The Legacy of Lerdahl and Jackendoff's *A Generative Theory of Tonal Music*

Bridging a significant event in the history of music theory and recent developments in cognitive music research

Niels Chr. Hansen

Not unlike many other representatives of North-American theory-based analysis, Fred Lerdahl and Ray Jackendoff's *A Generative Theory of Tonal Music*[1] (*GTTM*) has only received limited attention in Danish music theory. In fact, only a single reference to *GTTM* has appeared in previous issues of this journal, and that was in passing in a footnote.[2] Still, the theory has been cited extensively elsewhere (currently counting more than 2,300 hits in the Google Scholar Citation Index). Some familiarity with Lerdahl and Jackendoff's approach and its legacy thus seems crucial if one wants to fully assess the vast range of literature published in the wake of it. The primary aim of this paper is to provide such an overview by offering a qualitative account, tailored to a contemporary reader, of *GTTM*'s – at least in quantitative terms – irrefutable influence.

In recent years, aided by modern neuroimaging techniques, research in music cognition has gained territory in Denmark.[3] This field is characterized by vast degrees of interdisciplinarity based on hard-core empirical methodologies, but unmistakably dependant on cognitive theories to generate useful hypotheses. Such input has typically come from psychology where a 'cognitive revolution' took place during the 1950s and 1960s breaking with the previous dominance of behaviourism and establishing connections between psychology, linguistics, and the concurrently expanding field of computer science.[4] However, particularly from the 1980s and onwards, theoretical input for empirical music research increasingly came from music theory.

Celebrating its 25th anniversary in 2008, *GTTM* was an early example of such a contribution, representing an important step towards modelling the hierarchical properties of music cognition in humans. With *GTTM* Lerdahl and Jackendoff proposed a grammatical rule system of Western tonal music initiated by Leonard

1 Fred Lerdahl and Ray Jackendoff, *A Generative Theory of Tonal Music* (Cambridge: MIT Press, 1983). I would like to acknowledge Dr Michiel Schuijer for invaluable support when this project was initiated during my stay at Conservatorium van Amsterdam in 2008–9.

2 Anders Bonde, 'Algoritmisk mønsteridentifikation: Nogle betragtninger omkring computeranvendelse i musikanalytisk øjemed', *Danish Yearbook of Musicology*, 33 (2005), 77–105, n. 5.

3 Peter Vuust, 'Perception, Cognition and Learning: Cognitive Music Research at the Music Academies in Denmark', *Danish Yearbook of Musicology*, 36 (2008), 9–19.

4 Bernhard J. Baars, *The Cognitive Revolution in Psychology* (New York: The Guilford Press, 1986).

Bernstein's invitation to search for a musical grammar[5] and inspired by the laws of Gestalt psychology as well as by generative linguistics, the foremost representative of which was at that time Jackendoff's teacher Noam Chomsky.

The authors of *GTTM* argued that '[t]here is much more to music than the raw uninterpreted physical signal … a piece of music is a mentally constructed entity, of which scores and performances are partial representations by which the piece is transmitted'.[6] This altered the analytical perspective dramatically by changing the primary study object from musical structure to the listening process. Described in terms of Jean-Jacques Nattiez' tripartite distinction between the 'neutral' (the work itself), 'poietic' (the composer's intentions) and the 'esthesic' levels (the listener's cognitive percept, induced emotions, etc.),[7] *GTTM* changed the focus of traditional music analysis from the first two categories towards the last. Lerdahl and Jackendoff began considering music theory as 'a branch of cognitive science' and 'as the branch of theoretical psychology concerned with modelling the musical mind'.[8] In short, *GTTM*'s formulations had to be testable by the methods of experimental psychology, thus discarding hermeneutic approaches hitherto dominant in the field of music theory. In the following I will refer to this novel philosophical mindset as the 'cognitive paradigm'.

GTTM's recent anniversary seems like an apposite occasion for re-assessing the theory on the basis of its influence on music theory and cognitive research, and the timespan since its publication furthermore enables us to regard it from a comfortable historical perspective with the pleasant wisdom of hindsight. In sum, the current topicality of cognitive music research makes it more than just a curiosity for an audience of Danish theorists and musicologists to gain acquaintance with *GTTM*.

A brief overview of *GTTM*

With an abundance of available summaries,[9] I will restrain the following overview to the absolute basics and to specific concepts that will be taken up at later points.

In *GTTM* the authors distinguish between three kinds of accents: (1) *phenomenal accents* caused by e.g. changes in dynamics, *sforzandi*, long notes, harmonic changes and melodic leaps; (2) *structural accents* caused by melodic and harmonic points of gravity, especially in cadential contexts; and finally (3) *metrical accents* representing the relative importance of a given time-point in the inferred structure.

5 Leonard Bernstein, *The Unanswered Question* (Cambridge: Harvard University Press, 1976), which was reviewed by Jackendoff in 1977.

6 Lerdahl and Jackendoff, *GTTM*, 2.

7 Jean-Jacques Nattiez, *Music and Discourse: Toward a Semiology of Music* (New Jersey: Princeton University Press, 1990).

8 Fred Lerdahl, *Tonal Pitch Space* (Oxford: Oxford University Press, 2001), 4 and vii.

9 See e.g. Burton S. Rosner, 'Review', *Music Perception*, 2 (1984), 275–90; Eric F. Clarke, 'Theory, analysis and the psychology of music: A critical evaluation of Lerdahl, F. & Jackendoff, R. A generative theory of tonal music', *Psychology of Music*, 14 (1986), 3–16; Fred Lerdahl and Ray Jackendoff, 'An overview of hierarchical structure in music', *Music Perception*, 1/2 (1983), 229–52.

The theory supposes four components to influence music cognition: 'Grouping structure', 'metrical structure', 'time-span reduction' and 'prolongational reduction'. Figure 1 gives an example of how each of these four components is addressed in an analysis of the first eight bars from J. S. Bach's chorale 'Christus, der ist mein Leben'.

Grouping structure is notated with slurs beneath the score (see Figure 1) and refers to the generic term for motifs, phrases, themes, theme-groups, sections, and complete pieces into which the listener segments music while listening. In informal terms it tells us where to breathe when singing a melody. Segmentation mostly takes place intuitively, and the grouping component is thought of chiefly as idiom-independent and thus cross-cultural.

Metrical structure is the hierarchical pattern of strong and weak beats inferred by the listener on the music taking local phenomenal accents in the musical surface as its input. In later research this has been referred to as 'beat induction',[10] and informally metrical structure can be described as the pattern in which a conductor moves her baton or the listener taps his feet in time with the music. Since beats have no duration, metrical structure is notated with dots, the number of which determines the strength of a given metrical accent.

Time-span reduction constitutes a link between pitch and rhythm and is represented by a recursive tree and/or stave notation (as in Figure 1) indicating the relative structural accent of musical events.

Prolongational reduction is represented by another tree hierarchy (Figure 1, top) and/or stave notation (Figure 1, bottom) reflecting patterns of perceived tension and relaxation. Contrary to the bottom-up approach used in time-span reduction, prolongational reduction proceeds from global to local levels by a top-down procedure. Although the two reductions often correlate, contrasts between them add to the sense of tonal tension in a piece, thus constituting a major force in musical form. Sometimes prolongational trees conform to the 'normative structure' which is comparable to a Schenkerian *Ursatz*, but more generalized and far from aesthetically prescriptive.

GTTM is 'generative' in the sense that it proposes a finite set of rules enabling an infinite number of possible musical structures. Firstly, Well-Formedness Rules (WFRs) determine which structures are possible. Secondly, Preference Rules (PRs) establish factors influencing the listener's choice between different well-formed structures. The PRs reflect the 'Law of Prägnanz', a key tenet of Gestalt pscyhology stating that we automatically order our experience in a manner as simple, regular, and symmetric as possible, grouping objects according to e.g. the two interacting principles of 'proximity' and 'similarity'. In some cases there are also Transformational Rules (TRs) accounting for phenomena (e.g. elisions) conflicting with the well-formedness conditions by describing how an underlying structure can in some cases be transformed into an alternative surface structure. Despite the significance of TRs in linguistic grammar, they only play a peripheral role in *GTTM*.

10 For a review of rule-based models of beat induction until 1999, see Peter Desain and Henkjan Honing, 'Computational models of beat induction: The rule-based approach', *Journal of New Music Research*, 28/1 (1999), 29–42.

Figure 1. Hierarchical analysis of bb. 1–8 from J. S. Bach's chorale 'Christus, der ist mein Leben'. *Tonal Pitch Space* by Fred Lerdahl (2001), Fig. 1.18, pp. 22–23. By permission of Oxford University Press, Inc.

Typical points of criticism

After an overview of the analytical system I will now delve into some key concepts and tenets underlying *GTTM* to throw light upon typical points of criticism raised against the theory by later scholars. I will touch upon *GTTM*'s view on music, hierarchical properties of music listening, the rule system, issues of formalism, testability and its presumptions about universality, innateness, and the 'idealized' listener.

A simplified view on music

A well-known topic in musicological debate is the discussion whether analysis should focus on structural coherence or parsing. According to Lerdahl, 'nineteenth-century analytic approaches … tended to emphasise motivic, phrasal, and sectional parsings. Schenker, with his composed-out voice-leading structures, went to the opposite extreme … In *GTTM* this structural counterpoint is revealed through a comparison of its grouping and prolongational analyses'.[11] That is, *GTTM* was established as a claimed synthesis of a previously unresolved dichotomy. One can, however, question whether the partitioning into four separate components with no clear image of a single, 'final' representation really represents a synthesis. It is similarly questionable whether an attempt of unification is novel at all. The idea of simultaneously striving forces was already present in Schenker's theory where the composer was thought of as opposing Nature by composing-out the stable 'Chord of Nature' by means of counterpoint and prolongation.[12]

Additionally, *GTTM* is unable to cope with polyphonic textures. Although the authors were indeed conscious about this shortcoming,[13] the exclusively homophonic view on music is nonetheless the one underlying their theory. It is thus questionable whether *GTTM* is capable of dealing with the whole repertoire from which the authors draw their examples. The theory would in particular have difficulties with polyphonic textures in developmental passages from the Classical Era, and Lerdahl and Jackendoff provide no complete analysis of a sonata movement. In time-span reduction two or more notes can be 'fused', but all such examples refer to instances of pseudo-polyphony, thus insinuating that fusion is in fact an ad hoc compensation for *GTTM*'s inability to account for polyphony.

GTTM's unequivocal emphasis on hierarchical listening has by some scholars been considered an artefact of its reductive view on music as sheer scores which is all in all inconsistent with the authors' intention of investigating properties of music listening. In short, when listening to music, we listen to performances and not to

11 Lerdahl, *Tonal Pitch Space*, 24.

12 Heinrich Schenker, *Harmony* (Chicago: University of Chicago Press, 1980), 44: 'the [musical] system is to be considered, accordingly, as a compromise between Nature and Art, a combination of natural and artistic elements'.

13 Lerdahl and Jackendoff, *GTTM*, 37.

notations of it. Already one of *GTTM*'s first reviewers, Henry Cady, ascertained that tree notation works towards scores, but asked rhetorically whether it really reflects real-time, cognitive processes.[14] Music is a temporal art form and thus disables us from viewing an entire piece as a simultaneous whole. In Cady's view, Lerdahl and Jackendoff seem to avoid mentioning the process of listening, but then do assume it for many PRs.

Global versus local listening

Lerdahl emphasizes that *GTTM* 'provides structural descriptions not for how the music is heard as it unfolds in time but for the final state of a listener's understanding'.[15] Similarly central to *GTTM* is the 'Reduction Hypothesis' stating that '[t]he listener attempts to organize all the pitch-events of a piece into a single coherent structure, such that they are heard in a hierarchy of relative importance'.[16] The theory considers listening from a global perspective as a final state rather than locally as a continuous process; i.e., the listener's representation of music is assumed to take place retrospectively rather than consecutively.

Lerdahl and Jackendoff support their 'Reduction Hypothesis' by claiming that 'linear-motivic aspects of pitch structure cannot be given proper systematic treatment without a theory of the hierarchical structures within which they are heard'.[17] Although shared by aspects of Schenkerian theory, the validity of this tenet is not obvious.[18] Conversely, numerous theories of music cognition are based on consecutive violation and confirmation of expectancy.[19] Related ideas have appeared in humanistic musicology and music philosophy.[20]

The cognitive theories mentioned above have received support from empirical findings. Memory constraints influence musical processing by impeding the ability to listen hierarchically, and it has been shown that even expert listeners are unable to distinguish original pieces from altered versions ending in another key.[21] Huron argues that, when selecting preferred cognitive schemas, the mind always has to compromise between high predictive power and low information content.[22] For instance, pitch direction (i.e. contour) in melodies is structured

14 Henry Cady, 'Book review', *Psychomusicology*, 3/1 (1983), 60–67.

15 Lerdahl, *Tonal Pitch Space*, 5.

16 Lerdahl and Jackendoff, *GTTM*, 106.

17 Ibid. 117.

18 In Lerdahl and Jackendoff's defence, however, the strict hierarchy implied by the Reduction Hypothesis was introduced at a relatively late point in the inception of *GTTM*; e.g. absent from Fred Lerdahl and Ray Jackendoff, 'Toward a Formal Theory of Tonal Music', *Journal of Music Theory*, 21/1 (1977), 11–171.

19 E.g., Leonard B. Meyer, *Emotion and meaning in music* (Chicago: University of Chicago Press, 1956); Eugene Narmour, *The Analysis and Cognition of Basic Melodic Structures – The Implication-Realization Model* (Chicago: The University of Chicago Press, 1990); David Huron, *Sweet Anticipation* (Cambridge, Mass.: MIT Press, 2006).

20 Jerrold Levinson, *Music in the Moment* (Ithaca, NY: Cornell University Press, 1997).

21 Nicholas Cook, 'The perception of large-scale tonal closure', *Music Perception*, 5 (1987), 197–205.

22 Huron, *Sweet Anticipation*, 122–25.

due to the principle of 'regression to the mean' governing all central-tendency distributions. Nevertheless, research suggests that listeners apply the alternative mental representation 'post-skip reversal' where a large interval is expected to be followed by a change in direction because the latter representation has nearly as strong predictive power but considerably lower information content (listeners do not have to maintain information about all previous pitches necessary for constantly recalculating the mean). Moreover, theorists have argued that *GTTM*'s focus on global listening is inconsistent with memory constraints and lacks empirical support; another likely source of influence is Schenkerian analysis, which had already been dominant in American theory for some decades prior to *GTTM*'s appearance.[23]

Nevertheless, peculiarly, in one case Lerdahl and Jackendoff do refrain from strict hierarchical organisation due to lacking perceptual salience. Their hierarchical beat concept does not extend into global levels since they consider metrical structure as a relatively local phenomenon. Further levels beyond five or six are considered 'perceptually irrelevant'.[24]

In hierarchical models like *GTTM*, elements subsume or contain other elements. However, many musical phenomena are more likely related by association. This is true for motives and for chords related by 'substantial affinity' rather than 'functional affinity'.[25] Lerdahl and Jackendoff do hint at motivic associations, but still claim that 'they are not the grouping structure that he [the listener] hears', continuing '[b]ecause associational structure is not hierarchical … our theory has little to say about it'.[26] Nonetheless, if the authors acknowledge that listeners make such associations, they ought not ignore it.

As previously mentioned, time-span and prolongational reductions proceed in opposite directions. Reviewers have regarded this indecisiveness between 'bottom-up' and 'top-down' procedures as somewhat unsatisfactory.[27] Interestingly, this tension does not only exist between the four components of *GTTM*, but also evokes internal conflicts within the components where some PRs work in a global manner ('top-down') whereas others work locally ('bottom-up').[28] Such conflicts render the analytical result ambiguous and thus impede generation of falsifiable hypotheses.

23 This is e.g. argued by Zofia Helman, 'Von Heinrich Schenkers analytischer Methode bis zur generativen Theorie der tonalen Musik', *International Review of the Aesthetics and Sociology of Music*, 19/2 (1988), 181–95.

24 Lerdahl and Jackendoff, *GTTM*, 21.

25 Teresa Waskowska Larsen and Jan Maegaard, *Indføring i romantisk harmonik* (Copenhagen: Engstrøm og Sødring, 1981).

26 Lerdahl and Jackendoff, *GTTM*, 16–17.

27 Clarke, 'Theory, analysis'.

28 Keiji Hirata, Satoshi Tojo, and Masatoshi Hamanaka, 'Techniques for implementing the Generative Theory of Tonal Music', *ISMIR Tutorial*, http://ismir2007.ismir.net/proceedings/ISMIR2007_tutorial_hirata.pdf, 23 Sept. 2007, slide 67.

The rule system of *GTTM*

A consequence of the cognitive paradigm was the establishment of 'a crucial distinction between the principles by which a piece is composed and the principles by which it is heard',[29] referred to in later publications as the 'compositional' and 'listening grammar'.[30] The authors' choice to focus on the latter significantly changed the role played by rules in music theory. Traditionally, rules represented instructions on how to compose counterpoint and harmony. On the contrary, rules in *GTTM* are associated with the analytical listening process modelling human perception and cognition.

Prior to *GTTM*'s redefinition of rules, Sundberg and Lindblom presented a rule system capable of generating songs in the style of Swedish nursery rhymes.[31] However, even though referring to Chomsky, their understanding of the term 'generative' differed from Chomsky's generative grammar due to which the term should be understood in its mathematical sense referred to above rather than as a mechanistic algorithm generating sentences (or musical pieces).

Contrary to the WFRs, PRs do not constitute categorical demands. Rather their degree of fulfilment represents degrees of clarity and perceptive unambiguity.[32] The PRs were subject to severe criticism from some reviewers describing them as 're-writings of Gestalt laws that have been shown, in other contexts, to fall short of providing that theoretical framework. This criticism is not simply one of form, but of substance. Once preference rules are introduced, the theoretical apparatus becomes fatally flawed'.[33] They proceed to conclude that Schenker's 'rules are clear and considerably stronger in their assertions than those of Lerdahl and Jackendoff' criticizing Lerdahl and Jackendoff for ignoring the all-important voice-leading aspect of musical structure. In their reply, Lerdahl and Jackendoff acknowledged 'that the PR system [was] not yet predictive enough', but they excluded that PRs should represent a wrong kind of rule system by referring to the prominence of empirically supported PR-like principles in Gestalt psychology, theories of vision, music psychology, and theoretical linguistics.[34] To this one might add optimality theory, Bayesian inference, and 'goodness-of-fit' models which play prominent roles in other cognitive theories, computer programming, and machine learning.

29 Lerdahl and Jackendoff, *GTTM*, 298.

30 Fred Lerdahl, 'Cognitive Constraints on Compositional Systems', in John A. Sloboda (ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition* (New York: Oxford University Press, 1988), 231–59; Fred Lerdahl, 'Pitch-Space Journeys in Two Chopin Preludes', in Mari Riess Jones and Susan Holleran (eds.), *Cognitive Bases of Musical Communication* (Washington D.C.: American Psychological Association, 1992), 171–91.

31 Johan Sundberg and Bjorn Lindblom, 'Generative Theories in Language and Music Description', *Cognition*, 4 (1976), 99–122.

32 Lerdahl, *Tonal Pitch Space*, 6–7.

33 John Peel and Wayne Slawson, 'Review', *Journal of Music Theory*, 28/2 (1984), 271–94. Similar points were raised in another brief, but extremely critical, review by Christopher Longuet-Higgins, 'All in theory: The analysis of music', *Nature*, 304 (1983), 93.

34 Fred Lerdahl and Ray Jackendoff, 'A reply to Peel & Slawson's review of *A Generative Theory of Tonal Music*', *Journal of Music Theory*, 29/1 (1985), 145–60.

Formalism and testability

Many PRs are, as earlier mentioned, direct manifestations of the Gestalt principles of 'proximity' and 'similarity'. Gestalt psychologists were, however, vehemently criticized for lack of formalism.[35] Apparently, Lerdahl and Jackendoff wanted to counter such criticism with their very formalist rule system.

One example of formalism is the formulation of two separate TRs for 'elision' and 'overlap' in the grouping structure although it seems to be a difference of degrees rather than a categorical one. Furthermore, this distinction does not seem to make any substantial difference to the theory. Formalism also occurs when the authors provide multiple versions of rules. The first grouping PR is initially formulated as '[s]trongly avoid groups containing a single event' and subsequently in an 'alternative form' telling us to '[a]void analyses with very small groups – the smaller, the less preferable'.[36] Since, however, PRs are by definition flexible, there should be no need for the initial one. It is as if the authors feel obliged to provide strict formulations to comply with criticism from the community of generative linguistics although musical intuition tells them that music cognition cannot be modelled in such an inflexible manner. Their rigid approach to analysis might have misled them into unnecessary complexity.

Lerdahl and Jackendoff cannot exclude that some components of music cognition are explicable in simpler terms. E.g., although music is an art form, numerous empirical findings suggest that we also react intuitively to basic, notably threatening, acoustic features like extreme pitch, sudden loudness, and dissonance outside – *and* within – musical contexts.

Despite formalistic tendencies, Lerdahl and Jackendoff took some steps to accommodate the diversity of their audience by simplifying terminology and notation. In some respects, *GTTM* might be criticized for being imprecise – or even insufficiently rigid. Unlike musical set theory and related theories from the preceding decades, Lerdahl and Jackendoff did not adopt advanced terminology and notational devices from mathematics and logic. The only external device was tree notation from generative grammar. *GTTM*'s application of it was, however, described as 'purely musical', and in certain ways it did not conform to the linguistic version with trees representing '*is-a*' relations where two grammatical categories (e.g., a verb and a noun) go together to form a third grammatical category (in this case, a verb phrase).[37] Conversely, *GTTM* represents 'elaborational' relations where certain events coexist on various levels without being merged into another category.[38]

To some extent, the absence of advanced terminology and notational devices rendered *GTTM* more accessible to musicologists, theorists, and musicians than e.g. pitch-class set theory. By comparison, in the 1960s, set theorists felt obliged to

35 Lerdahl and Jackendoff, *GTTM*, 306.
36 Ibid. 43.
37 Ibid. 113.
38 Lerdahl and Jackendoff, 'An overview' (1983).

conform to the dominant science paradigm. Their theories were primarily associated with this paradigm, and minor inconsistencies were considered as unacceptable flaws rather than as acceptable traits characterizing the specific application of set-theoretical concepts to the musical domain. On the contrary, *GTTM*'s authors acknowledged that their application of generative grammar was necessarily less rigid than the original one simply because 'music is not tied down to specific meanings and functions, as language is'.[39]

Despite *GTTM*'s disapproved tendency towards formalism, some concepts were in fact insufficiently defined. In PRs of all components, for instance, great significance is assigned to 'parallelism'. Nonetheless, lacking a strict definition of this concept, such PRs are difficult to apply in analysis[40] and nearly impossible to implement in computational models.[41]

Furthermore, some PRs are mutually dependent in a manner making them difficult to handle in practice. *GTTM*'s problem of 'circular definitions' has been mentioned by commentators[42] and is a frequent criticism towards many strands of music analysis. A common example is the arbitrariness of segmentation in pitch-class set analysis, and in Schenkerian analysis structural notes are sometimes selected simply due to their capability of demonstrating the concept of the *Urlinie*. Interdependence of individual components also conflicts with another key tenet, namely *GTTM*'s ability to generate empirically testable hypotheses.[43]

Finally, the authors' remarks on brain localization seem rather tentative. Although this was due to the basic state of neuroimaging techniques in the early 1980s, it is still remarkable that they do not outline more direct connections between *GTTM* and cognitive neuroscience. If one does not know what findings to expect from the theory, then *GTTM* cannot be characterized as hypothesis-generating.

Universality, innateness, and the 'idealized' listener

From its very title it is evident that *GTTM* is a theory of Western, tonal music. The authors substantiate their focus by stating that 'one cannot hope to address in any deep way the question of musical universals without first developing a precise theory of at least one complex musical idiom'.[44] Still, they claim universality of their theory with the exception of a few idiom-specific rules. This viewpoint is ascribable to the cognitive paradigm, and may be regarded as a novelty in music theory where e.g. Schenkerian and set theory apply to tonal and atonal music, respectively. In the wake of universality claims, the question arises whether cognitive capacities for music are innate. Lerdahl and Jackendoff tend to think so; perfectly in lines with

39 Lerdahl and Jackendoff, *GTTM*, 9.
40 Clarke, 'Theory, analysis'.
41 Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, 'ATTA: Automatic Time-Span Analyzer based on extended *GTTM*', *Proceedings of ISMIR*, 2005, 358–65.
42 E.g., Hirata et al., 'Techniques for implementing'.
43 Lerdahl, *Tonal Pitch Space*, vii.
44 Fred Lerdahl and Ray Jackendoff, 'An overview of hierarchical structures in music', in Stephan M. Schwanauer, *Machine Models of Music* (Cambridge: MIT Press, 1993), 289–312.

Chomsky's generative grammar where innateness was similarly claimed for knowledge of grammatical structure.[45]

Reviewers have criticised *GTTM*'s authors for making non-falsifiable claims of universality and innateness based on intuition rather than on intercultural research and for drawing arbitrary distinctions between universal and idiom-specific PRs.[46] Lerdahl and Jackendoff themselves openly admit their 'own ignorance of other [musical] idioms'.[47] Presenting a theory of tonal music, Lerdahl and Jackendoff are not strictly obliged to prove their claims; generating falsifiable hypotheses would suffice. However, even if all hypotheses were true, universality would still not be proven since *GTTM* only addresses Western, tonal music. Thus, claims of universality seem to represent a forced conclusion lacking adequate empirical support.

GTTM presupposes an 'experienced listener' although, '[i]n reality no two listeners are exactly alike, nor are any two hearings by the same listener'.[48] Moreover, 'the grammar deals explicitly with only those aspects of heard structure that are hierarchical'.[49] However, the simplified view of the listener and musical structure conflicts with *GTTM*'s universality claims. A potential falsification of a hypothesis generated by *GTTM* could always be explained away by lack of experience on the part of the listener or by influence from non-hierarchical aspects of music listening. This is a severe threat to *GTTM* from a theory-of-science perspective. Also, prescriptive statements about the 'correct' way of listening tend to result from the assumptions of an idealized listener. One questions whether such prescriptions belong in a scientific theory of human cognition.

Furthermore, though Lerdahl and Jackendoff consider their rule system innate, yet, by distinguishing between 'experienced' and 'inattentive' listeners, *GTTM* does seem to recognize the effect that exposure and experience have on music cognition. However, they still exclude that a listener 'is somehow capable of inferring the organisation that the composer, through his compositional method, has consciously built into the piece' and that 'a listener, through experience, acquires serial principles in such a way as to be able to comprehend the serial structure of novel pieces in the idiom'.[50] It is thus unclear whether the rule system is constant or influenced by experience.

The latter is assumed e.g. in Huron's 'ITPRA-theory' according to which our mental representations of music are internalized through 'statistical learning', i.e. repeated exposure to the probabilistic properties of music.[51] However, it is unclear how Lerdahl and Jackendoff define an experienced listener, and it seems paradoxical if, on his way to becoming 'experienced', the listener can only train the principles of hierarchical organization of which most are already assumed to be innate.

45 Anonymous, 'Noam Chomsky', *Encyclopaedia Britannica Online*, www.britannica.com/EBchecked/topic/114218/Noam-Chomsky, accessed 26 June 2009.
46 Rosner, 'Review'.
47 Lerdahl and Jackendoff, *GTTM*, 279.
48 Lerdahl, *Tonal Pitch Space*, 5.
49 Lerdahl and Jackendoff, 'An overview' (1983).
50 Lerdahl and Jackendoff, *GTTM*, 298–89.
51 Huron, *Sweet Anticipation*.

Reception and legacy

The vast number of citations mentioned in the introduction to this paper necessitates a clear decision on sampling criteria prior to delving into this huge data material in an attempt to review *GTTM*'s impact on subsequent scholarship and analytical practice. I will therefore restrain myself to publications where *GTTM* was much more than just a peripheral reference, but played a key role. Moreover, *GTTM* opened several possible paths to pursue in subsequent research. I will structure this discussion according to the following five subcategories:

1. Empirical testing
   ↓
2. Extension and further refinement
   ↓
3. Rule quantification
   ↓
4. Computational implementation
   ↓
5. Computational application for other purposes

In my view, a certain serial order exists where one subcategory naturally leads to the next. This has been indicated in the figure above. The serial order is, however, not strictly imperative. One may begin one's quest from any of the categories, and e.g. one might suppose the perceptual relevance of the rule system, base one's extension on other things than empirical findings, accept the theory as it is and proceed directly to rule quantification, make a computer implementation without rule quantification using instead aleatoric operations, etc. Furthermore, some argue that music theory should remain within the traditional humanistic paradigm, thus circumventing altogether the series outlined above.[52]

Finally, I will distinguish between endeavours of Lerdahl and Jackendoff and those of other researchers.

Later work by Lerdahl and Jackendoff

(1) *Empirical testing*. The authors' own empirical testing of *GTTM* has been limited to tests of tonal tension predicted by Lerdahl's later *Tonal Pitch Space Theory* (see below).[53]

(2) *Extension and further refinement*. Despite the considerable criticism outlined above, Lerdahl and Jackendoff never withdrew parts of their theory. E.g., the

52 Justin London, 'Lerdahl and Jackendoff's Strong Reduction Hypothesis and the Limits of Analytical Description', *In Theory Only*, 13/1 (1997), 3–29; Clarke, 'Theory, Analysis'.

53 Emmanuel Bigand, Richard Parncutt, and Fred Lerdahl, 'Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training', *Perception & Psychophysics*, 58/1 (1996), 124–41; Fred Lerdahl and Carol L. Krumhansl, 'Modeling Tonal Tension', *Music Perception*, 24/4 (2007), 329–66.

tenth-anniversary summary was practically unchanged,[54] and though it seems non-intuitive that passing and neighbour notes are always subordinate to a single note, Lerdahl still argued for strict branching in 1997 proposing a method for conflict solving by calculation.[55]

Jackendoff also extended *GTTM*. Exploring how the rule system works in a listener's mind, he proposed a parallel multiple-analysis model for real-time, mental processing and discussed its advantages in comparison with two serial models.[56] He furthermore claimed this process to take place independently from long-term memory and hypothesized it to be partly responsible for musical affect.[57] Focusing on real-time processing, he probably tried to soften the initial problematic focus on final-state listening. Thus, he generalized the operational sphere of *GTTM* beyond retrospective, hierarchical listening.

One of Lerdahl's subsequent extensions of *GTTM* was a hierarchical organization of timbre by ways of applying prolongational analysis to timbral dimensions.[58] He argued that timbre is often regarded a secondary musical parameter because it is typically organized in an associational manner and not – yet at least – hierarchically like pitch and rhythm.

Moreover, Lerdahl derived grouping, metrical, and prolongational structure from phonology and prosody in a poem by Robert Frost.[59] This analysis was subsequently used for substantiating his hypotheses on commonalities and differences between language and music processing. These hypotheses lend themselves directly to cognitive psychologists and neuroscientists for theory building and empirical testing.

In another study, Lerdahl adopted the composer's perspective by introducing 'compositional' and 'listening grammar'.[60] An average listener is unable to establish a mental representation mirroring the compositional algorithm underlying many serial pieces. Suggesting certain 'cognitive constraints' for compositional systems, he argued that music cognition can inform compositional practice by mending the gap between the two grammars without resorting to tonal nostalgia.

Replacing the 'stability conditions' of the time-span component with 'salience conditions' and adding PRs for atonal prolongation, Lerdahl developed an extension of *GTTM* capable of dealing with atonal music in a hierarchical manner.[61]

54 Lerdahl and Jackendoff, 'An overview' (1993).

55 Fred Lerdahl, 'Issues in Prolongational Theory: A Response to Larson', *Journal of Music Theory*, 41/1 (1997), 141–55.

56 Ray Jackendoff, 'Musical processing and musical affect', in Jones and Holleran, *Cognitive Bases of Musical Communication*, 51–68.

57 Jackendoff argued that musical affect cannot be accounted for by expectancy alone since well-known music being heard in one's head ('musical imagery') also evokes emotions.

58 Fred Lerdahl, 'Timbral hierarchies', *Contemporary Music Review*, 2 (1987), 135–60.

59 Fred Lerdahl, 'The Sounds of Poetry Viewed as Music', *Annals of the New York Academy of Sciences*, 930 (2001), 337–54; Fred Lerdahl, 'Two Ways in Which Music Relates to the World', *Music Theory Spectrum*, 25/2 (2003), 367–73.

60 Lerdahl, 'Cognitive Constraints'; Lerdahl, 'Pitch-Space Journeys'.

61 Fred Lerdahl, 'Atonal Prolongational Structure', *Contemporary Music Review*, 3 (1989), 65–87; Fred Lerdahl, 'Prolonging the inevitable', *Revue Belge de Musicologie / Belgisch Tijdschrift voor Muziekwetenschap*, 52 (1998), 305–9.

Moreover, he illuminated shortcomings of previous attempts to apply Schenkerian and set theory to this repertoire. E.g., the latter relates sets in an associational manner, but ignores relations between individual set members. An alternative to replacing the stability conditions is to define them in further detail. This quest underlies Lerdahl's *Tonal Pitch Space* theory (henceforth *TPS*) containing four components: pitch space, surface tension, attraction models, and prolongational structure.[62]

(3) *Rule quantification*. Applying the 'distance algorithm' from *TPS* to information from *GTTM*'s prolongational reduction, Lerdahl made the sole attempt of quantification on the part of the authors themselves by calculating tonal tension.

(4) *Computational implementation*. Lerdahl and Seward lately took steps towards computer implementation of *GTTM* and *TPS* which was already proposed by Lerdahl in 2001.[63] This endeavour is, however, still in its infancy, and examples of (5) *computational application* are absent altogether from Lerdahl and Jackendoff's work.

In sum, Lerdahl and Jackendoff did indeed work along some of the five proposed lines of development. However, rather than refinement in the form of rule specification and quantification, they worked primarily at extending the theory towards atonal repertoires, the timbral domain, harmony, and poetry.

Contributions by other researchers

(1) *Empirical testing*. As encouraged by Lerdahl and Jackendoff themselves, *GTTM*'s rule system has been tested empirically in a vast number of studies. Bigand alleged to provide evidence in support of expert and non-expert listeners' ability to distinguish prolongational structures from one another.[64] However, Bigand's research methods were questionable because he did not control all factors with sufficient rigour, and it is unclear how his experiment supported Lerdahl and Jackendoff's theory specifically more than it supported a general ability to extract underlying harmonies from melodic contexts – a common tenet of much theory, including Schenker's.

In a later study, Bigand and Parncutt modelled tension using *GTTM*, *TPS* and a sensory-psychoacoustical model by Parncutt himself and compared perceived tension in listeners. Local models seemed to account for listener ratings more accurately than global, hierarchical ones.[65]

Dibben similarly tested perceptual salience of hierarchical structure by comparing the degree of experienced similarity between an original melody and either a correct time-span reduction or an incorrect one where surface events were chosen in prefer-

62 Lerdahl, *Tonal Pitch Space*; preliminary version published as Fred Lerdahl, 'Tonal Pitch Space', *Music Perception*, 5 (1988), 315–50.

63 Fred Lerdahl and Rob Seward, 'Toward a computer implementation of the *GTTM/TPS* analytic system', unpublished manuscript, 2008.

64 Emmanuel Bigand, 'Abstraction of two forms of underlying structure in a tonal melody', *Psychology of Music*, 18 (1990), 45–59.

65 Emmanuel Bigand and Richard Parncutt, 'Perceiving musical tension in long chord sequences', *Psychological Research*, 62 (1999), 235–54.

ence of deep-structure elements.[66] Dibben's results provided support for Lerdahl and Jackendoff's claims about hierarchical organization, but failed to support Lerdahl's hypotheses about hierarchical structures in atonal music. Also findings by Palmer and Krumhansl support the presence of a strong hierarchical component in mental representation for musical metre.[67]

Deliège tested *GTTM*'s rules for grouping structure concluding that, compared with non-musicians, musicians perform segmentation which is more consistent with the theory; partly due to better melodic memory.[68] Apparently, this finding is consistent with *GTTM*'s focus on an 'experienced' listener.

Some empirical support for certain grouping PRs was obtained by Pearce, Müllensiefen and Wiggins who evaluated several statistical and rule-based computational models of grouping by comparing their performance to phrase boundaries in Germanic folk melodies detected by experts.[69] However, alternative models also provided reliable predictions, some of them outperforming those of *GTTM*.

Peretz similarly investigated perceived grouping, more specifically in French folktunes and particularly with respect to parallelism, change in register, and length.[70] Although Lerdahl and Jackendoff claimed the opposite, it is indeed likely that memory of lyrics predisposed for specific melodic parsing due to familiarity of this repertoire to the vast majority of subjects.

Conversely, Clarke and Krumhansl focused on repertoire without lyrics, asking subjects to segment Stockhausen's *Klavierstück IX* and Mozart's *Fantasie in C-minor* (KV275).[71] Most segmentation criteria reported by subjects were in fact consistent with *GTTM*'s PRs, namely as to proximity, change, and parallelism.

Palmer and Krumhansl found melodic phrase judgments of a J. S. Bach fugue theme to be based on mutually independent pitch and temporal information.[72] Thus, they concluded that Lerdahl and Jackendoff's time-span reduction could be considered a relatively good predictor for phrase judgements.

Dodson's study from 2002 represents a border case between subcategory (1) and (2), reacting towards both the criticism of *GTTM*'s focus on scores and the increas-

66 Nicola Dibben, 'The cognitive reality of hierarchic structure in tonal and atonal music', *Music Perception*, 12/1 (1994), 1–25.

67 Caroline Palmer and Carol L. Krumhansl, 'Mental representations for musical meter', *Journal of Experimental Psychology: Human Perception and Performance*, 16 (1990), 728–41.

68 Irène Deliège, 'Grouping conditions in listening to music: An approach to Lerdahl and Jackendoff's grouping preference rules', *Music Perception,* 4 (1987), 325–60.

69 Marcus T. Pearce, Daniel Müllensiefen, and Gerraint A. Wiggins, 'A Comparison of Statistical and Rule-Based Models of Melodic Segmentation', in *ISMIR: Proceedings of the Ninth International Conference on Music Information retrieval* (Drexel University, Philadelphia, 2008), 89–94.

70 Isabelle Peretz, 'Clustering in music: An appraisal of task factors', *International Journal of Psychology*, 24 (1989), 157–78.

71 Eric F. Clarke and Carol L. Krumhansl, 'Perceiving musical time', *Music Perception*, 7 (1990), 213–51.

72 Caroline Palmer and Carol L. Krumhansl, 'Independent temporal and pitch structures in determination of musical phrases', *Journal of Experimental Psychology: Human Perception and Performance*, 13 (1987), 116–26.

ing influence of performance analysis.[73] Based on empirical data from quantitative performance analysis, he extended *GTTM*'s tripartite view on accents by adding a category of performance-controlled 'phenomenal micro-accents' comprising both 'dynamic' and 'agogic micro-accents'. Hypermetrical contraction and completion were thus formalized as TRs.

(2) *Extension and further refinement*. Deliège's findings provided support for her proposed concept of 'postponed segmentation' supposing segmentation to take place after – and not before – a new duration or articulation is introduced.[74] This modification did, however, not apply to acoustic changes in register, dynamics, or timbre. Furthermore, she suggested a few new grouping PRs – e.g., 'segmentation in relation to change in harmony'. This PR emphasizes structural instead of phenomenal accents, and it was probably excluded from *GTTM* because the authors wanted to avoid issues of harmonic stability in considerations on grouping structure.

Based on their own empirical tests of perceived structural boundaries in Western popular music, Bruderer, McKinney, and Kohlrausch similarly proposed new grouping PRs for timbre, tempo, and rhythm changes.[75] In the same issue of *Musicae Scientiae*, Lartillot addressed *GTTM*'s insufficient specification of associational structures by presenting a promising rule-based formalization of motivic parallelism which is allegedly suitable for computational implementation.[76]

In their refined PR system for analysing metrical structure and harmony, Temperley and Sleator modified a metrical WFR calling for an isochronous tactus level into a PR to account for metrical changes, recitativo style, fermatas, etc.[77] Marsden similarly introduced a representational framework improving certain shortcomings of *GTTM* and Schenkerian analysis.[78]

London and Clarke both criticized the binary logic of *GTTM*'s transitive subordination and claims of absolute recursivity implied by its Reduction Hypothesis.[79] They suggested that PRs should instead be rated differently on different hierarchical levels because different principles are salient on various levels. This would allow non-recursive groupings that may interlock between hierarchical levels. London proposed, instead, a 'Weak Reduction Hypothesis' reflecting the ambiguities of music listening.

73 Alan Dodson, 'Performance and hypermetric transformation: An extension of the Lerdahl-Jackendoff Theory', *Music Theory Online*, 8/1 (2002).

74 Deliège, 'Grouping conditions'.

75 Michael J. Bruderer, Martin F. McKinney, and Armin Kohlrausch, 'The Perception of Structural Boundaries in polyphonic representations of Western popular music', *Musicae Scientiae*, Discussion Forum 5 (2010), 273–313.

76 Olivier Lartillot, 'Reflections Towards a Generative Theory of Musical Parallelism', *Musicae Scientiae*, Discussion Forum 5 (2010), 195–229.

77 David Temperley and Daniel Sleator, 'Modeling Meter and Harmony: A Preference Rule Approach', *Computer Music Journal*, 23/1 (1999), 10–27.

78 Alan Marsden, 'Generative Structural Representation of Tonal Music', *Journal of New Music Research*, 34/4 (2005), 409–28.

79 London, 'Lerdahl and Jackendoff's'; Clarke, 'Theory, analysis'.

(3) *Rule quantification*. Systematic quantification of four individual PRs was attempted by Frankland and Cohen.[80] They found that parsing choices could nearly be explained exhaustively by the PRs for attack point and length, whereas rules for slur/rest and register only had marginal influence. Additionally, they suggested modifications of some rules, formalizing e.g. the 'postponed segmentation' previously observed by Deliège who addressed the relative strength of rules by creating systematic conflicts between pairs of rules in her stimuli.[81] Nevertheless, these results only showed preliminary tendencies and most certainly called for further investigation.

(4) *Computational implementation*. Various computational models of *GTTM* have been implemented. Nord made such an attempt in his Ph.D. thesis,[82] but did not, however, cover the complete rule system, using a too straightforward and over-simplified approach that repeatedly transformed 'preference' into 'necessity'.[83] Baker implemented certain *GTTM* rules for grouping structure and time-span reduction drawing on research from the field of Artificial Intelligence and some competing theories to *GTTM*.[84] A syntactic processing algorithm called *Automated Grouping Analysis System* (AGA) and a knowledge-based recognition algorithm termed *Grouping Analyser with Frames* (GRAF) were proposed. Finally, in his book on music cognition Temperley presented the *Melisma Music Analyzer*.[85] Based on PRs, this software analyses six components: metre, melodic phrase, counterpoint, pitch spelling, harmony, and key. Unlike later models developed by a Japanese research team (see below), the weight of parameters was here fixed rather than adjustable.

(5) *Computational application*. Some computational applications of *GTTM* have served the purpose of modelling expressive performance. Todd used the time-span component to model expressive timing,[86] and Arcos and Mantaras developed a system capable of generating expressive performance.[87] The strategy was to imitate human performance integrating, amongst others, knowledge from Narmour's *Implication-Realization Model* and metrical structure, time-span, and prolongational reduction from *GTTM*.

80 Bradley W. Frankland and Annabel J. Cohen, 'Parsing of melody: Quantification and testing the local grouping rules of Lerdahl & Jackendoff's (1983) "Generative theory of tonal music"', *Music Perception,* 21/4 (2004), 499–543.

81 Deliége, 'Grouping conditions'.

82 Timothy Arlan Nord, 'Toward Theoretical Verification: Developing a Computer Model of Lerdahl and Jackendoff's Generative Theory of Tonal Music', Ph.D. thesis (University of Wisconsin, Madison, 1992).

83 According to Hirata et al., 'Techniques for implementing'.

84 Michael J. Baker, 'An artificial intelligence approach to musical grouping analysis', *Contemporary Music Review*, 3/1 (1989), 43–68.

85 David Temperley, *The Cognition of Basic Musical Structures* (Cambridge: MIT Press). Software is available at www.link.cs.cmu.edu/cbms/.

86 Neil Todd, 'A Model of expressive timing in tonal music', *Music Perception*, 3/1 (1985), 33–57.

87 Josep Lluís Arcos and Ramon López de Mantaras, 'Combining AI Techniques to Perform Expressive Music by Imitation', *AAAI Workshop: Artificial Intelligence and Music* (California: AAAI Press, 2000), 41–47.

Likewise, Widmer proposed a rule-based, performance-rendering system applying rules to structural information obtained from *GTTM*'s grouping and metrical analysis and time-span reduction.[88] This information was also balanced by Narmour's model to account for surface phenomena. In conclusion, the author found such surface structure to be somewhat more decisive to expressive performance than the deeper structures referred to by *GTTM*.

Japan – an Asian hot spot of generative music theory

In recent years an extensive research programme related to *GTTM* has arisen in Japan. Because these publications span over several subcategories, I will treat them jointly and, as far as possible, in chronological order.

Rather than starting with (1) *empirical tests* of Lerdahl and Jackendoff's original theory, members of the Japanese research group have used experimental results for validating their own extensions and computational implementations of *GTTM*. The first challenge that they addressed was to (2) *extend GTTM*'s applicability to polyphonic textures. Initially, they developed a strategy for polyphonic grouping by means of Voronoi-diagrams.[89] Subsequently, steps were taken towards a polyphonic time-span component, and as (5) *application* music-summarization software was produced along the lines of Information Technology by removing excerpts with high degrees of time-span similarity.[90] Related to this, data from time-span analysis were used as annotation to sound files, thus improving the prospects of retrieval, reproduction, and sharing of music.[91] Furthermore, they developed a method of creating intermediary melodies between two different melodies[92] and constructed software making alternative arrangements of piano pieces, similarly with the aid of time-span trees.[93]

In two cases *GTTM* was used for modelling expressive performance and music cognition. In one case time-span reduction was used to model expressive interpretations of computer-performed musical scores,[94] and in the model of melodic expectancy developed by Hamanaka and colleagues the next note in a melody was

88 Gerhard Widmer, 'Modeling the Rational Basis of Musical Expression', *Computer Music Journal*, 19/2 (1995), 76–96; Gerhard Widmer, 'Learning Expressive Performance: The Structure-Level Approach', *Journal of New Music Research*, 25 (1996), 179–205.

89 Masatoshi Hamanaka and Keiji Hirata, 'Applying Voronoi-diagrams in the automatic grouping of polyphony', *Information Technology Letters*, 1/1 (2002), 101–2.

90 Keiji Hirata and Shu Matsuda, 'Interactive music summarization based on Generative Theory of Tonal Music', *Journal of New Music Research*, 32/2 (2003), 165–77.

91 Keiji Hirata, Shu Matsuda, Katsuhiko Kaji, and Katashi Nagao, 'Annotated music for retrieval, reproduction, and sharing', in *Proceedings of International Computer Music Conference* (2004), 584–7.

92 Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, 'Melody expectation method based on *GTTM* and TPS', in *Proceedings of ISMIR* (2008), 107–12.

93 Keiji Hirata and Tatsuya Aoyagi, 'Computational Music Representation Based on the Generative Theory of Tonal Music and the Deductive Object-Oriented Database', *Computer Music Journal*, 27/3 (2003), 73–89.

94 Keiji Hirata and Rumi Hiraga, 'Ha-Hi-Hun plays Chopin's Etude', in *Working Notes of IJCAI-03 Workshop on Methods for Automatic Music Performance and their Applications in a Public Rendering Contest* (2003), 72–73.

predicted.[95] Contrary to many statistical models working on the musical surface, this model also takes deep structure into account by using information on melodic stability derived from *GTTM*.

Moreover, the Japanese research group contributed significantly to (4) *computational implementation*. Hamanaka and colleagues constructed two algorithms assigning grouping and metrical structure to monophonic music with adjustable parameters for relative strength of PRs.[96] These were later adopted as part of the *Automatic Time-Span Analyzer* (ATTA) software capable of performing grouping, metrical and time-span analyses of monophonic music using adjustable variables and a MusicXML-file as input.[97]

Later the ATTA was extended to a full-automatic version (FATTA) where optimal parameters were set automatically using a feedback loop that even integrates some of the interdependent PRs.[98] FATTA represented a first step towards a more dynamic implementation of GTTM overcoming the limitations of Lerdahl and Jackendoff's stable, rule-based approach. However, despite correlation with empirical data, the cognitive validity of FATTA's feedback mechanism in consecutive, real-time listening seems relatively doubtful.

(2) *Extension* and (3) *rule quantification*. Both the ATTA and the FATTA were based on an extended version of *GTTM* ('ex*GTTM*') distinguished by '[a] implementing new parameters for resolving rule conflicts, supplementing implicit/lacking concepts, and developing a working algorithm (especially an algorithm for acquiring hierarchy); [b] adding full externalisation and parameterisation, [c] coping with restrictions in implementing GTTM, and finally [d] aiming at generating as many correct results for humans as possible'.[99]

ATTA and FATTA also distinguish themselves by taking deep structures into account. However, the implementations of *GTTM* presented by Hamanaka and colleagues only handled monophony, disregarded harmony, ignored some PRs, did not implement prolongational reduction, and established no feedback-loop from time-span reduction back to grouping and metrical analysis.

In conclusion, neither Western nor Japanese researchers have adhered strictly to the serial pattern suggested by the five subcategories, but rather worked in parallel sometimes starting elsewhere than from empirical testing and occasionally moving counter to the serial order. As for the Japanese research group, this has led to advanced computational applications which may correlate with – but have not arisen

95 Hamanaka et al., 'Melody expectation'.

96 Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, 'Automatic Generation of Grouping Structure Based on The *GTTM*', in *Proceedings of ICMC* (2004), 141–44; Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, 'Automatic Generation of Metrical Structure Based On *GTTM*', in *Proceedings of ICMC* (2005), 53–56.

97 Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, 'Implementing "A Generative Theory of Tonal Music"', *Journal of New Music Research*, 35/4 (2007), 249–77. Software accessible at http://staff.aist.go.jp/m.hamanaka/atta/.

98 Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, 'FATTA: Full Automatic Time-span Tree Analyzer', in *Proceedings of ICMC* (2007), 153–56.

99 Hirata et al., 'Techniques for implementing'.

from – empirical data and fail to take all PRs and the full interaction between them into account. Only very few Westerners have spanned across more subcategories and have thus not always felt obliged to draw the full consequences of findings obtained by their peers. Noticeably, the prolongational component has been lacking in all computational models, and thus far no one has managed to provide a complete externalization and parameterization integrating all WFRs, TRs and PRs. Furthermore, empirical findings point towards the importance of surface structure and non-hierarchical properties of real-time listening which ultimately may render such a quest irrelevant. The complexity and ambiguities of *GTTM* have simply made it impossible to exhaust any single of the five subcategories. Hence, all endeavours, particularly into the last few subcategories, involve the dangerous risk that one might be theorizing on fallible assumptions.

Influence of GTTM on music analysis and the theory curriculum

Eventually, *GTTM*'s impact on teaching music analysis and on the theory curriculum will be discussed. Owing to the obvious similarities between *GTTM* and Schenkerian analysis, the camp of Schenkerians would certainly not be the worst place to look for signs of *GTTM* influence. I will only outline a few possible directions here and leave it up to others to explore this connection in further depth. Recently, Lerdahl expressed that he considers Schenker's analytical system as a 'proto-generative theory' in the sense that music is viewed as hierarchical elaboration and transformation of the underlying *Ursatz* where 'the same elaborative and transformational principles apply recursively at all levels'.[100] Like Schenker, *GTTM* considers structural significance a matter of syntax and not of surface salience, and *GTTM*'s prolongational reduction is probably the hierarchical component most comparable to Schenkerian reductions in conceptual and notational terms.[101]

Interestingly, the limitations brought about by the presumed hierarchical listening make the critical reader query whether *GTTM*'s reductions elaborate musical structure beyond Schenkerian theory. If not, then *GTTM* is nothing but a sheer methodological specification. However, by refraining from quantification of their PRs, Lerdahl and Jackendoff still fall short of justifying *GTTM*'s superiority to traditional Schenkerian analysis.

Despite all similarities, *GTTM* also differs from Schenkerian theory in various ways. E.g., metrical structure, time-span reduction, and to some extent grouping structure add a temporal dimension absent from traditional Schenkerian theory. Allen Forte already pointed towards this shortcoming as an 'unsolved problem in music theory'.[102] Understood as a methodological specification of Schenkerian analysis, *GTTM* did contribute to the solving hereof. However, *GTTM* is only rarely quoted by Schenkerians, and its influence on Schenkerian theory seems to have been only peripheral and momentary. Another reason for the incompatibility between

100 Fred Lerdahl, 'Genesis and Architecture of the *GTTM* Project', *Music Perception*, 26/3 (2009), 187–94.
101 Lerdahl and Jackendoff, *GTTM*, 231.
102 Allen Forte, 'Schenker's Conception of Musical Structure', *Journal of Music Theory*, 3/1 (1959), 1–30.

Schenkerian theory and *GTTM* is arguably the fact that the former seems to remain persistently at the 'poietic' and 'neutral' levels whereas, conversely, the latter defined itself unambiguously as an 'esthesic' practice.

In sum, *GTTM* has not been widely acknowledged by analysts and is hardly used as a methodological analysis textbook in humanistic musicology. *GTTM*'s position among German and Finnish scholars and analysis teachers illustrates this perfectly: About a decade ago, Cornelius Bradter declared the *GTTM* project for dying,[103] two years later Heikki Valkonen upheld the death sentence,[104] and recently Wolfgang Just signed the death certificate based on similar arguments to the ones put forth in the initial reviews.[105]

There are, however, also other possible reasons for the limited educational applicability of *GTTM*. First of all, it was not intended as a textbook, and there are no exercises and practically no instructions in how to apply the theory in actual analysis. The authors used relatively few musical examples (primarily themes by Mozart, Beethoven, and J. S. Bach), and no single analysis of a complete piece occurs. Even in earlier versions of the theory and in later summaries by other authors these excerpts still recur as the only ones in use.[106]

Calling for four interdependent, simultaneous approaches, *GTTM* reaches immense complexity, making the analytical process extremely time-consuming.[107] E.g., in time-span analysis one ideally needs to have not only the metrical and grouping structure in mind, but also the interdependent prolongational reduction. Thus, no logical order of handling the four components exists. Also, due to visually similar notation forms and considerable interdependence, there is a risk that novices will mix up the two reduction types.

As pointed out by Hamanaka and colleagues, even a ten-note melody provides millions of possible time-span trees.[108] It would be unrealistic to imagine an analyst calculating all possible, well-formed trees and then evaluating the interaction of PRs for each of them. Rather, she would base her analysis on musical intuition. Then *GTTM* would indeed account for this intuition by circular reasoning, but it would not provide a practically applicable analytical system. Since *GTTM* basically formalizes musical intuition, but tends towards complexity, using intuition in itself usually leads to similar results. Moreover, it takes considerable amounts of intuition just to administer the PRs. That is, even if the analyst was highly familiar with the rules of *GTTM*, he would still rely heavily on musical intuition which was exactly what the

103 Cornelius Bradter, *Die generative Theorie der tonalen Musik. Grundlagen und Entwicklungsimpulse durch F. Lerdahl und R. Jackendoff* (Beiträge zur Musikpsychologie, 2; Münster: LIT Verlag, 1998).

104 Heikki Valkonen, 'Lerdahl and Jackendoff Revisited – *A Generative Theory of Tonal Music*', University of Jyväskylä, www.cc.jyu.fi/*heivalko/articles/lehr_jack.htm (2000), no longer accessible online.

105 Wolfgang Just, *Die generative Theorie tonaler Musik nach Lerdahl und Jackendoff – Darstellung und Kritik* (Darmstadt: GRIN Verlag, 2007).

106 Even I declare myself guilty of this by including Figure 1 as an initial illustration in this paper.

107 Clarke, 'Theory, analysis'.

108 Hamanaka et al., 'ATTA: Automatic'.

theory intended to objectify and formalize. Furthermore, it has been shown that *GTTM* is based on a simplified view on music, that in some respects it is unnecessarily formalistic whereas in others it is not explicit enough. Finally, its claims of universality, presumptions about the listener, and preoccupation with global, hierarchical listening seem to disregard other cognitive theories and empirical findings.

Additionally, developments in contemporary cognition research pose certain limitations to the relevance and future applicability of *GTTM*'s rule-based approach. In cognitive modelling, rule-based models have largely been 'out-ruled' by models acquiring knowledge through unsupervised statistical learning. Rens Bod already challenged rule-based segmentation models by showing that listeners' grouping analysis of 1,000 songs from the Essen Folksong Collection corresponded to occurrence-frequencies of motives in the general repertoire even though some motives deviated strongly from the Gestalt principles of proximity, similarity, and parallelism.[109] Similarly, in the earlier mentioned study by Pearce, Müllensiefen and Wiggins, a computational model was presented that placed grouping boundaries before unexpected notes in a melody, once transitional probabilities had been internalized from a given training corpus.[110] This unsupervised model performed remarkably well even though it integrated no predefined music-theoretical rules. Importantly, probabilistic models of music cognition allow researchers to account for cultural differences, which Lerdahl and Jackendoff were rather ambiguous about. Thus, the prominence of statistical learning in contemporary cognition research might ultimately make a dead-end of further attempts of quantifying *GTTM* for computational purposes.

Despite all this, from a historical perspective, Lerdahl and Jackendoff's *A Generative Theory of Tonal Music* did leave significant, though somewhat indirect, imprints on the discipline by playing an important role in the introduction of the cognitive paradigm assigning new meaning to the concept of rules in music theory and placing listening grammar (i.e. the 'esthesic level') in a hitherto unseen key position. Similarly, a closer – and certainly long-lasting – link was established between music theory and empirical psychology, which has encouraged empirical research with music-theoretical implications and all in all implies great potential for future research.

Interestingly, besides from two review replies and Lerdahl's *Tonal Pitch Space* theory, Lerdahl and Jackendoff never published in the traditional music theory journals again (*Journal of Music Theory*, *The Musical Quarterly*, *Perspectives of New Music*), although they had done so prior to 1983. Instead they turned towards the new-established *Contemporary Music Review*, the similarly new and empirically oriented *Music Perception* and dedicated themselves to writing book chapters. To the extent that *GTTM* is actually referred to in present-day theory, it is primarily used to justify simple claims about hierarchical organization of musical structure rather than unfolding its detailed rule system. Thus, although Lerdahl and Jackendoff had no monopoly on such ideas – neither in music theory nor in general – the cognitive

109 Rens Bod, 'Memory-Based Models of Melodic Analysis: Challenging the Gestalt Principles', *Journal of New Music Research*, 30/3 (2001), 27–37.

110 Pearce et al., 'A Comparison'.

paradigm was indeed promoted by the publication of *GTTM*, and this seems to have spurred a significant relocation of the academic 'battleground' for generative music theory away from traditional theory in the direction of music cognition, psychology, empirical research, computational modelling, and cognitive neuroscience. In these years where music cognition research is gaining an increasingly steady foothold within the musicological sphere – in Denmark as well as abroad – Fred Lerdahl and Ray Jackendoff's *A Generative Theory of Tonal Music* is similarly likely to achieve a more pronounced position than hitherto within the Danish music theory canon.

Summary

Fred Lerdahl and Ray Jackendoff's *A Generative Theory of Tonal Music* (*GTTM*) has only received limited attention in Danish music theory. Yet, its influence is irrefutable in terms of introducing the 'cognitive paradigm', which changed analytical focus from musical structure to the listening process. Recently, music cognition research has gained territory in Denmark, thus warranting a re-assessment of GTTM and its legacy.

This paper provides an overview of GTTM outlining typical points of criticism. These include a simplified view on music, an unresolved conflict between global and local listening, an occasionally underspecified rule system, and unsubstantiated claims of universality and innateness based on intuition rather than cross-cultural research. GTTM's reception and legacy is discussed in terms of 1) empirical testing, 2) theoretical refinement, 3) rule quantification, 4) computational implementation, and 5) application. Empirical findings have repeatedly emphasized the significance of surface structure and non-hierarchical, real-time listening, and models acquiring knowledge through unsupervised, statistical learning have largely replaced rule-based ones in cognitive modelling. This allows researchers to account for cultural differences, which Lerdahl and Jackendoff were strongly ambiguous about. Moreover, *GTTM* has not been widely acknowledged by analysts, is hardly included in the theory curriculum, and is primarily cited by present-day theorists to justify simple claims about hierarchical organization. Nevertheless, GTTM was instrumental in establishing a link between music theory and psychology, which has encouraged empirical research with music-theoretical implications within the fields of music cognition, experimental psychology, computational modelling, and cognitive neuroscience.