result should be replicated using a control group with no learning experience before any strong conclusions are drawn.

DISCUSSION

Statistical Segmentation in Music

The theories of statistical and Gestalt approaches to segmentation may be related in various ways. It is relevant to discuss some of these here. First, one may interpret Gestalt models as predicting degrees of perceived salience or accentuation at a low-level, leading to representations of the melodic surface over which statistical models may operate. A second possibility is that the two models are both valid (making the same experimentally corroborated predictions) but operate at different levels of explanation (Pearce & Wiggins, 2006). Finally, it may be argued that the two models are directly comparable by examining melodic stimuli where they make conflicting predictions regarding the perception of grouping boundaries (e.g., on small but rare melodic intervals). In the remainder of this paper, we shall discuss potential areas of conflict between the two models with a view to empirical corroboration of one or the other.

Subject to further empirical corroboration, the theory of statistical segmentation of melody offers the possibility of relating two areas of research on music perception: expectation and grouping. In spite of the artificial nature of the materials used in the implicit learning experiments reviewed above (on which grounds it might be argued that they tell us little about the perception of actual music), it would be surprising if the empirically demonstrated ability to use statistical properties of tone sequences for segmentation did not play a significant role in music perception.

The purpose of this section is to discuss what this role may

[1] Manning and Schütze (1999) note that pointwise mutual information is biased in favour of low-frequency events inasmuch as, all other things being equal, I will be higher for digrams composed of low-frequency events than for those composed of high-frequency events. In statistical language modelling, pointwise mutual information is sometimes redefined as $count(xy)I(x, y)$ to compensate for this bias.

ISBN 88-7395-155-4

be and how it may be demonstrated experimentally. Since research to date has focused on the implicit learning of pitch and pitch interval structure, we assume that these dimensions of the musical surface (and related ones such as contour) are the most likely candidates for finding a relationship between grouping structure and expectations based on statistical learning (see also Narmour, 1990). In the following sections, we focus on the question of whether such learning mediates the influence of rhythmic, metrical and tonal structure on melodic boundary detection and, if not, how these influences may interact with the influence of expectations based on statistical learning.

Rhythmic Structure

There is a wealth of evidence that the perception of grouping boundaries is significantly influenced by temporal structure in the form of:

- the presence of rests or pauses (Deliège, 1987; Deutsch, 1980; Frankland & Cohen, 2004);
- the presence of notes with relatively long duration or IOI (Deliège, 1987; Dowling, 1973; Frankland & Cohen, 2004; Jusczyk & Krumhansl, 1993; Krumhansl & Jusczyk, 1990; Peretz, 1989).

Furthermore, these aspects of temporal structure are often found to account exclusively for the perception of segment boundaries with no need to postulate an additional influence of Gestalt principles for pitch proximity or similarity (Deliège, 1987; Frankland & Cohen, 2004; Peretz, 1989). However, this may depend on the recent musical experience of the listener, whether the listeners' attention is focused on pitch or temporal structure and the extent to which these two dimensions are coherent in suggesting boundaries at the same locations (Boltz, 1999). According to Narmour (1990), events convey a greater degree of closure if they are followed by a rest or preceded by a shorter event.

However, if we use the statistical approach to identify grouping boundaries on the basis of rhythmic information (e.g., IOIs preceding a melodic event), the model would tend to predict boundaries at transitions from long to short IOIs as well as from short to long IOIs. (Although future research should examine to what extent this behaviour is also exhibited by *n*-gram models where $n>2$). In practice, research has demonstrated that humans perceive boundaries only in the latter case (Deliège, 1987; Frankland & Cohen, 2004). It seems likely that relatively long intervals between the onsets of sounding events can be sufficient to disrupt the ongoing processing of musical structure and prompt the perception of a grouping boundary.

Where then are the putative effects of expectancy on boundary perception to be found? Following research on implicit learning of the statistical structure of tone sequences, we suggest that these effects may arise as a result of the processing of absolute pitch, pitch interval and tonal structure in a melodic stimulus. In this case, a statistical model will find large melodic intervals more unexpected than smaller ones since the former are much less frequent in most musical styles than the latter. In this case, Gestalt models similarly predict boundaries where large melodic intervals occur (Cambouropoulos, 2001; Lerdahl & Jackendoff, 1983; Tenney & Polansky, 1980).

However, empirical research has typically found, at best, weak influences of melodic interval size on boundary perception (Frankland & Cohen, 2004; Peretz, 1989). Since this may be a result of a concordance between pitch and temporal boundary structure in the materials used or simply the dominant influence of temporal proximity in boundary perception, future research should control for these possibilities by using distinct stimulus sets consisting of isochronous melodies, melodies that are coherent with respect to temporal and pitch structure and melodies that are incoherent in this regard.

Metrical Structure

There is also evidence for the influence of metrical structure inasmuch as perceived grouping boundaries tend to be aligned to strong metrical accent locations (Palmer & Krumhansl, 1987a, 1987b; Stoffer, 1985) and annotated phrase boundaries in folk music tend to conform to metrical parallelism (Temperley, 2001). According to Narmour (1990), events convey a greater degree of closure to the extent that they occur in stronger metrical location than the preceding event.

The influence of metrical structure and representation on the perception of grouping boundaries does not appear to be directly amenable to the statistical learning approach, at least not as it has been formulated here. Nonetheless, the influences of metrical structure must be controlled for in any experimental examination of the approach and ultimately, the manner in which induced regularities in pitch structure interact with perceived metrical structure in the perception of grouping structure must be elucidated.

Tonal-harmonic Structure

In Western tonal music, phrase endings are commonly associated with a move to more tonally stable notes in the prevailing key. According to Narmour (1990) a movement to a more tonally stable tone in a melody results in a greater sense of closure associated with that tone. This phenomenon has been studied in more depth by a number of researchers as discussed below.

Tan et al. (1981) conducted a study of harmonic influences on perceptual organisation in which listeners were presented with isochronous two-phrase melodies and asked to indicate whether a two-note sequence (the probe) had occurred in the melody. The participants included both musicians and non-musicians and the melodies varied according to whether the first phrase ended with a full cadence (a perfect cadence ending on the tonic) or a semicadence (an imperfect cadence ending on the dominant). The critical probes were taken from one of three locations in the melody: first, ending the first phrase; second, straddling the phrase boundary; and third, beginning the second phrase. As predicted by Tan et al., the results demonstrated that probes in the second position

ISBN 88-7395-155-4

were more difficult to recognise than those in other positions. This effect was found to be much stronger for the musicians than for the non-musicians. Furthermore, the results showed other effects of musical training. For non-musicians, the effect of probe position was no different for full cadences and semicadences. However, the musicians not only showed the strongest probe position effect in the full cadence condition but also exhibited strikingly better performance on last-phrase probes than on first-phrase probes in this condition (but not the semicadence condition).

In subsequent research, Boltz (1991) conducted a melody recall experiment in which musically trained participants were asked to listen to unfamiliar folk melodies and then recall them using music notation. The melodies varied in two dimensions: first, according to whether phrase endings (determined by metrical and intervallic structure) were marked by tonic triad members; and second, according to whether temporal accents coincided with the melodic phrase structure. The performance metric was the percentage of recalled notes forming the correct absolute interval with the preceding note. The results demonstrated that performance decreased when temporal accents conflicted with phrase boundaries and that the marking of phrase boundaries by tonic triad members resulted in significantly better performance but only when these boundaries were also marked by temporal accents.

In a subsequent analysis, Boltz (1991) classified the errors into three classes: those due to missing notes; those due to incorrect interval size but correct contour; and those due to incorrect interval size and incorrect contour. For melodies with coherent temporal accents, a significantly large proportion of the errors were contour-preserving but this was not the case for melodies with incoherent temporal accents. Finally, Boltz conducted an error analysis specifically at phrase endings which demonstrated that for coherent melodies with phrases marked by tonic triad members, a large proportion of the recalled notes were tonic triad members (correct or incorrect). For incoherent melodies with phrases marked by tonic triad members, however, the phrase-final notes were most frequently recalled as non-tonic triad members or missing notes.

Povel and Jansen (2002) report experimental evidence that goodness ratings of entire melodies depend not so much on the overall stability of the component tones (Krumhansl & Kessler, 1982) but the ease with which the listener is able to form a harmonic interpretation of the melody in terms of both the global harmonic context (key and mode) and the local movement of harmonic regions. The latter process is compromised by the presence of non-chord tones to the extent that they cannot be assimilated by means of anchoring (Bharucha, 1984) or by being grouped as part of a run of melodic steps. Povel and Jansen argue that the harmonic function of a region determines the stability of tones within that region and sets up expectations for the resolution of unstable tones.

Although it has not to our knowledge been empirically examined in these terms, it is intuitively plausible that the influence of implied harmonic movement on perceived melodic grouping is mediated by a process of implicit statistical learning. There is some evidence that the cognitive representation of tonal hierarchies depends on statistical learning (Krumhansl, 1990). Furthermore, phrase endings are often consistently marked by a highly constrained set of tonally stable scale degrees, which tend to be highly probable and expected, generating a relatively greater degree of uncertainty about the next tone (which will therefore be less expected when it does arrive).

Experimental Approaches

In future research, it will be important to clarify the relationship between information-theoretic influences resulting from statistical learning of regularities in pitch and pitch interval, and the influence of rhythmic structure (e.g., temporal proximity), metrical structure (e.g., alignment to strong beats and parallelism) and tonal-harmonic structure (e.g., closure) on the perception of grouping. Either these factors must be controlled for, such that they do not differ between stimuli, or they must be experimentally manipulated as independent variables. If such manipulations are deemed undesirable (perhaps in the interests of ecological validity), these potential influences on perceived grouping structure must at least be recorded and examined during analysis and modelling of the data.

The methodological question of choosing a paradigm for probing the points at which listeners perceive melodic segment boundaries is one that deserves attention. Asking participants to explicitly indicate boundary locations while listening to a melody is the most popular method but tends to yield effects of musical training (Deliège, 1987; Frankland & Cohen, 2004; Peretz, 1989) perhaps because "musicians appear to be more skilled than nonmusicians when examining their musical percept" (Peretz, 1989, p. 174). Results obtained using this method may suffer from the confounding effects of explicit musical training. The probe recognition paradigm (Dowling, 1973; Frankland & Cohen, 2004; Peretz, 1989; Tan et al., 1981) does not require the explicit examination of the musical percept but has yielded mixed findings regarding the effect of musical training. Furthermore, while some research has found convergence between the two paradigms (Frankland & Cohen, 2004) other research has failed to do so (Peretz, 1989).

The implicit learning paradigm used by Saffran et al. (1999) and others has the advantage of specifically controlling the relevant experience of the listener and, like the probe recognition method, does not depend on an explicit examination of the musical percept. These studies generally use isochronous non-tonal stimuli and in future development of this research it will be important to examine other dimensions of the melodic surface. For example, it would be interesting to know whether statistical regularities in event durations are learnt in the same manner or whether temporal proximity governs perceived grouping structure. In addition, it will be important to develop stimuli that explicitly distinguish between a grouping strategy based on pitch proximity and one based on statistical learning. We would also like to know how pitch and temporal structure (both statistical and Gestalt components) interact in affecting perceived grouping boundaries (Boltz, 1999; Palmer & Krumhansl, 1987a, 1987b). It would also be appropriate to examine how the influence of statistical learning is affected by manipulating metrical and tonal-harmonic structure. Indeed it is plausible that statistical learning of sequential relative pitch structure (i.e., scale degree) would influence perceived grouping boundaries.

In addition to examining the relationship between statistical learning and the perception of different dimensions of the melodic surface, future research should examine the limits of higher order statistical learning of tone sequences. Furthermore, the results obtained using the implicit learning paradigm are generally compatible with the claim that listeners compute a number of closely related statistics including transition probability, mutual information and conditional entropy (Creel et al., 2004). In future research, it will be important to attempt to distinguish between the use of these different statistics, as well as the hypothesis that boundaries are perceived at points of high uncertainty (low entropy), and identify which of these models best characterises the behaviour of listeners. Finally, in all this proposed research, the importance of including control groups with no training should not be forgotten (as emphasised by Tillmann & McAdams, 2004).

As discussed above, it is difficult to gauge the implications of the research on implicit statistical learning of tone sequences for research investigating the cognitive processing of music drawn from existing repertoires. Given this, it may prove fruitful to examine a relationship between perceived grouping structure and expectancy in a more exploratory fashion. In particular, it would be possible to identify perceived grouping boundaries in a set of melodic stimuli using the probe recognition paradigm and then experimentally examine expectations at the resulting boundary locations. The theory presented here predicts that listeners would exhibit a degree of predictive uncertainty at these locations or would experience a relatively great degree of expectancy violation at these boundary locations.

It is worth considering which experimental paradigm should be used to examine expectations in such a study. The continuation-tone rating paradigm typically used to examine melodic expectations (e.g., Schellenberg, 1996) should be avoided since it may elicit schematic tonal expectations specifically related to melodic closure since the melody is paused to allow the listener to respond (Aarden, 2003). Furthermore, the continuation generation paradigm (e.g., Thompson, Cuddy, & Plaus, 1997) is also unsuitable since the continuations generated are beyond the control of the experimenter complicating the assessment of uncertainty. This also holds true for the betting paradigm used by Manzara, Witten, and James (1992) which only yields results pertaining to the expectedness of the correct tones in a melody as does the study of reaction times in retrospective contour judgements (Aarden, 2003). Meanwhile, paradigms that involve the indication of some perceived quality such as expectedness (e.g., Eerola, Toiviainen, & Krumhansl, 2002) while listening to a melody do not provide sufficiently fine granularity in time or in terms of the relevant dimensions of the musical surface. Finally, all these methods differ in the extent to which they require the listener to conduct an explicit analysis which, given the present focus on implicit statistical learning, it is desirable to minimise. It will be important to develop a paradigm that addresses these shortcomings of existing experimental methods in the context of the present theoretical concerns.

We are currently engaged in the design of experiments that will allow the empirical examination of some of these issues in our continuing research in this area.

ACKNOWLEDGEMENTS

We would like to thank Daniel Müllensiefen for useful comments on an earlier draft of this paper. The research reported here was funded by EPSRC grant GR/S82220/01.

REFERENCES

Aarden, B. (2003). *Dynamic Melodic Expectancy*. Doctoral dissertation, Ohio State University, Columbus, OH.

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, *16*, 485–518.

Boltz, M. G. (1991). Some structural determinants of melody recall. *Memory and Cognition*, *19*(3), 239–251.

Boltz, M. G. (1999). The processing of melodic and temporal information: Independent or unified dimensions? *Journal of New Music Research*, *28*(1), 67–79.

Brent, M. R. (1999a). An efficient, probabilistically sound algorithm for segmentation and word discovery. *Machine Learning*, *34*(1-3), 71–105.

Brent, M. R. (1999b). Speech segmentation and word discovery: A computational perspective. *Trends in Cognitive Science*, *3*, 294–301.

Brochard, R., Dufour, A., Drake, C., & Scheiber, C. (2000). Functional brain imaging of rhythm perception. In C. Woods, G. Luck, R. Brochard, F. Seddon, & J. A. Sloboda (Eds.), *Proceedings of the Sixth International Conference of Music Perception and Cognition*. Keele, UK: University of Keele.

Cambouropoulos, E. (2001). The local boundary detection model (LBDM) and its application in the study of expressive timing. In *Proceedings of the International Computer Music Conference* (pp. 17–22). San Francisco: ICMA.

Chater, N. (1996). Reconciling simplicity and likelihood principles in perceptual organisation. *Psychological Review*, *103*(3), 566–581.

Chater, N. (1999). The search for simplicity: A fundamental cognitive principle? *The Quarterly Journal of Experimental Psychology*, *52A*(2), 273–302.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998).

Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*(10), 406–416.

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

Deliège, I. (1987). Grouping conditions in listening to music: An approach to Lerdahl and Jackendoff's grouping preference rules. *Music Perception*, *4*(4), 325–360.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception and Psychophysics*, *28*(5), 381–389.

Dowling, W. J. (1973). Rhythmic groups and subjective chunks in memory for melodies. *Perception and Psychophysics*, *14*(1), 37–40.

Eerola, T., Toiviainen, P., & Krumhansl, C. L. (2002). Real-time prediction of melodies: Continuous predictability judgements and dynamic models. In C. Stevens, D. Burnham, E. Schubert, & J. Renwick (Eds.), *Proceedings of the Seventh International Conference on Music Perception and Cognition* (pp. 473–476). Adelaide, Australia: Causal Productions.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Frankland, B. W. & Cohen, A. J. (2004). Parsing of melody: Quantification and testing of the local grouping rules of Lerdahl and Jackendoff's *A Generative Theory of Tonal Music*. *Music Perception*, *21*(4), 499–543.

Jusczyk, P. W. & Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infant's sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(3), 627–640.

Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. Oxford: Oxford University Press.

Krumhansl, C. L. & Jusczyk, P. W. (1990). Infant's perception of phrase structure in music. *Psychological Science*, *1*(1), 70–73.

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

Lerdahl, F. & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Liegeoise-Chauvel, C., Peretz, I., Babai, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain*, *121*(10), 1853–1867.

Manning, C. D. & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.

Manzara, L. C., Witten, I. H., & James, M. (1992). On the entropy of music: An experiment with Bach chorale melodies. *Leonardo*, *2*(1), 81–88.

Meyer, L. B. (1957). Meaning in music and information theory. *Journal of Aesthetics and Art Criticism*, *15*(4), 412–424.

Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures: The Implication-realisation Model*. Chicago: University of Chicago Press.

Palmer, C. & Krumhansl, C. L. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(1), 116–126.

Palmer, C. & Krumhansl, C. L. (1987b). Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing and familiarity. *Perception and Psychophysics*, *41*(6), 505–518.

Pearce, M. T. & Wiggins, G. A. (2004). Rethinking Gestalt influences on melodic expectancy. In S. D. Lipscomb, R. Ashley, R. O. Gjerdingen, & P. Webster (Eds.), *Proceedings of the Eighth International Conference of Music Perception and Cognition* (pp. 367–371). Adelaide, Australia: Causal Productions.

Pearce, M. T. & Wiggins, G. A. (2006). Expectancy in melody: The influence of context and learning. *Music Perception*, *23*(5), 377–405.

Peretz, I. (1989). Clustering in music: An appraisal of task factors. *International Journal of Psychology*, *24*(2), 157–178.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain*, *113*(4), 1185–1205.

Povel, D. J. & Jansen, E. (2002). Harmonic factors in the perception of tonal melodies. *Music Perception*, *20*(1), 51–85.

Saffran, J. R. (2003). Absolute pitch in infancy and adulthood: The role of tonal structure. *Developmental Science*, *6*(1), 37–49.

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

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

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

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

Schellenberg, E. G. (1996). Expectancy in melody: Tests of the implication-realisation model. *Cognition*, *58*(1), 75–125.

Seger, C. A. (1994). Implicit learning. *Psychological Bulletin*, *115*(2), 163–196.

Stoffer, T. H. (1985). Representation of phrase structure in the perception of music. *Music Perception*, *3*(2), 191–220.

Tan, N., Aiello, R., & Bever, T. G. (1981). Harmonic

ISBN 88-7395-155-4

structure as a determinant of melodic organization. *Memory and Cognition*, *9*(5), 533–539.

Temperley, D. (2001). *The Cognition of Basic Musical Structures*. Cambridge, MA: MIT Press.

Tenney, J. & Polansky, L. (1980). Temporal Gestalt perception in music. *Contemporary Music Review*, *24*(2), 205–241.

Thompson, W. F., Cuddy, L. L., & Plaus, C. (1997). Expectancies generated by melodic intervals: Evaluation of principles of melodic implication in a melody-completion task. *Perception and Psychophysics*, *59*(7), 1069–1076.

Tillmann, B. & McAdams, S. (2004). Implicit learning of musical timbre sequences: Statistical regularities confronted with acoustic (dis)similarities. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*(5), 1131–1142.

ISBN 88-7395-155-4