Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

The Cognition of Grouping Structure in Real-time Listening of Music. A GTTM-based Empirical Research on 6 and 8-year-old Children

Dimitra Koniari,[*1] Costas Tsougras[#2]

[*]*Department of Music Science and Art, University of Macedonia, Greece*
[#]*Department of Music Studies, Aristotle University of Thessaloniki, Greece*
[1]dkoniari@europe.com, [2]tsougras@mus.auth.gr

ABSTRACT

The aims of the present study are: a) to investigate how children of average ages 6 and 8 segment a musical piece during real-time listening, b) to compare children's indicated segment boundaries with boundaries obtained by the segmentation of the piece by adults (musicians and nonmusicians), and c) to compare the adults' and children's segmentation profiles to the structural boundaries predicted in a previous study by a full analysis of the piece according to the principles of GTTM. 70 children participated in the empirical study, of average age 6 and 8 (n =35 individuals for each Grade), as well as 50 adults (25 musicians and 25 nonmusicians). The performed boundaries were placed into two categories, depending on whether or not they were predicted by the analysis of the piece using the Generative Theory of Tonal Music (GTTM). Participants indicated a maximum of 38 segment boundaries. 16 corresponded to the boundaries predicted by the analysis of the piece with the use of GTTM, and 22 were not. The deviations in the range of values obtained from the 38 segment boundaries are also justified by the theory's principle of hierarchy, by the GS and TSR preference rules, and by the idiomatic features of the selected piece. The results suggest that even by the age of 6, children can perceive the grouping structure of a piece in accordance to the general laws expressed by the GTTM and by the age of 8 year-old children are almost perfect 'experienced listeners' of their musical culture, in accordance to the GTTM's principles.

I. INTRODUCTION

According to the Generative Theory of Tonal Music (GTTM) (Lerdahl & Jackendoff, 1983), an experienced to the tonal musical idiom listener intuitively organizes the musical sounds (the musical surface) of a piece into four types of hierarchical structures: grouping structure (GS), metrical structure (MS), time-span reduction (TSR), and prolongational reduction (PR). The mental organization of the musical surface on the level of GS has been characterized as "the most basic component of musical understanding" (Lerdahl & Jackendoff, 1983: 13) because it is responsible for the intuitive segmentation of the musical surface into hierarchically organised structures, such as motive, theme, theme-group, section etc. In that way, it helps the listener to perceive the basic units of a musical piece, that become an important input for the construction of other, more complicated kinds of musical structure.

The grouping component of the GS consists of two sets of rules: Grouping Well-Formedness Rules (GWFRs) and Grouping Preference Rules (GPRs). The first set of rules determines the grammatically correct possible structure of grouping patterns and the second establishes which among the formally possible structures correspond to the listener's actual intuitions. For this reason, the second set of rules is the important one from a cognitive perspective and is going to be further considered for the purposes of the present study.

Lerdahl & Jackendoff describe 7 GPRs (1983: 43-55). Among them, two rules, GPR2 (Proximity) and GPR3 (Change) stem from the general Gestalt laws of proximity and similarity, and specify possible boundary locations at instants in the musical flow where relatively contrastive discontinuities along one or several musical dimensions can be observed by the listener. For instance, the proximity rule sets a group boundary in a sequence of notes if there is a slur, rest, or a change in the duration of the attack point between adjacent notes (Lerdahl & Jackendoff, 1983: 45). Accordingly, the change rule sets a boundary in a sequence of notes when a change in register, dynamics, pattern of articulation, and length of notes (as well as in timbre or instrumentation) occurs (Lerdahl & Jackendoff, 1983: 46). The rest of GPRs determine how to combine the outcome of GPR2 and GPR3 at higher levels of the grouping hierarchy or how to relate the perception of boundaries to other features such as TSR (Time-Span Reduction), PR (Prolongational Reduction), symmetry, or parallelism/repetition of a group of notes.

While the above rules reflect the grouping intuitions of the ideal listener, an important number of experimental studies have been devoted to test, validate empirically, and/or develop, the above rules, with normal listeners, focusing, to the authors' knowledge, exclusively to adults, with or without musical training (musicians and nonmusicians) (Deliège, 1987; Clarke & Krumhansl, 1990; Frankland & Cohen, 2004; Bruderer, McKinney & Kohlrausch, 2009, Ordoñana & Laucirica, 2010).

Deliège (1987) was the first to assess the validity of the GTTM rules with adult listeners, musicians and nonmusicians. For every grouping rule of the GTTM, she used very short excerpts from the classical music repertoire and asked participants to segment them. The obtained segmentation performances largely supported the GPRs and she argued that the process of segmentation is not necessarily influenced by musical training, as she found no obvious differences between musicians' and nonmusicians' segmentation patterns. However, she commented that the validity of the GPRs proved to be more homogeneous in the musicians, justifying the notion of "experienced listener" proposed by Lerdahl & Jackendoff (GTTM: 3). She also found that an important contributor to perceptual boundaries in music has been shown to be the change in timbre. Finally, she introduced an additional rule to complement the existent GPRs: change in instrumental and/or sound density.

Clarke & Krumhansl (1990), using two pieces for piano, an atonal composed by Stockhausen and a tonal composed by Mozart, found also that adult listeners segmented the

presented musical pieces in broad agreement with the grouping rules of GTTM (1983).

Frankland and Cohen (2004) were the first to attempt a quantification of four of the components of Grouping Preference Rules 2 and 3: *rest* (GPR 2a), *attack-point* (GPR 2b), *register-change* (GPR 3a), and *length-change* (GPR 3d). They analyzed two well-known children's melodies and an unknown tonal melody and asked from musically experienced listeners to segment them. They then compared the obtained perceptual profiles with the predictions of the quantified rules. From the four rules, only two, *attack-point* (long note in between two short notes) and *rest*, were needed to account for the perceptual segmentations.

Bruderer, McKinney & Kohlrausch (2009) investigated the perception of structural boundaries in MIDI monophonic representations of six popular music songs and compared the obtained perceptual boundaries to boundaries predicted by two models and the four GTTM rules that were quantified by Frankland and Cohen (2004). They found a moderate correlation between the perceptual and predicted boundaries which was higher when the rule *change in timbre* (this rule is not a part of GPR 3, but it is mentioned as one of its possible extensions, see GTTM: 46) was considered. However, the most salient cue of the four quantified GPRs has been proved to be the *rest* rule (GPR 2a).

Ordoñana & Laucirica (2010) studied the construction of groupings, according to Lerdahl and Jakcendoff's GWFRs and GPRs (1983), in the performance of the first movement of Hindemith's sonata for flute and piano by three intermediate flute students and three famous concert performers. After analysing the score of the piece with the use of GTTM's grouping rules, they compared their analysis to the groupings (group beginnings and endings) carried out in the different performances of the piece by the two groups of musicians. The authors observed that there were some differences in the time-points selected for segmentation by the professional musicians and the students. However, all of them respected the set of rules formulated by Lerdahl and Jackendoff (1983).

II. AIMS

All the above-mentioned studies addressed grouping perception of adult listeners, with or without musical training. The present study addresses grouping perception of 6- and 8-year-old children. It will be focused only on the formation of the GS through the procedure of segmentation, i.e. the spontaneous decomposition of the musical flow into segments, by the above mentioned ages of children, and on its comparison with the same of adults, musicians and nonmusicians.

The used stimulus is a short piece for piano, by the Greek composer Yannis Constantinidis (1903-1984). The melody of the piece is based on a Greek folk tune, although not a very well-known one, presented with a harmonization that combines diatonic modality and twentieth century compositional techniques (Tsougras, 2003 & 2010).

The aims of the present study are: a) to investigate the selection of structural boundaries in real-time listening by children of average ages 6 and 8, b) to compare the boundaries chosen by children with boundaries obtained by the segmentation of the piece by adults (musicians and nonmusicians), and c) to compare the adults' and children's segmentation profiles to the structural boundaries predicted -in a previous study- by a full analysis of the piece according to the principles of GTTM (Tsougras, 2003).

III. METHOD

A. Participants

70 children participated in the present experiment, male and female pupils in Grades 1 and 3 of Primary school, of average age 6 (*n*=35, 17 boys and 18 girls, age ranged from 5 years and 11 months to 6 years and 10 months, with an average of 6 years and 4 months), and 8 (*n*=35, 17 boys and 18 girls, age ranged from 7 years and 11 months to 8 years and 10 months with an average of 8 years and 4 months), as well as 50 adults, students at the University of Macedonia. 25 adults had no musical training (12 male and 13 female, age ranged from 19 to 23 years with an average of 21.08 years; called 'nonmusicians') and 25 were students at the Department of Music Science and Art (12 male and 13 female, age ranged from 19 to 23 years with an average of 21.04 years; called 'musicians'). Musicians' musical training varied from 10 to 17 years of continuous study of a musical instrument, with an average of 15.08 years.

None of the children had any kind of formal musical training and none of the participants had ever listened before to the presented piece.

B. Stimulus

The music stimulus was Yannis Constantinidis's (1903-1984) piece no XXXVII from his collection for piano *44 Children's Pieces on Greek Melodies* (also known as *Greek Miniatures for Piano,* composed in 1949-1951 and published in 1957). The duration of the piece was 64 seconds, as performed by Domna Evnouhidou on the CD edition of the work by LYRA (Constantinidis, 1995: CD 2, track 4). The piece follows the normative *Theme and Variations* form (described as *Tema con variazioni* by the composer). The musical theme is a Greek folk tune, "*Τέσσερα και πέντε*" ("*Four and five"*) from Kastellorizo island, with binary phrase structure (an 8-bar period consisting of an antecedent 4-bar phrase and a consequent 4-bar phrase). It is presented initially as a monophonic melody and it is followed by 6 variations that have the length of the melody and differ in their harmonization and/or accompaniment. The piece ends with a coda containing repetitions of the beginning of the theme (Tsougras, 2003: 214-215). This piece was chosen among the other pieces of the collection for two main reasons: a) because its relatively simple form makes the musical structure conceivable by young children, b) because the original tune is little known outside Kastellorizo, making it suitable for an experiment carried out in Thessaloniki.

C. Apparatus

The experiment was carried out by the use of a special computer program, called *MusicalTests* (Neokleous & Ttofas, 2007), created for the experimental needs of the present study. This software allows each participant to listen to the piece through headphones and to indicate the time-points of segmentation while listening in real time, by pressing the space button on the computer keyboard. In a hidden text file,

the software records -in seconds and milliseconds- the temporal position of each pressing of the space button by each participant.

D. Procedure

Each participant was assessed alone with the experimenter in a quiet room. The experiment lasted 5 minutes and the experimental procedure was the following: Firstly, participants were told that they would listen to a piece that was very long and needed to be segmented into smaller sections. Then, they were asked to listen to it carefully and indicate the time-points where the piece could be segmented, i.e. the points where the music seemed to come to an end, producing a meaningful -for the listener- musical unit.

Participants were asked to go through the listening and segmenting procedure three times. The first two trials served for familiarisation with the piece and the experimental procedure. In the present study, only the third performance was considered for data analysis. This decision was made after taking into account (GTTM: 8-11) that a listener can construct a full hierarchical mental representation of a musical structure only after listening to the whole piece. So when he/she listens to it for a second or third time he/she can focus more clearly on the details of the musical structure in real-time. Segmentation boundaries from the first and the second listening/segmentation procedures will be used in the future for subsequent research on real-time perception of segmentation boundaries.

IV. RESULTS

The data collected through the experiment revealed differences between the time-points of the boundary indications obtained from the participants. For this reason, we accumulated -across participants- segmentations that were close in time and the resulting clusters have been relocated to the end of the note that was playing at the time-point of the first segmentation indication of each cluster. By this procedure, participants' segmentation indications were regrouped into 38 general segmentation boundaries (Fig. 1).

A comparison was made between these boundaries and the GS segmentation carried out -in a previous study- as part of a complete GTTM analysis of the piece (Tsougras, 2003: 214-220). The 38 obtained segmentation boundaries were placed into two categories, depending on whether or not they corresponded to the boundaries predicted by the analysis of the piece with the use of GTTM at level e of the Grouping Structure (4-bar phrase level). 16 segmentations corresponded to the boundaries predicted by the analysis of the piece and 22 were not. All of these boundaries were not necessarily chosen by each participant. However, the segment boundaries that were predicted by the analysis of the piece with the use of GTTM were chosen more often than the others. The deviations in the range of values obtained from the 38 segment boundaries can be justified by the following possible reasons: 1) the boundaries that correspond to phrase level (e) are enhanced by local or global melodic/harmonic cadences at the Time-Span Reduction (TSR), by the phrase symmetry preference rule (GPR5), and by the binary regularity metrical preference rule (MPR10), making them stronger boundary candidates, 2) certain boundaries correspond not only to the grouping boundaries of phrase level (e), but to higher

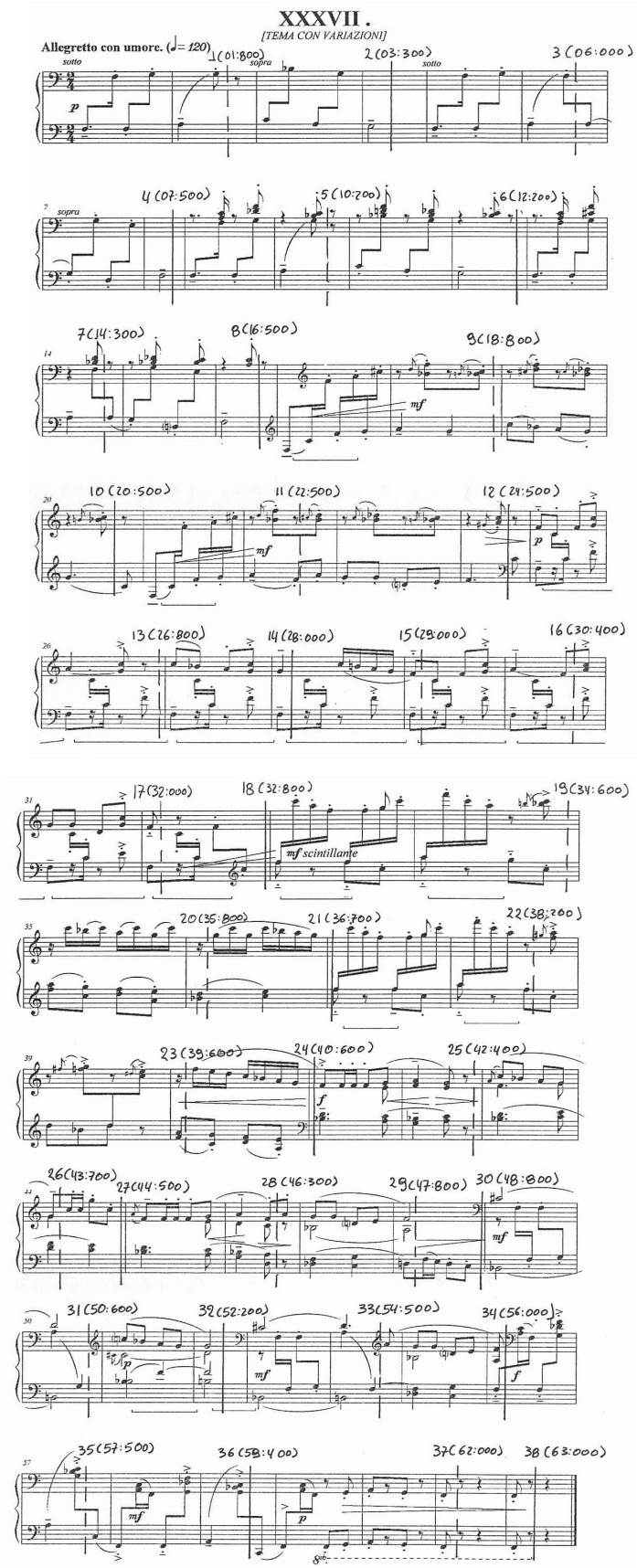

Figure 1. The piece XXXVII from the "44 Children's Pieces on Greek Melodies" by Yannis Constantinidis. Steady lines indicate boundaries predicted by the GTTM analysis of the piece (level e of the GS) and perceived by the participants. Doted lines indicate segment boundaries (corresponding to level f of the GS) perceived by the participants.

hierarchical levels also (period level d for b. 8, 16, 24, 32, 40, 48, 56 and sectional level c for b. 24, 40 and 56), making them stronger candidates, 3) the boundaries that correspond to sub-phrase level (f) tend to be more ambiguous, due to the conflict of different preference rules, such as note length (GPR3d) or parallelism (GPR6), and 4) certain boundaries are more ambiguous due to less stable harmonization (e.g. b. 48 features a deceptive cadence) or grouping elision (the onset of the coda at b. 56).

In more detail, participants of the group of 6-year-old averaged 57.88% of boundaries (169 segmentation points) in accordance with the GTTM phrase-level predictions and 42.12% (123 segmentation points) not in accordance. At the 8-year-old group, a steady reduction in the frequency of segmentations at non predicted boundaries is observed (27.46%: 95 segmentation points) as well as a steady increase in the frequency of the predicted ones by the GTTM (72.54%: 251 segmentation points). 8-year-old children responded significantly more in accordance with the GTTM rules than 6-year-old children did (p=0.001).

Nonmusicians averaged 81.73% of responses (170 segmentation points) in accordance with the GTTM analysis and 18.27% of responses (95 segmentation points) not in accordance. Musicians averaged 95.42% of responses (200 segmentation points) in accordance and only 4.58% of responses (11 segmentation points) not in accordance (Fig. 2 and Tables 1 &2).

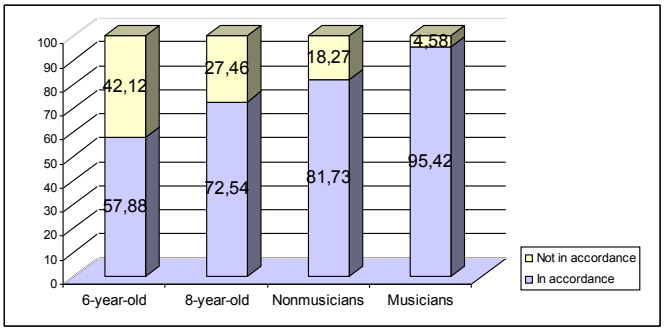

Figure 2. Percent of segmentation in boundaries predicted by the analysis of the piece with the use of GTTM at level e of the GS (bottom part of the column) and not predicted (upper part of the column)

Table 1. Distribution parameters for the segmentation decisions at boundaries that are in accordance with the GTTM's GPR rules predictions at level e of the GS

	6-year-old	8-year-old	Nonmus	Musicians
n	35	35	25	25
Total	169	251	170	229
Mean	4.83	7.17	6.80	9.16
Mean Dist	2.56	2.15	2.3	2.91
Mode	5	7	6	8
minimum	1	3	3	4
maximum	11	11	12	14

Table 2. Distribution parameters for the segmentation decisions at points that deviate from the GTTM's GPR rules predictions at level e of the GS

	6-year-old	8-year-old	Nonmus	Musicians
n	35	35	25	25
Total	123	95	39	11
Mean	3.51	2.71	1.56	0.44
Mean Dist	2.39	1.62	1.45	0.71
Mode	3	2	1	0
minimum	0	0	0	0
maximum	8	7	5	2

V. CONCLUSIONS

The comparison of participants' segmentation profiles with the boundaries predicted by the analysis of the piece with the use of GTTM rules showed that the predicted phrase-level boundaries were indicated by nearly all participants, although in a different percentage by each group. Out of the 16 segmentation boundaries that were predicted by the GTTM rules, segmentation indications of the 6-year-old group were in accordance with them in a percentage of 57.88% and of the 8-year-old group in a percentage of 72.54%. In the groups of adults the percentage of chosen segmentation boundaries in accordance with the structural boundaries predicted by the GTTM arises significantly. Nonmusicians' segmentation responses are at 81.73% in accordance with GTTM predictions and musicians' responses at 95.42%. However, the fact that the remaining segmentation indications, that correspond to lower hierarchical levels, are not in accordance with the phrase-level GTTM Grouping Structure analysis does not mean that these segmentation marks do not follow the Grouping Rules formulated by Lerdahl and Jackendoff (1983). The GRs are followed, but the inferred grouping structure is more ambiguous, due to the conflict of preference rules at these sub-phrase levels.

The auditory processes governing grouping and segmentation of sounds are candidate to be considered as automatic, independent of the listener's musical skill and, consequently, to represent universal aspects of music perception (Stevens & Byron, 2009). Even 6-and 4½-month-old infants have shown sensitivity to phrase structure in music. In a relevant experiment, infants tend to listen for significantly longer periods of time to segmented versions with relatively long notes and downward pitch contours, over incorrectly segmented pieces (Krumhansl & Jusczyk, 1990). Deliège (1987) suggested that the psychological processes involved in the process of segmentation while listening is not strongly influenced by musical training. Additionally, Lerdahl & Jackendoff (1983) have noted that their rules for grouping seem to be idiom-independent, that is, a listener needs to know relatively little about a musical idiom in order to assign grouping structure to pieces in that idiom (1983: 36).

However, there is evidence that not all aspects of auditory grouping are universal, that some forms of auditory grouping depend on musical skill. For instance, it has been shown that musicians and nonmusicians differ in their ability to preattentively group consecutive sounds according to good continuation of pitch but not according to pitch similarity (van Zuigen, et al, 2004). In the present study we worked with children which had no form of musical training. The aim was to investigate the cognition of listeners who are not biased from music education. However, even the mere exposure to music can influence listeners' music perception through the process of implicit learning (Bigand & Poulin-Charronnat, 2006; Tillmann et al, 1998).

Finally, our results suggest that even by the age of 6, children can perceive the grouping structure of a piece in accordance to the general laws expressed by the GTTM. This ability is strengthened by the age. 8-year-old children are almost perfect experienced listeners of their musical culture and follow the GTTM GS rules during music listening. The present research is restricted to the presentation of the

percentage of each group's segmentation indications in relation with segmentation boundaries that can be predicted by the GTTM grouping rules. The detailed analysis of the possible mappings between the participants' segmentations and the GTTM's prediction will be used for subsequent research. Additionally, future research should continue in this vein using also children of the same ages with formal musical training.

ACKNOWLEDGMENT

The authors thank all study participants for their cooperation. Special thanks go to the children for their joyful participation and to their parents and teachers for their consent. The research has been carried out with the approvement and permission of the Hellenic Pedagogical Institution.

REFERENCES

Bigand, E., & Poulin-Charronnat, B. (2006). Are we «experienced listeners»? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100*, 100–130.

Bruderer, M., McKinney, M.F., & Kohlrausch, A. (2009). The perception of structural boundaries in melody lines of Western popular music. *Musicae Scientiae, 13*, 273–313.

Clarke, E. F., & Krumhansl, C. L. (1990). Perceiving musical time. *Music Perception, 7*, 213-251.

Constantinidis, Y. (1957). *Greek miniatures for piano (vol. 1-3)*. New York, Rongwen Music.

Constantinidis, Y. (1995). *Yannis Constantinidis - The Piano Works* (Evnouhidou, D., piano) [CD]. Athens: LYRA, 0174/75.

Deliège, I. (1987). Grouping Conditions in Listening to Music: An Approach to Lerdahl and Jackendoff's Grouping Preference Rules. *Music Perception, 4*, 325–360.

Deliège, I., Mélen, M., Stammers, D., & Cross, I. (1996). Musical schemata in real-time listening to a piece of music. *Music Perception, 14*, 117-160.

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*, 499–543.

Krumhansl, C. & Jusczyk, P. W. (1990). Infants' Perception of Phrase Structure in Music. *Psychological Science, 1*, 70-73.

Lartillot, O., & Ayari, M. (2008). Segmenting Arabic modal improvisation: Comparing listeners' responses with computer predictions. In *Proceedings of the Fourth Conference on Interdisciplinary Musicology (CIM08)*. Greece: Aristotle University of Thessaloniki. Retrieved November 25, 2009, from http://web.auth.gr/cim08/.

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

Neokleous, A. & Ttofa, P. (2007). *Musical Tests: Musical experiments for children* [in Greek]. Unpublished Graduation Thesis. University of Macedonia: Department of Applied Informatics.

Ordoñana, J.A., & Laucirica, A. (2010). Lerdahl and Jackendoff's Grouping Structure Rules in the Performance of a Hindemith Sonata. *The Spanish Journal of Psychology, 13*, 101–111.

Stevens, C., & Byron, T. (2009). Universals in Music Processing. In S. Hallam, I. Cross, & M. Thaut (Eds.) *Oxford Handbook of Music Psychology* (pp. 14-23). Oxford: Oxford University Press.

Tillmann, B., & Bigand, E. (1998). Influence of global structure on musical target detection and recognition. *International Journal of Psychology, 33*, 107–122.

Tillmann, B., Bigand, E., & Madurell, F. (1998). Local versus global processing of harmonic cadences in the solution of musical puzzles. *Psychological Research, 61*, 157–174.

Tsougras, C. (2003). *Yannis Constantinidis "44 Greek miniatures for piano". Analysis with the use of the Generative Theory of Tonal Music* [in Greek]. Thessaloniki: University Studio Press.

Tsougras, C. (2010). The application of GTTM on 20th century modal music: Research based on the analysis of Yannis Constantinidis's "44 Greek miniatures for piano". *Musicae Scientiae*, *Discussion Forum 5*, 157-194.

van Zuigen, T.L., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2004). Grouping of sequential sounds – An event-related potential study comparing musicians and nonmusicians. *Journal of Cognitive Neuroscience, 16*, 331–338.