Beyond the Regular: A Formalization of Non-Isochronous Metrical Structure

MSc Thesis *(Afstudeerscriptie)*

written by

Tom Hendriks
(born September 28th, 1992 in Nijmegen, The Netherlands)

under the supervision of **Bastiaan van der Weij MSc** and **Prof Dr Henkjan Honing**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:
December 20th, 2016

Members of the Thesis Committee:
Bastiaan van der Weij MSc
Prof Dr Henkjan Honing
Dr Maria Aloni
Dr Raquel Fernandez
Dr Makiko Sadakata
Prof Dr Benedikt Löwe (chair)

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

Meter perception is the process of inferring metrical structure, a hierarchical and regular mental framework of beats, from an auditory signal. Research on meter perception revolves around the question how listeners perform this task. Formal theories provide a complete abstract representation of metrical structure. Thereby, they provide a clear overview of its key properties and can be implemented into cognitive models, which may in turn clarify the cognitive processes behind meter perception.

Most existing formal theories of metrical structure disregard patterns that contain non-isochronous (unequally spaced) beats, while many musical pieces from non-Western musical cultures, such as Balkan or African, induce such meters. Recently, London (2012) proposed a theory of metrical structure that is grounded in empirical perceptual studies and cross-cultural, by incorporating non-isochronous metrical structure. However, this theory is not formal: it is not completely and unambiguously specified.

This thesis is concerned with formalization of the theory of London. Formalization reveals multiple ambiguities and inconsistencies within the theory of London, which are evaluated by defining additional rules to the formalization and analyzing the effect of these rules on the space of possible meters. In future research, the proposed formalization may be implemented into a cognitive model of meter perception. Such a cognitive model may provide insight into the cognitive processes behind meter perception within a cross-cultural paradigm.

Acknowledgments

First and foremost, I would like to thank Bastiaan van der Weij for supervising my trip on the somewhat bumpy road of completing this thesis. Without you, I would not have been able to make as much out of it as is now laying before us. Thank you very much for our meetings, our discussions and your helpful feedback on all those drafts. Also, I would like to thank Henkjan Honing for his background supervision and especially the road block that made me steer back to the highway, including some very instructive direction signs along the way. A grateful side note goes out to my academic mentor Floris Roelofsen for scarce but meaningful meetings.

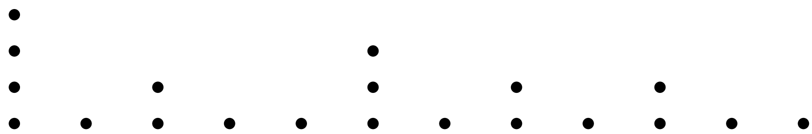
Furthermore, I would like to thank my thesis committee for taking the time to read this thesis and attending my defense. It is an honor!

Another grateful acknowledgment goes out to Justin London, who helped me a great deal to understand the spirit of his work through his comprehensive and witty response to my e-mail with numerous nosy questions. You truly inspired me to get the most out of the thesis.

I would also like to thank Eva van Boxtel and Yfke Dulek for their helpful comments regarding the formal mathematical parts of this thesis. Without your help, Chapter 4 would still be a complete maze. Of course, any remaining errors are my responsibility only.

Also many thanks to the entire Music Cognition Group for their feedback, discussions and good times. In general, I would like to thank everyone involved in providing me with this great subject. Thanks to this thesis, I now know how I should call these weird meters I played with my former progressive rock band Sane (3-2, 2-2-3, 3-3|3-2...). Conversely, this also means that I had a little bit of a 'predisposition' for appreciating non-isochronous meters. The group improvisation workshops by the wonderful Nicoline Snaas added to this appreciation even more ("Oh... look at that!"). More importantly, these workshops helped me find a little bit of peace when I needed it most.

Last, but most certainly not least, I would like to thank my friends and family for their love and support through my bumpy road, which, especially during these last months, has seemed more like a rock I lived under. I want to thank Quico in particular for supporting this hermitage and for pulling me out of it when needed, back into the real world or relaxation mode. You have been more helpful and supporting than you can imagine!



"I know the pie- ces fit"

— Tool, *Schism* (2001)

Contents

Chapter 1

Introduction

Any person without sensory or cognitive disabilities can 'feel the rhythm' of many musical pieces (Patel, 2008; Honing, 2012). This feeling is related to an underlying regular pattern that is called *metrical structure* by musicologists and cognitive scientists. One of the core questions within research on cognition and perception of rhythm is how listeners can 'feel' this metrical structure. This question has led to many different theories of *meter perception*: the way in which a particular metrical structure (a *meter*) is induced in the listener upon hearing a sound signal. The 1950s witnessed the rise of *formal* theories of cognition in general (e.g., Chomsky, 1956, 1957), followed in the 1970s by formal theories on meter perception in particular (Longuet-Higgins, 1978).[1] A formal theory of metrical structure provides a complete, abstract representation of metrical structure and specifies rules that unambiguously decide which patterns can serve as a meter. Formal theories have many advantages; they provide a clear overview of key properties of metrical structure and they can be implemented into cognitive models of meter perception. In turn, cognitive models can visualize and clarify the cognitive processes behind meter perception.

Due to their focus on Western classical music, many theories on meter perception require that all pulses (or beats) in metrical structure are *isochronous* (i.e., equally spaced). However, many musical pieces from non-Western musical cultures, such as Balkan or African, induce a metrical structure that contains non-isochronous beats as well. Most existing formal theories do not account for these meters, while some current theories that do account for them, are not formal. A cognitive theory of metrical structure should be: (1) as simple as possible, while capturing all relevant details; (2) formal, that is, completely and unambiguously specified; (3) grounded in empirical perceptual studies and cross-cultural (and therefore incorporating non-isochronous metrical structure[2]). Most existing formal theories of metrical structure do not meet requirement (3) by disregarding non-isochronous metrical structure. Recently, London (2012) proposed a theory of metrical structure that does meet requirements (1) and (3): it is grounded in empirical research on perception and incorporates non-isochronous metrical structure within a theory that is simple with respect to the subject matter. However, the approach of London is not formal: it leaves some details unspecified and thereby does not meet requirement (2). In order to meet all three criteria, this thesis proposes a formalization of London's theory. This formalization differs from related work by providing an analysis of aspects in London's theory that are underspecified. In the current project, we provide a direct formalization of London's well-formedness constraints. Then, we show how this formalization generates meters that are not well-formed according to London's theory. We address the problems of underspecification

[1]Indeed, it was Longuet-Higgins (1973) who coined the term *cognitive science* as a collective name for the interdisciplinary field of sciences that is most likely to be enriched by artificial intelligence studies; see also Pearce and Rohrmeier (2012).

[2]Empirical studies prove that non-isochronous metrical structure is indeed perceivable (e.g., Hannon & Trehub, 2005a, 2005b; Hannon, Soley, & Ullal, 2012; Kalender, Trehub, & Schellenberg, 2013).

and inconsistency by proposing additional constraints and analyzing their effect on the set of permissible meters. Before formalization, we provide an overview of literature, as well as an overview of aspects that have been left open by London's proposal and which need to be resolved in order to construct a formalization.

Before theories of meter perception can be discussed, we need to know what meter actually is. Section 1.1 presents different views on the definition of meter. Subsequently, Section 1.2 provides an overview of the current thesis.

1.1 What is meter?

The definition of meter is a starting point for many books and articles about meter perception and it is not trivial; researchers wield different meter definitions that do not always agree. This section discusses some definitions of meter in literature, which properties of meter they generally agree about and in which they differ.

In most literature, meter is defined as a mental framework of beats or pulses that is, to some extent, hierarchical and regular. The notion of a mental framework refers to the idea that meter is a form of structure in the mind of the listener (instead of the musical score, composer or performer) that guides perception of rhythmic stimuli. The idea of meter as a mental phenomenon is generally shared among recent literature on meter perception (most explicitly in Clarke, 1999; Honing, 2012, 2013; more implicitly in Lerdahl & Jackendoff, 1983; Povel & Essens, 1985; Palmer & Krumhansl, 1990; Temperley, 2007, 2013; London, 2012; Vuust & Witek, 2014). It relates to the often discussed distinction between meter and rhythm: while rhythm refers to the actual sound signal, meter is induced in, or constructed by, the listener (e.g., Clarke, 1999; Honing, 2013).

Ideas of hierarchy and regularity mainly originate from the first formal theories on meter perception by Longuet-Higgins (1976, 1978) and Lerdahl and Jackendoff (1983). Lerdahl and Jackendoff (1983) define meter as "a regular pattern of strong and weak beats" (p. 12) that is inferred by the listener upon hearing a sound signal. Respective theories propose that hierarchy and regularity help listeners to actually use meter as a framework; when the listener perceives some beats as stronger than others, they can predict this regular alternation and use this abstract, mental knowledge to make sense of unfamiliar melodies (Palmer & Krumhansl, 1990).

The exact form of the metrical framework differs among different researchers. Longuet-Higgins (1976), Lerdahl and Jackendoff (1983) and Palmer and Krumhansl (1990) propose multi-leveled hierarchies of pulses that are fully regular, while in Povel and Essens (1985), this mental framework resembles a single isochronous sequence of beats ('clock') that is subdivided only one level below, in a less regular fashion. Here, the respective theories and models come into play. Depending on these theories, meter definitions of different researchers are sometimes more strongly hierarchical and regular, and sometimes less.

Lastly, it is a point of discussion whether the metrical framework is discrete. Desain (1992) argues that meter is not a symbolic notion, but should be seen as a continuous structure of attentional peaks, that is, an *expectancy curve*. Likewise, Large and Kolen (1994) propose that meter perception is a dynamic *entrainment process* that builds upon resonance, as derived from ideas on *dynamic attending*. This theory proposes that perception, attention and memory are inherently rhythmical; these processes entrain to music and generate 'anticipatory pulses of attention' (Large & Kolen, 1994; Jones & Boltz, 1989). Large and Kolen (1994) deliberately present no definition of meter that is revised to their theories, as they identify unresolved issues in their model, so it cannot yet be linked to a theory of meter. Later work does introduce definitions of meter within this paradigm. For instance, Vuust and Witek (2014) regard meter as the result of the interplay between external periodicities and internal attending processes.

Meter definitions using entrainment and attending are often associated with models involving continuity, subsymbolism and flexibility (e.g., Desain, 1992; Large & Kolen, 1994; Vuust &

Witek, 2014). As it seems natural to think of a mental framework as a discrete structure (for instance in the form of a tree, line or grid), the ‘entrainment definitions’ might seem incompatible with the idea of meter as a mental framework. But this is not necessarily the case: ‘entrainment definitions’ of meter provide information on the *cognitive mechanisms behind* meter, rather than defining *what meter is*. The mental framework of meter might not be discrete, but it can still be seen as a framework that guides rhythm perception. Even in this continuous view, the other discussed properties of meter still hold as well: meter is still a mental construct, entrainment involves oscillations and therefore periodicity (a less strict form of regularity), and when there are multiple oscillations, a hierarchy emerges of stronger and weaker peaks. Moreover, in resulting cognitive models, continuous curves can be abstracted to discrete frameworks. The theory of London (2012) is an example of this. London defines meter as a form of entrainment behavior that can be learned and allows listeners to synchronize their perception to rhythms they hear. However, London argues that the corresponding attentional peaks map to time points. Thereby, London designs a representation of meter that is explicitly discrete, even though he takes ideas of entrainment and attending into account.

In summary, the generally agreed definition of meter is a mental framework of beats, which is, to some extent, hierarchical and regular, and may or may not be discrete. Among different books and articles, there are nuances in this definition in terms of hierarchy, regularity and discreteness.

1.2 Overview

This thesis is set out as follows. Chapter 2 provides an overview of relevant theories and models of metrical structure and argues that a cognitive model of metrical structure should be hierarchical and incorporate non-isochronous meters. It concludes with a discussion of the theory of London (2012), which incorporates both these requirements, followed by the proposal of formalization. Chapter 3 discusses related work in formalization of metrical structure and evaluates the differences between these formalizations and the current formalization. Chapter 4 introduces the formalization, along with a derivation of its agreement with the well-formedness rules as defined in London. The chapter also introduces additional rules to address inconsistencies within London’s theory. Chapter 5 presents an analysis of the formalization, along with a discussion of the space of possible meters, and the effect of the proposed additional rules on this space. Chapter 6 contains a conclusion and suggestions for further research, in which the formalization may be implemented in a cognitive model that can be tested empirically.

Chapter 2

Theories and models of metrical structure

Music theoretic accounts of meter date back into at least the eighteenth century (for an overview, see Mirka, 2009). This thesis focuses on studies in the more recent cognitive tradition, that is, from the late 1960s onward. In these last 50 years, many different theories of meter perception have been proposed. Some theories are tested with models or experiments, while others only propose a metrical representation that could be implemented in a model.

This chapter discusses a selection of prominent theories and models of meter perception. The chapter is divided into four sections. Section 2.1 contains a discussion of theories and models that are strongly hierarchical. Section 2.2 presents a weakly hierarchical model. Section 2.3 discusses evidence for strongly hierarchical representations. Section 2.4 discusses the theory of non-isochronous metrical structure by London (2012) and concludes with the current proposal of formalization. The previous chapter discussed nuances in meter definition with regard to hierarchy, regularity and discreteness. These factors recur in the theories and models discussed in the current chapter, and they will occasionally be used to differentiate between the theories and models.[1] Note that theories of *meter perception* are often built on theories of *beat induction*, which is the perception of a single recurring beat instead of a hierarchy of beats (for accounts of beat induction, see for instance Povel & Essens, 1985; Desain & Honing, 1999; Todd, O'Boyle, & Lee, 1999; Patel, 2008; Honing, 2012; Todd & Lee, 2015).

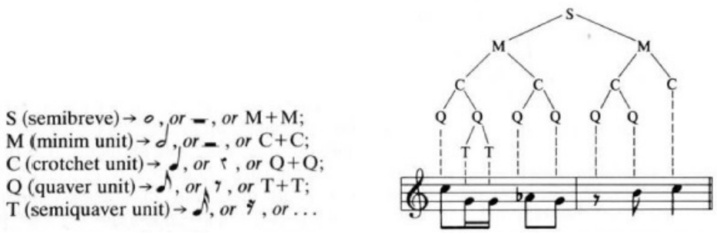
2.1 Strongly hierarchical models

This section contains a discussion of strongly hierarchical models. This discussion includes theories on metrical grammars and metrical grids, which are early formalizations of meter perception and metrical structure, respectively. The section concludes with a brief overview of three other strongly hierarchical models that are more formal.

2.1.1 Metrical grammars

Early accounts for formalization of meter perception have arisen in the late 1960s and focus on the similarities of music and language. Simon (1968) states that appreciation of music relies on finding and understanding patterns. Inspired by findings in linguistics, Simon (1968) proposes that music has an underlying hierarchical structure, like language (Chomsky, 1957, 1965). Simon presents an algorithm that determines this underlying structure by looking at durations only. While Simon's algorithm works on a large timescale and on phrase structure

[1] Other differences between theories and models concern whether and how they incorporate performance factors, such as phrasing and expressive timing. Those factors are beyond the scope of the current project, but an overview of models especially designed to incorporate expressive timing can be found in Temperley (2013).



(a) (b)

Figure 2.1: (a) Example of a set of context-free realization rules for a $\frac{4}{4}$ meter, from Longuet-Higgins (1978, Figure 5, p. 150); (b) Example of a rhythm generated by realization rules, from Longuet-Higgins (1978, Figure 3, p. 150).

instead of meter, the idea that durational values are the most important factor for metrical interpretation of rhythm, recurs in other work.[2] Longuet-Higgins and Steedman (1971) define an algorithm based on durational values to find the time signature of Bach fugues. Longuet-Higgins (1976) uses the same idea to design an algorithm that transcribes live performance of classical melodies into musical notation. Longuet-Higgins and Lee (1982) design a model that imitates a listener who creates hypothetical groupings of notes through their relative lengths only. This model places bars in musical scores by working incrementally from left to right through the score, while following elementary rules based on the relative note lengths (Longuet-Higgins & Lee, 1982).

The assertion of an underlying structure in music is also found in the formal theory of *metrical grammar* by Longuet-Higgins (1978). In contrast to Simon's (1968) focus on phrase structure, Longuet-Higgins (1978) aims to develop a formally precise syntactic theory of *meter*. To reach this goal, Longuet-Higgins (1978) proposes to use computational science as the language to describe the complexity of perceptual and cognitive processes in the human mind, in the same way as differential calculus describes theoretical physics. Longuet-Higgins (1978) argues that the mental representation of music is a structure like *syntactic structures* in linguistics, where a sentence is not regarded as a sequence of words, but rather as a structure held together by syntactic relations (Chomsky, 1957). In the same way, rhythm can be described syntactically as a tree structure in which every node either represents a note or a rest, or branches into other nodes (Longuet-Higgins, 1978). Every rhythm is generated by a musical *grammar* that represents the meter (Longuet-Higgins, 1978); see Figure 2.1.

The formal theory of Longuet-Higgins (1978) resonates in the formal model of Longuet-Higgins and Lee (1984). This model particularly focuses on Simon's (1968) implicit claim that listeners can arrive at a rhythmic interpretation through relative durations of notes only.[3] Longuet-Higgins and Lee (1984) question how listeners perform this task easily, even though every sequence of note values is in principle rhythmically ambiguous and can be metrically interpreted in an infinite number of ways. Longuet-Higgins and Lee (1984) regard the work presented by Longuet-Higgins and Steedman (1971), Longuet-Higgins (1976) and Longuet-Higgins and Lee (1982) as partial solutions of the problem posed by this question. In order to obtain a complete account for the problem, Longuet-Higgins and Lee (1984) propose a generative theory based on Lindblom and Sundberg (1969)[4], who argue there must be rules that govern how melodies are built up and propose a *generative grammar* (Chomsky, 1957) containing tree

[2] Simon (1968) does not make this idea explicit, but notes that the model makes use of relative durations of notes only: "nous avons seulement utilisé les durées des notes" (p. 33).

[3] Other clues like accents, staccato and legato, tonal relationships and lyrics can also be used to arrive at a rhythmic interpretation, but these clues are not necessary, according to Longuet-Higgins and Lee (1984).

[4] In the text, Longuet-Higgins and Lee (1984) refer to 'Lindblom and Sundberg (1972)', but that article is not retrievable, while there is an article 'Lindblom and Sundberg (1970)' in the reference list of which the version from 1969 does fit the content (the 1970 version is not retrievable either).

structures that represent these rules. Longuet-Higgins and Lee (1984) transpose the theory of Lindblom and Sundberg (1969) from beyond bar level to the rhythms of individual bars. This approach is based on ideas of Martin (1972), who argues that rhythm in speech (but also music) is hierarchical by definition and proposes a formal description of binary trees with accent levels. In line with Longuet-Higgins (1978), Longuet-Higgins and Lee (1984) argue that a time signature actually designates "a grammar consisting of a set of context-free realization rules" (p. 427).

Although every sequence of notes is rhythmically ambiguous, sometimes it is clear to every listener which metrical interpretation of a given sequence is 'natural' (Longuet-Higgins & Lee, 1984). Therefore, Longuet-Higgins and Lee (1984) argue that rhythm perception must be tightly constrained by assumptions of what is a 'natural' interpretation. To account for this, Longuet-Higgins and Lee (1984) provide a formal definition of syncopation. Now, the 'natural' metrical interpretation of a sequence is the metrical grammar that realizes an unsyncopated rhythm with respect to this grammar (Longuet-Higgins & Lee, 1984). The listener grasps this interpretation through the notion of a *regular passage*, a sequence of bars that are all generated by the same standard meter, without syncopations (Longuet-Higgins & Lee, 1984). Through an algorithm, the metrical interpretation can now be built up from the shortest intervals in a formal way (Longuet-Higgins & Lee, 1984).[5]

The theory of Longuet-Higgins (1978) and the model of Longuet-Higgins and Lee (1984) involve a strongly hierarchical and regular sense of meter. Regularity results from explicitly only taking 'standard meters' into account: meters in which all subdivisions on every hierarchical level are equal. For this, no explicit motivation is given. The need for hierarchical representations is motivated by Longuet-Higgins (1978) through linguistic arguments along the lines of Chomsky (1957, 1965). In addition, multiple authors argue that models which use Markov processes are not suited for formalization of musical patterns (Simon, 1968; Slawson, 1968), in the same way that these models cannot adequately describe grammatical structure either (as argued by Chomsky, 1956, 1957). Contemporary arguments for hierarchical processing of music emphasize how temporal processing in general makes use of hierarchical principles, independent of whether the sequences are linguistic, visual, or letter and number sequences (Simon & Sumner, 1968; Vos, 1973; Palmer & Krumhansl, 1990). A third class of arguments is music-theoretic, such as the analogy with higher-level hierarchical structures such as sonatas and movements (Longuet-Higgins, 1976), or the analogy with perceptual hierarchy in tonality (Palmer & Krumhansl, 1990).

In contrast to Longuet-Higgins (1978) and Longuet-Higgins and Lee (1984), recent theories of meter perception tend to focus less on linguistic components of meter and more on a motor component. As will be discussed in Section 2.4.1, involvement of this motor component constitutes theories of meter perception that are grounded in neuroimaging studies and fit well within present general theories of cognition. In contrast, theories with linguistic motivations provide a more abstract level of description.

2.1.2 Metrical grids

While theories of metrical grammars formalize the process of *meter perception*, theories of *metrical grids* focus on the formalization of *metrical structure* only. After such a formalization, further research develops models that show how this metrical grid is actually induced in the listener. Among theories of metrical grids, the influential book *A Generative Theory of Tonal Music* by Lerdahl and Jackendoff (1983) is a starting point.

Before defining the metrical grid, Lerdahl and Jackendoff (1983) distinguish *grouping* (spontaneous segmentation of the sound signals) and *meter* (the inferred regular pattern of strong

[5]Longuet-Higgins and Lee (1984) provide an extension to the algorithm that parses the rhythmic structure of passages that *are* syncopated. This extension is based on an additionally proposed rule for phrasing, which is less elaborated.

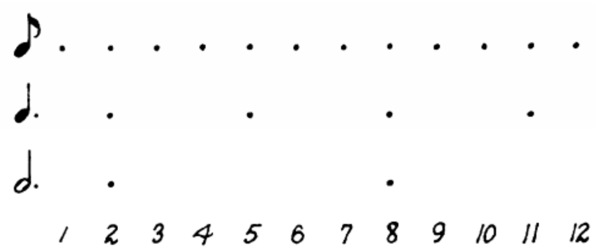


Figure 2.2: Dot notation, from Lerdahl and Jackendoff (1983, example 4.1, p. 68).

and weak beats). Metrical structure is defined separately, as “the regular, hierarchical pattern of beats to which the listener relates musical events” (Lerdahl & Jackendoff, 1983, p. 17). Lerdahl and Jackendoff argue that grouping is common to many areas of human cognition and happens for all music, while metrical structure can only be inferred for some music.

Lerdahl and Jackendoff (1983) also disambiguate the notion of accent in music and define three different forms: *phenomenal accent* (any emphasized moment, such as attack points, leaps and sudden changes), *structural accent* (an accent caused by gravity in the melodic or harmonic flow) and *metrical accent* (a strong beat with respect to metrical context). These accents relate; for instance, phenomenal accent serves as input for extrapolation of a regular pattern of metrical accents, the *metrical pattern* (Lerdahl & Jackendoff, 1983). However, metrical accent is special, as it is only a mental construct that is relative to the metrical pattern (Lerdahl & Jackendoff, 1983). In turn, this mental metrical pattern is inferred from (but not identical to) patterns of actual accentuation in music (Lerdahl & Jackendoff, 1983). Once a listener has inferred a metrical pattern, this pattern is only renounced in the face of strong counterevidence (Lerdahl & Jackendoff, 1983).

In Lerdahl and Jackendoff (1983), elements that make up the metrical pattern are infinitesimal beats and are represented by dots (following Imbrie, 1973; and Komar, 1971); see Figure 2.2. Since meter entails periodic alternations of strong and weak beats, Lerdahl and Jackendoff infer there must be a metrical hierarchy that involves two or more levels of beats (inspired by Yeston, 1976). In this hierarchy, called the *metrical grid*, strength and level of beats are related: a *strong beat* on some level is also a beat (weak or strong) on a higher level (Lerdahl & Jackendoff, 1983).

The notation of Lerdahl and Jackendoff (1983) is strongly hierarchical, but does not necessarily presuppose a regular beat. However, regularity arises from the formulation of *metrical well-formedness rules* for metrical structure of tonal music (Lerdahl & Jackendoff, 1983). For instance, beats must be equally spaced at all metrically important levels, as this is “the norm in tonal music” (Lerdahl & Jackendoff, 1983, p. 20). Similar reasoning through normativity appears throughout the book and is not backed by perceptual or cognitive arguments. Instead, Lerdahl and Jackendoff appeal to intuition; things ‘are heard’ in a certain way or simply ‘must be’. The approach of Lerdahl and Jackendoff is a first, intuitive attempt to formally define metrical structure. However, it does yield a metrical representation that is innovative, as it is provides insight into the metrical hierarchy than the traditional prosodic notation of precursors such as Cooper and Meyer (1960), in which the relation between metrical level and beat strength is not clear. According to London (2012), the dot notation of Lerdahl and Jackendoff (1983) “has become ubiquitous in musical analysis” (p. 79).

Lerdahl and Jackendoff (1983) expand their formalization of metrical structure to meter perception through definition of *metrical preference rules.* Whereas well-formedness rules decide which metrical interpretations are possible, preference rules govern listeners’ preference of certain interpretations over others. Together, the well-formedness rules and preference rules form a model: when one applies the rules to a rhythm, one gets the most preferred metrical interpretation of that rhythm. However, Lerdahl and Jackendoff do not specify details or relative weights of the contributing factors they mention (this is remarked as well by Palmer &

Krumhansl, 1987a; and Clarke, 1999). Therefore, the model of Lerdahl and Jackendoff is not a formal model, whereas the algorithm of Longuet-Higgins and Lee (1984) is. However, arguments of both Lerdahl and Jackendoff (1983) and Longuet-Higgins and Lee (1984) for correctness or 'naturalness' of certain interpretations appeal to intuition only.

Initially, the metrical well-formedness rules of Lerdahl and Jackendoff (1983) seem restricting, and it seems that they are claimed to be universal. However, later Lerdahl and Jackendoff argue that these rules can be altered in order to fit other metrical idioms than classical Western music. The degree of regularity of Lerdahl and Jackendoffs representation is therefore not fixed, but depends on the exact formulation of the metrical well-formedness rules. This nuance in the theory of Lerdahl and Jackendoff is rarely quoted. Indeed, when the application of Lerdahl and Jackendoffs theory in empirical research will be discussed in the following paragraphs, it will show that its most restricted version is referred to, as this fits the majority of classical Western music.

2.1.3 Other strongly hierarchical models

After the theories of Longuet-Higgins (1978) and Lerdahl and Jackendoff (1983), other researchers have developed models that are also strongly hierarchical, discrete and regular, but more formal. For instance, Parncutt (1994) provides a quantitative model in which perceived meter is the simultaneous perception of different isochronous pulses. The model determines salience of the pulses in phenomenal accents through occurrence (number of events matching the isochronous template) and tempo (how closely the pulse train approximates 100 beats per minute)[6] and superimposes the most salient pulse templates to create a metrical hierarchy (Parncutt, 1994). Metrical accents derived from this hierarchy agree well with experimental results (Parncutt, 1994). The filter-based model of Todd (1994) proposes frequency-domain filters that output a set of periodicities. A culture-specific top-down process then matches the output with metrical patterns (Todd, 1994). As argued by Clarke (1999), both the models of Parncutt and Todd are inspired by the idea that meter perception also has a motor component. In Section 2.4.1, we will see how recent findings account for further substantiated models of meter perception with a motor component.

The final model that is important with regard to the current project, is the probabilistic model of Temperley (2007), that utilizes the metrical grid of Lerdahl and Jackendoff (1983); see Figure 2.3. According to Temperley, deriving the meter from a given rhythmical pattern of onsets is equal to aligning the correct metrical grid to the pattern. In order to find the most probable metrical grid, Temperley uses Bayes' Theorem: the probability of a certain grid given some onset pattern is proportionate to the probability of that onset pattern given the grid, multiplied by the probability of that grid in general. To calculate the latter two factors, the model uses a generative approach: it calculates the probability that the onset pattern *is generated by* the grid and the probability of different possible grids being generated at all (Temperley, 2007). By training the model on a data set, the required parameters to calculate both factors are established. Temperleys model yields positive results, especially on the lower levels of the metrical hierarchy. Another advantage of probabilistic models is that they can be used to simulate effects of enculturation on rhythm perception. Van der Weij, Pearce, and Honing (2016) present such a model.

2.2 A weakly hierarchical model

Apart from strongly hierarchical theories of metrical structure, theories have been proposed that postulate a smaller amount of levels. This section discusses the clock model of Povel and

[6] A pulse is more salient when its rate more closely approximates 100 beats per minute (Parncutt, 1994).

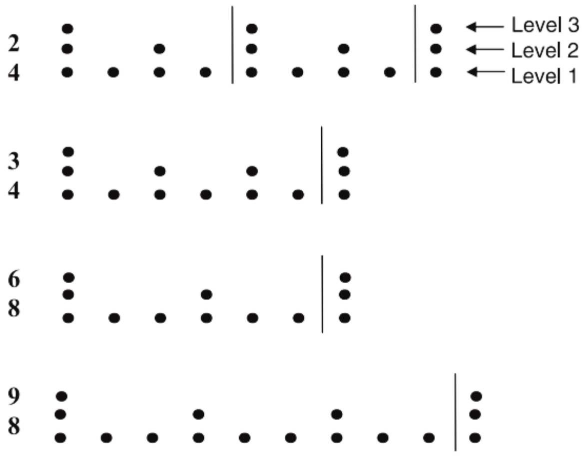


Figure 2.3: Metrical grids pertaining to time signatures, from Temperley (2007, Fig. 3.2, p. 25).

| | | . . . | . . | . . | | . .

Figure 2.4: A temporal pattern, after Povel and Essens (1985, Fig. 2, p. 416).

Essens (1985). The next section will argue that strongly hierarchical models are favored over weakly hierarchical models.

The model of Povel and Essens (1985) is an extension to Povel (1981), who does not assert a multi-leveled hierarchy of beats, but rather one beat level and one subdivision level below. In Povel and Essens, the notion of an *internal clock* is central.[7] The listener uses this clock to specify temporal structure in auditory patterns (Povel & Essens, 1985). The clock is hierarchical in the following sense: it has a unit (equally spaced pulsing) that has a location (or phase) and subdivisions (Povel & Essens, 1985). In the model of Povel and Essens, temporal patterns are represented by isochronous frames of potential onsets, such that there is either an onset or silence at every point on the frame (see Figure 2.4). Now, finding the correct meter of a given rhythm consists of aligning the clock that best fits the pattern and determining the best fitting subdivision.

The model of Povel and Essens (1985) falls apart into a clock model and a subdivision model. In the clock model, the fit of the clock is mainly determined by accents. This is in accordance with Lerdahl and Jackendoff (1983), but while Lerdahl and Jackendoff formulate many preference rules, Povel and Essens incorporate only three clues, based on tone onsets only (without information on pitch, duration or loudness). These clues are provided by Povel and Okkerman (1981) and are as follows: every isolated tone receives an accent; in a series of two tones, the second tone receives an accent; in a series of three or more tones, the first and final tone both receive an accent. The model of Povel and Essens now uses these rules to indicate the accented notes in every onset pattern. Then, the model extracts the best clock from the resulting accent patterns by collecting counterevidence for every clock (Povel & Essens, 1985). This counterevidence is determined by how many 'clock ticks' coincide with unaccented events or silences, including a parameter that weights the penalty of the latter relative to the former (Povel & Essens, 1985). The model also takes into account that the duration of the clock unit should be a divisor of the stimulus period, such that it keeps in phase (Povel & Essens, 1985).

The subdivision model of Povel and Essens (1985) attempts to find the correct subdivisions of a clock and relies on arguments about coding complexity. Povel and Essens presume that note onsets between clock ticks are represented in the model by a code that represents how the

[7] Povel and Essens (1985) avoid music-theoretic terms like 'beat', as they intend to regard temporal patterns in general, rather than musical patterns only.

clock is divided into (potentially irregular) subdivisions. Povel and Essens further suppose that the efficiency of the code for subdivisions is inversely related to the number of symbols needed in the code. Equal intervals give the simplest coding (as an equal division can be represented by fewer symbols), non-regular intervals give a more complex coding (Povel & Essens, 1985).

The idea of coding makes the model of Povel and Essens (1985) intrinsically different from models like Longuet-Higgins and Lee (1984) and Lerdahl and Jackendoff (1983). In the latter two models, the rhythm is perceived as relating to a metrical framework that is already in the mind, whereas in Povel and Essens, the rhythm is heard only within in a single beat framework (the clock) and is further coded by subdivisions that exactly describe the rhythm. Essens (1995) describes an additional difference: in Povel and Essens, hierarchy is based on a higher level time unit (i.e., the clock or beat), whereas in Longuet-Higgins and Lee, the hierarchy is built up from the smallest interval.[8] Apart from this difference, the model of Povel and Essens (1985) can be seen as weakly hierarchical (the subdivision level cannot be further divided into different levels) and less regular (a clock might be subdivided into unequal time spans).

The model of Povel and Essens (1985) is tested through different experiments. Povel and Essens themselves test the model by creating multiple permutations from one temporal template. In these permutations, only the order of inter-onset intervals (the duration between the onsets, or attack points, of two beats) is changed, such that type or number of certain intervals does not affect the results (Povel & Essens, 1985). Now, the model groups these patterns into nine categories on the basis of how strongly the best clock per pattern is induced (Povel & Essens, 1985). This induction strength decreases per category and depends on how many clock ticks coincide with unaccented elements or silence (Povel & Essens, 1985). In one experiment, Povel and Essens find that both learning time and reproduction deviation increase per category (with the exception of one outlier). In another experiment, the best clock is additionally induced by adding a low-pitched isochronous sequence, which significantly improved learning time and reproduction accuracy (Povel & Essens, 1985). Again, clock strength categories are also a significant factor between results (Povel & Essens, 1985). However, in this experiment, all clocks had a time unit of four, which might imply a bias in stimuli patterns. It should also be noted that reproduction paradigms require conscious attention while this process may be unconscious and effortless (as remarked by Huron, 2006). In all, these two experiments indicate that the clock model of Povel and Essens is at least an accurate model of beat induction.

A third experiment in Povel and Essens (1985) tests the subdivision model; essentially the meter perception part of the model. This experiment is based on the finding that the same sequence combined with different additional low-pitched clocks is often judged as different (Povel & Essens, 1985). According to Povel and Essens, this could be due to the different coding of the subdivisions of the clock. To test this hypothesis, Povel and Essens design pairs of sequences combined with additional clocks of either length three or four and test the correspondence of theoretic complexity predictions with complexity judgments by participants. Povel and Essens find that theoretic predictions are close to actual judgments, but as Essens (1995) later comments, it does not take interference by the induction strength of the *clock* into full account. Therefore, the experiment of Povel and Essens does not imply whether the subdivision model is accurate on itself. Essens (1995) resolves this confound in an improved experiment that separates the fit of the clock (*C-score*) and subdivision complexity. In an immediate reproduction task, Essens (1995) finds that only C-score has a significant effect. In contrast to Povel and Essens (1985), Essens (1995) concludes that the subdivision model does not capture coding of temporal sequences well.

Shmulevich and Povel (2000) aim to improve the subdivision model of Povel and Essens (1985) as follows: they formalize the coding complexity aspect by weighting the different possi-

[8] Essens (1995) mentions that Essens and Povel (1985) found no support for the hypothesis that the smallest interval is used for structuring the sequence. Because of this, Essens argues that the model of Povel and Essens (1985) is favored over the model of Longuet-Higgins and Lee (1984).

bilities of subdivision and adding a score for repetitions. The parameters are determined by a complexity judgment task in Essens (1995) and optimized to increase correlation (Shmulevich & Povel, 2000). Shmulevich and Povel test the subdivision model by comparing its results in a complexity judgment task with two other measures of complexity: Tanguiane (1993) and Lempel and Ziv (1976). The latter two measures have low correlation scores, while the model of Shmulevich and Povel scores high. Shmulevich and Povel conclude that their measure is robust, as the parameters found from one set of patterns yield high predictive power for another set. This is true, but it might also imply that the measure of Shmulevich and Povel was 'best prepared' of all measures, as it was trained on similar data. Indeed, the Tanguiane measure and the Lempel and Ziv measure also show low correlation with Essens's (1995) study, which was used to train Shmulevich and Povel's (2000) model. Furthermore, the measure of Shmulevich and Povel is the only measure that is based on an empirically tested model of rhythm perception to begin with. It might therefore be that the high correlation score of the Shmulevich and Povel measure is mainly caused by the fact that it is fitted to similar data. Altogether, the study by Shmulevich and Povel (2000) attempts to prove the cognitive reality of the subdivision model of Povel and Essens (1985), but multiple factors undermine the strength of the evidence.

In summary, the clock model of Povel and Essens (1985) is supported by different experiments, but the cognitive reality of the subdivision model is doubtful. The established part of the model of Povel and Essens is therefore mainly a model for beat induction. This part does not account for hierarchy and therefore falls short of being a good cognitive model for meter perception.

2.3 Evidence for strongly hierarchical representations

This section discusses evidence for strongly hierarchical representations. First, studies are presented that support the models of Longuet-Higgins and Lee (1984) and Lerdahl and Jackendoff (1983) in particular. The section concludes with general evidence for the cognitive reality of a strongly hierarchical representation of meter.

One of the predictions of Longuet-Higgins and Lee (1984) was recently confirmed in an EEG study: in participants without advanced music training, Ladinig, Honing, Háden, and Winkler (2009) found that syncopation on a more salient position evokes a stronger response. In both attentive and electrophysiological preattentive conditions (as obtained through Mismatch Negativity responses in event-related brain potentials), participants are better at detecting syncopations on stronger metrical positions than at weaker metrical positions (Ladinig et al., 2009).[9]

The theory of Lerdahl and Jackendoff (1983) also performs well in models and experiments, even when restricted for regularity as mentioned in Section 2.1.2. Palmer and Krumhansl (1987a) investigated whether participants judged segments of musical phrases in a Bach fugue as complete and compared different conditions in which either the pitch pattern or the temporal pattern was preserved. The judgments correlated significantly with the theory of Lerdahl and Jackendoff on metrical structure (Palmer & Krumhansl, 1987a). In a similar experiment, there was no correlation, but Palmer and Krumhansl (1987a) argue that this might be due to the average trial length being too short to firmly establish a metrical structure in the listener. Furthermore, Palmer and Krumhansl (1987b) found agreement with metrical predictions of Lerdahl and Jackendoff in three out of four phrase judgment tasks of classical music that was either altered in pitch or temporal pattern.

In an extensive and controlled study, Palmer and Krumhansl (1990) find several sources of evidence for the representation of Lerdahl and Jackendoff (1983). First, frequency distributions of note events in compositions correlated with the theoretically predicted 'handprint' of meter

[9]The result for the preattentive condition in slightly weakened in Ladinig, Honing, Háden, and Winkler (2011).

by Lerdahl and Jackendoff, independent from composer or style (Palmer & Krumhansl, 1990). Second, goodness-of-fit judgments (rating the fit of single temporal events in a given metrical context), correlated significantly with the music-theoretic predictions (Palmer & Krumhansl, 1990). Third, a discrimination task (correctly remembering whether two beats relative to a context beat are the same or different) showed that discrimination judgments were more accurate for events in metrically strong than metrically weak locations (Palmer & Krumhansl, 1990). Furthermore, Palmer and Krumhansl (1990) found that the respective compositional, perceptual and memory evidences are all highly correlated with each other and are reinforced by musical experience: in both the goodness-of-fit judgment and discrimination task, musicians discriminated more hierarchical levels than non-musicians (and also more than was suggested by the instruction). In summary, Palmer and Krumhansl (1990) show that perception and memory of temporal relationships is dependent on the meter that is suggested by the temporal context, and argue that this reflects a mental hierarchical framework of accents.

In all, multiple empirical studies support the strongly hierarchical models of metrical grammars and metrical grids, while we have seen in Section 2.2 that there is only empirical support for the *beat induction aspect* of the weakly hierarchical model of Povel and Essens (1985). Apart from this, there is more evidence for the cognitive reality of a strongly hierarchical representation of metrical structure. Thul and Toussaint (2008) tested the outcome closeness of different mathematical models to human measures. Models and human performance were tested on widely different data sets: Povel and Essens (1985), Shmulevich and Povel (2000), Essens (1995) and Fitch and Rosenfeld (2007). Thul and Toussaint (2008) found that models that use a metrical hierarchy of weights (such as Longuet-Higgins & Lee, 1984, Smith & Honing, 2006, Keith, 1991 and Toussaint, 2002) to calculate syncopation more closely model the measure 'human meter complexity' (how well people track the underlying metrical beat, or pulse, of a rhythm) than other models. However, these models less closely model human reproduction quality. Whereas the *data* of both Povel and Essens (1985) and Shmulevich and Povel (2000) have been used in Thul and Thoussaint's assessment, the *models* of these papers were not tested. This is regrettable, since this could have provided supplementary indication of the cognitive validity of these models with respect to the strongly hierarchical models. Additionally, it should be noted that the study of Thul and Toussaint (2008) evaluates perceptual or performance complexity measures *derived from* models instead of the cognitive reality of the models *themselves*.

Strongly hierarchical representations of metrical structure are not only supported by behavioral studies (Palmer & Krumhansl, 1990), but also by electrophysiological evidence. In the earlier mentioned article by Ladinig et al. (2009), a significantly different event-related brain potential was found for strong and weak syncopations. This indicates that there is a level on which omission of a beat evokes a strong response, a level on which omission evokes a weak response and a level on which it evokes no response – hence, a hierarchy of at least three distinctive levels. As Ladinig et al. (2009) found, this also holds for listeners without extensive music training. In another study of event-related brain potentials, Schaefer, Vlek, and Desain (2011) did not only find a significant difference in signals between accented and unaccented events, but also further differentiation of unaccented events.

In conclusion, there is much empirical evidence that a strongly hierarchical approach is favorable over a weakly hierarchical approach. Both behavioral and electrophysiological evidence points to cognitive reality of multi-leveled hierarchical metrical structure. Furthermore, the weakly hierarchical model of Povel and Essens (1985) does not describe meter perception more accurately than the strongly hierarchical models.

2.4 Non-isochronous metrical structure

This section is concerned with the theory of London (2012), as presented in the book *Hearing in Time: Psychological Aspects of Musical Meter*. Like the theories of Longuet-Higgins (1978) and

Lerdahl and Jackendoff (1983), the theory of London involves a strongly hierarchical representation of meter. However, Longuet-Higgins (1978) and Lerdahl and Jackendoff (1983) mainly focus on Western classical music, while London takes an approach that is cross-cultural and grounded in empirical evidence. Furthermore, London's theory investigates which structures are perceivable as meters through formulation of well-formedness constraints, along with many examples of perceivable meters.

This section first presents the motivation and perceptual grounding for London's (2012) theory. Subsequently, we evaluate London's motivations with regard to recent findings on the motor component of meter perception. Then, the theory and representation of London will be explained. The final subsection argues why London's theory best fits the goals of the current thesis. It also discusses the ways in which the theory of London is not formal and how this leads to underspecifications and inconsistencies.

2.4.1 Motivation and perceptual grounding

As discussed in Section 1.1, London (2012) defines meter as a form of entrainment behavior that can be learned and allows listeners to synchronize their perception to rhythms they hear. This view is inspired by time-continuous entrainment models such as Large and Jones (1999), in which the mind acts to rhythm as a resonating system. Such a system entrains to periodicities in music and thereby generates peaks in attentional energy or expectation (Large & Jones, 1999; London, 2012). Large and Palmer (2002) extend this single beat view to a system in which multiple oscillations of different periodicities combine into a metrical expectancy curve with multiple peaks (London, 2012). London (2012) takes this system as a starting point for his entrainment theory and explains how the attentional peaks of the system mark infinitesimal time points on which musical events of relatively greater salience are expected. Through identification of these time points, London constructs a discrete model motivated by time-continuous processes.

Before constructing a representation of meter, London (2012) explores the temporal constraints on perception of periodicities. Backed by literature on cognition and perception (such as cortical processing), London states that entrainment only occurs for periodicities from about 100 ms to 5-6 seconds. Moreover, a sense of beat is only felt in the subrange of 200-250 ms to about 2 seconds, with a preference region around 600 ms, or 100 beats per minute (London, 2012). London theorizes that these differences, along with structuring, account for "hierarchically integrated cycles of attention and expectation" (p. 46) and therefore a hierarchical structure with multiple levels.

London's (2012) ideas about meter are not only motivated by these temporal constraints, but also by neurobiological research. While in the 1960-80s the origin of meter was mainly sought in language (as we have seen in Section 2.1.1), more recently there is also attention to the idea that meter has a motor component.[10] Evidence for this idea has been found in various studies. In an fMRI study, Chen, Penhune, and Zatorre (2008) found that listening to musical rhythms activates multiple motor areas in the brain. These areas are not only related to motor control and learning (basal ganglia) and integration of sensory and motor information (cerebellum), but also execution of movement (pre-motor cortex and supplementary motor area), and they were activated even in the absence of movement (Chen et al., 2008; London, 2012). In an MEG study, Iversen, Repp, and Patel (2009) found that metrical interpretation had an effect in the beta range response, which likely plays a role in motor processing. Additionally, several studies (Phillips-Silver & Trainor, 2005, 2007, 2008; Trainor, Gao, Lei, Lehtovaara, & Harris, 2009; Wang & Tsai, 2009) found that stimulation of the vestibular system (by head movement or directly through galvanic stimulation) influences rhythm perception. On top of that, Todd and Lee (2015) argue through their comprehensive survey of brain research that for every brain area

[10]There is still much research on the relation between music and language as well, see for instance Mithen (2007) or Patel (2008).

associated with rhythm perception, there is a close correlation with the vestibular sensory-motor network.[11]

As noted by Phillips-Silver and Trainor (2007), the way in which metrical beat is extracted from body movement could be explained through the dynamic attending theory of Jones and Boltz (1989). This way, the just proved importance of motor aspects is represented in entrainment theories like London's (2012), who concludes that meter is "a kind of sensorimotor entrainment" (p. 48) or "a kind of virtual motion" (p. 132). In general, entrainment is one of the promising aspects in London's approach, as also more recent EEG research found that rhythmic stimuli elicited spontaneous emergence of an internal representation of beat, possibly indicating neuronal entrainment (Nozaradan, Peretz, & Mouraux, 2012). The idea of entrainment adds the important nuance that meter perception is an automatic process: even though it happens in the listener, it does not require their conscious effort (London, 2012, p. 68). However, London also argues that this entrainment behavior is highly practiced (p. 4), just like motor behavior is highly practiced in general. Furthermore, it has both conscious and subconscious aspects, depending on the time frame (London, 2012, note 5.3, p. 202).

2.4.2 Evaluation of the motivation

In the previous subsection, we have seen that the motivation for London's (2012) theory relies partially on research on the motor component of meter perception. Meanwhile, there are recent developments in this area that are not taken into account by London. These recent developments go further than motor areas in the brain and the vestibular system alone and might challenge London's views. In order to judge whether London's motivation still holds, this subsection discusses these recent developments.

Recent accounts on the motor component of meter perception are related to the introduction of theories on *embodied cognition* within music cognition by Iyer (2002).[12] Embodied cognition is the idea that "cognition is an activity that is structured by the body situated in its environment" (Iyer, 2002, pp. 388–389). In this idea, perception is based on the sensory-motor capacities of the body and has evolved together with motor action (Iyer, 2002). Applied to music, and in our case, to beat and meter, the theory postulates that "we may use our bodily movements to help parse the metric structure of music" (Toiviainen, Luck, & Thompson, 2010, p. 59). The major addition here is the role of the *whole body* on top of the role of the vestibular system and motor areas in the brain.

An important theory that incorporates embodied cognition is the *sensory-motor theory* of Todd and Lee (2015). Originally formulated by Todd et al. (1999), before the emergence of embodied cognition within music cognition, it asserts that the internal representation of the musculoskeletal system mediates beat induction, even when one does not actually move. Partially in contrast to the idea of London (2012), this theory claims that beat induction is not a passive process but rather a sensory-guided action (Todd et al., 1999).[13] Indeed, Todd and Lee argue that their explanation of beat induction is incompatible with oscillator-based approaches such as Large and Kolen (1994) and Large and Jones (1999). Instead, Todd and Lee propose that beat induction is mediated through two distinct sensory-motor circuits, in which oscillators are unnecessary from both an explanatory and an evolutionary viewpoint. This is in conflict with the ideas of London, who argues that metrical entrainment is "a form of coupled oscillation or resonance" (p. 48). As we have seen, London earlier also cites the studies of Large and Jones (1999) and Large and Palmer (2002) to illustrate this view. However, it is not clear whether oscillators are essential to London's theory and representation, so it can be argued that the

[11] For a more complete and comprehensive overview of this brain research, see London (2012) or Todd and Lee (2015).

[12] Seemingly independently from Iyer (2002), Leman (2007) proposes a similar theory.

[13] Note that Todd et al. (1999) and Todd and Lee (2015) focus on beat induction; no account for meter perception is given.

sensory-motor theory does not directly challenge the ideas of London.

Todd and Lee (2015) mention several studies that stress a link between musical beats and the body. For instance, Styns, van Noorden, Moelants, and Leman (2007) show that music influences the rate in which people walk. In a perceptual task, Todd, Cousins, and Lee (2007) found that 16 % of variation in preferred beat rate can be predicted from anthropometric factors, suggesting direct influence of the body to perception of rhythm. Additionally, Dahl, Huron, Brod, and Altenmüller (2014) found that body dimensions influence preferred dance tempo. Todd and Lee (2015) link these results to McAuley, Jones, Holub, Johnston, and Miller (2006). While McAuley et al. (2006) themselves do not offer an explanation for the finding that children have shorter preferred beat periods than adults, Todd and Lee (2015) propose this may be because they have smaller bodies. However, as Repp (2007a) remarks, this entails that there should be a sex difference as well, since women are significantly smaller than men. This difference was not found in either McAuley et al. (2006) or Todd et al. (2007), challenging the results of those studies (Repp, 2007a). However, Dahl et al. (2014) did report this sex difference and proved in a second experiment that this sex difference is fully due to the corresponding height difference. It remains unclear why this difference did not appear in the studies of McAuley et al. (2006) and Todd et al. (2007). Lastly, in a kinetic analysis of body movements, Toiviainen et al. (2010) showed that pulsations on different levels of metrical hierarchy can simultaneously be embodied in music-induced movement (with faster levels in the extremities of the body and slower levels in the central parts). This finding is especially relevant to the current study, as this is a statement of meter rather than beat only. However, it does not necessarily entail that movements help parse metrical structure, as postulated by Toiviainen et al. (2010) . Hence, the collection of mentioned articles does not constitute convincing evidence for the sensory-motor theory with respect to metrical structure.

In addition to practical evidence, Todd and Lee (2015) introduce multiple theoretic motivations for the sensory-motor theory. However, many of those are too speculative to be accepted as evidence. For instance, Todd and Lee claim that the theory might predict why some animal species exhibit beat induction (e.g., humans and birds) and some do not (e.g., primates): as humans and birds are both bipedal, their typical bodily movement patterns (*eigenmovements*) are different from those in primates, which are quadrupedal. As remarked by Bregman, Iversen, Lichman, Reinhart, and Patel (2013), these results are also predicted by the *vocal learning theory* of Patel (2006), that states that only species capable of complex vocal learning have the capacity to synchronize their movements to a musical beat. However, as Bregman et al. (2013) note, informal observations have shown that horses, which are vocal non-learners, occasionally move in synchrony with a musical beat as well. Whether horses actually synchronize is yet to be scientifically tested, but this challenges the vocal learning theory (Bregman et al., 2013). Now, Todd and Lee argue that the sensory-motor theory can account for this through a special property of a horse's body: its pendular long neck. However, this account is speculative and requires further elaboration; it entails that any bipedal species and any species with a long neck synchronizes to a beat, from kangaroo to giraffe (unless they are withheld by other reasons).[14] In conclusion, more research is needed before either theory can be deemed more plausible.

Regarding the early state of the sensory-motor theory and its lack of strong evidence, a more nuanced view may be favored. In such a view, meter perception does have a motor component that might be rooted in embodied cognition, but to what extent is yet to be determined. Such a combined view can be seen in Naveda and Leman (2011), who hypothesize that meter might stem from a musical-choreographic form in dance, but has further developed independent of body. In a case study of dance (with topological gesture analysis) and music recordings of Afro-Brazilian samba, Naveda and Leman (2011) find that typical aspects of meter, such as symmetry, periodicity and preference rules (for tempo and distribution of metrical levels),

[14] On the other hand, vocal learning theory has the same problem of having to prove capacity of synchronization for a wide range of vocal learners like bats and dolphins (Patel, 2014).

reflect properties in dance and human body morphology. Naveda and Leman (2011, p. 492) explicitly do not claim that dance is fully responsible for meter, but argue that meter and dance do affect each other. The idea of involvement of dance (without the full sensory-motor theory) suggests a more nuanced view on the role of embodied cognition on meter perception. This view less undermines the oscillator part of London's (2012) theory than the sensory-motor theory of Todd and Lee (2015) does; regardless of whether or not oscillators are essential to London's (2012) theory.

The 'nuanced dance view' was further supported by Lee, Barrett, Kim, Lim, and Lee (2015), who found enhanced meter perception in participants viewing a dance video. Lee et al. (2015) also showed larger enhancement for participants who were familiar with the choreography. Furthermore, this view is in line with findings that dance is used as a means of maintaining non-isochronous meters, such as Balkan *aksak* (Fracile, 2003), Romanian, Bulgarian and Macedonian dance (Proca-Ciortea, 1969; Rice, 1994, 2000; Singer, 1974) and Norwegian *springar* (Haugen, 2015). In turn, non-isochronous meters fit well into the theory of London (2012).

In contradiction to the earlier discussed argument of Todd and Lee (2015), it should be noted that the sensory-motor theory is not necessarily incompatible with oscillator-based approaches. In the theory of Todd and Lee (2015), oscillators might seem unnecessary to explain beat induction, but this does actually not preclude that entrainment processes in the brain do play a role in beat induction or meter perception. Furthermore, in many of the aforementioned articles, instead of the brain, now the body functions as a resonating system (e.g., Styns et al., 2007; Dahl et al., 2014). This way, many views on embodied meter perception still have an entrainment aspect and are therefore indeed compatible with the theory of London (2012). So even when embodied cognition to great extent influences meter perception, this does not challenge the theory of London. Indeed, embodied cognition reinforces the aforementioned idea that periodicity rather than regularity is central to meter perception. As we have seen in the previous subsection, this latter idea fits into London's theory of meter perception.

In conclusion, recent developments on the motor component of meter perception do not necessarily undermine the theory of London (2012). Rather, ideas on embodied cognition fit within London's point of view that meter perception is a from of entrainment behavior.

2.4.3 Theory and representation

In his theory, London (2012) proposes the following representation of metrical structure. The hierarchically integrated cycles mentioned in Section 2.4.1 are represented within a cyclical representation of meter (London, 2012). In the corresponding diagram, the outer circle represents one full period of the meter (a measure) over which time flows clockwise (London, 2012); see Figure 2.5. On this circle, dots represent peaks of 'entrained sensorimotor attention', or "temporal targets for actual or virtual motor behaviors" (London, 2012, p. 83), in line with Section 2.4.1. The twelve o'clock position marks the *downbeat* (the strongest beat) of the meter (London, 2012). The cycle that contains all dots (i.e., the outer circle) is called the *N cycle*, in which N is the number of dots on the cycle, or its *cardinality* (London, 2012). There can be multiple cycles in one metric pattern; the meter of Figure 2.5 has two cycles apart from the N cycle (London, 2012). These cycles are called *subcycles* and they represent the way in which a set of attentional peaks coordinates (London, 2012). Note that this is roughly equivalent to a linear metrical representation consisting of different levels, such as in Lerdahl and Jackendoff (1983). Hence, London (2012) argues that the proposed cyclical representation of meter relates attentional models to strongly hierarchical notions of metrical structure. One of the cycles carries the *tactus*, which is the most salient pulse (London, 2012), or in motor terms, the pulse to which a listener would tap their foot (Lerdahl & Jackendoff, 1983).[15] This cycle is called the *beat cycle* and may be the N cycle, but is more often a subcycle, as the tactus level usually

[15] For a more elaborate discussion of the tactus, see London (2012, p. 30–33).

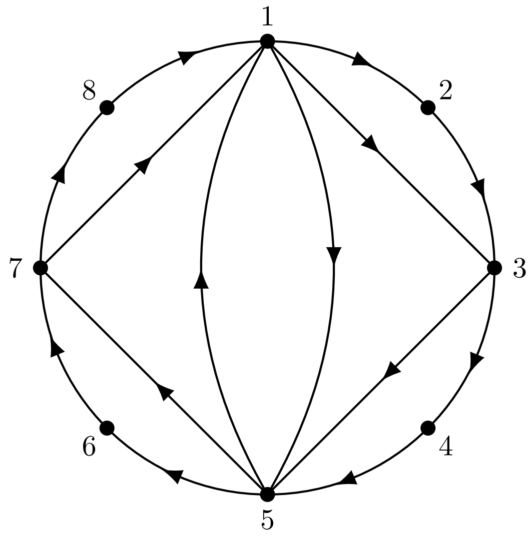


Figure 2.5: Cyclical representation of meter with higher levels drawn within the circle, after London (2012, fig. 5.6, p. 87). Arrows indicate the direction of temporal flow.

includes at least one level of subdivision (London, 2012). Another cycle that gets a specific name is the *half-measure*. This cycle divides the N cycle in two halves, but these halves need not necessarily be equal (London, 2012). Analogously, a cycle that divides the N cycle in three thirds (which need not necessarily be equal either), is called a *third-measure* (London, 2012). Finally, London (2012) adds that each attentional peak has some amount of temporal spread, although this spread is constrained by the number of cycles on which this beat occurs.

For the determination of which meters can and cannot exist, London (2012) is inspired by the metrical well-formedness rules of Lerdahl and Jackendoff (1983) that define which meters are musically *possible* instead of which meters are *typical* for a specific style. However, while the well-formedness rules of Lerdahl and Jackendoff are based on intuitions, London (2012) defines a set of constraints that are grounded in *perception* and *structure* (such as global combinatoric constraints), thereby addressing the problem of appeal to intuition only. By defining constraints in this way, London (2012) aims to include the widest possible range of meters from both Western and non-Western cultures. We have seen that this was possible as well in the theory of Lerdahl and Jackendoff, but it was not developed in detail.

London's (2012) *well-formedness constraints* are defined on the cyclical representation of meter. The first group of constraints ("WFC 1.1-4" in London, 2012) is temporal. These constraints are grounded in perception as mentioned before: inter-onset intervals on the N cycle are at least $\approx$ 100 ms, the inter-onset intervals on the beat cycle should all lie between $\approx$ 400 ms and $\approx$ 1200 ms, and the maximum duration for the full cycle is $\approx$ 5000 ms (London, 2012). The second group of constraints ("WFC 2.1-3") are minimal structural requirements on the beat cycle: there must be one, which must involve at least two beats (London, 2012). The third group of constraints ("WFC 3.1-3") concerns 'formal requirements', namely that all cycles are continuous, have the same total period and are all in phase (London, 2012). These are requirements of entrainment: if these constraints are violated, patterns are no longer hierarchically coherent; for instance, a downbeat is not articulated at all levels, or subcycles do not properly nest (see London, 2012, pp. 92–94). There is also one special constraint ("WFC 3.4") that requires that time points on a subcycle are not adjacent on the next lowest cycle (London, 2012). London argues that this constraint ensures that beats and subdivisions on the same cycle are not mixed up. In all, London argues that the above well-formedness constraints collectively "create and maintain the hierarchic integrity of the meter" (p. 97). The well-formedness constraints of London are reprinted in Section 4.2.

The well-formedness constraints by London (2012) make it possible that there are meters that contain non-isochronous pulses, henceforth *non-isochronous meters.*[16] Rhythms that induce non-isochronous meters in listeners are relatively rare in Western music, but do occur in other musical cultures, such as Balkan (Proca-Ciortea, 1969; Singer, 1974; Rice, 1994; Fracile, 2003), African (Kauffman, 1980; Locke, 1982, 1998; Patel, 2008), Arabian (Touma, 1996), Indian (Morris, 1998; Clayton, 2008) and Scandinavian music (Kvifte, 2007; Haugen, 2015). Furthermore, comparative studies with adults and infants show that there is no innate predisposition for isochronous meters; experience and enculturation overrule the supposed cognitive complexity of non-isochronous meters (Hannon & Trehub, 2005a, 2005b; Hannon et al., 2012; Kalender et al., 2013). This means that non-isochronous meters are not 'harder to grasp' than common Western meters. Since the notation of London makes non-isochronous meters possible and gives them the same status as isochronous meters, this representation is more cognitively plausible and less culture-specific than the others. This makes the metrical representation of London also less regular than some of the other discussed representations. As we have seen in the Section 2.4.1, the focus of London is on periodicities of pulses rather than their regularity.[17]

Of course, not all non-isochronous patterns can serve as meters. London (2012) realizes this and therefore investigates which properties make some non-isochronous meters "'regular enough' to function as a timing framework for sensorimotor-based entrainment" (p. 121). This investigation is spread out over Chapters 8 and 9 of London's book and it consists of perceptual constraints, examples and the notion of *maximal evenness.* Maximal evenness is the mathematical distribution of beats in a relatively regular or even way (London, 2012, p. 131). The notion was originally developed by Clough and Douthett (1991) in the context of scales and interval classes (London, 2012).

In order to explain the effect of maximal evenness, London (2012) borrows two more terms from studies on scale structure: *ambiguity* and *contradiction.*[18] These two notions relate to the different ways in which pitch distance between two tones can be measured: *generic* interval size and *specific* interval size (London, 2012, p. 130; Rahn, 1991, p. 34). Generic size refers to the number of scale degrees between the tones; for example, two for both C-E (C to D and D to E) and D-F (D to E and E to F) in the diatonic collection (London, 2012; Rahn, 1991). Specific size refers to the number of semitones between the tones; for example, four for C-E and five for D-F (London, 2012; Rahn, 1991). As London (2012, pp. 129–130) explains, ambiguity occurs in scale structure when the pitch intervals between two pairs of tones have the same specific size but contrasting generic sizes. London provides the following example of ambiguity within the diatonic collection: the specific size of both F-B and B-F is six semitones, but interval F-B has a generic size of three scale degrees (F to G, G to A and A to B; it is an augmented *fourth*), while B-F has a generic size of four scale degrees (B to C, C to D, D to E and E to F; it is a diminished *fifth*). Contradiction is a more severe form of discrepancy, in which one interval is larger in terms of specific size, while the other is larger in terms of generic size (London, 2012). This may happen for exotic scales, see Rahn (1991, p. 36).

Ambiguity and contradiction can be translated to the domain of metrical structure, in which units of the N cycle form the *specific* intervals and all higher-level periodicities form the *generic* intervals (London, 2012). These higher-level periodicities are beat intervals, and categories of

[16]In literature, these meters are also called 'complex', 'irregular', 'odd', 'additive' or 'asymmetric'. However, not all meters that fit the above definition are necessarily perceived as complex or irregular, as we will see later. Neither can all of these meters be described by an odd number, additive patterns or asymmetric figures. Because of these reasons, this thesis follows London (2012) in calling these meters non-isochronous.

[17]Sethares and Staley (2001) showed earlier that focusing on periodicities makes for a good account of non-isochronous meters.

[18]This ambiguity as a *property of meter* is not to be confused with London's (2012) 'metric ambiguity', which refers to duration patterns that "may give rise to different metric construals on different listening occasions" (p. 106). The metrical phenomena of ambiguity and contradiction should also not be confused with ambiguities and contradictions within London's theory itself, which will be explored in Section 5.1. Wherever confusion may arise, it will be specified which notion of ambiguity or contradiction is referred to.

beat intervals will hereafter be referred to as *beat classes*. For instance, as explained in London (2012, p. 131), when a beat of 2 units (two N cycle inter-onset intervals) occurs on the same level as a beat of 4 units (which can be subdivided as 2+2), ambiguity arises: the same specific interval of 4 units may amount to a different generic interval of one beat (the 4-unit beat) or two beats (two 2-unit beats). Verify that a contradiction would arise in more extreme cases, such as the coexistence of a 5-unit beat and a 2-unit beat on the same level. London (2012) argues that maximal evenness precludes ambiguity and contradiction within metrical structure and ensures that the different levels of the formalization remain distinct. This way, it ensures that different beat intervals still fall within the same temporal range, such that beats and subdivisions do not get mixed up perceptually (London, 2012). Apart from this, maximal evenness ensures that attentional energy is optimally distributed, and it optimizes periodic motor behavior (London, 2012). Thereby, maximal evenness is also motivated through the motor component of meter perception (London, 2012).

London (2012) favors maximal evenness over the stricter rule that on every level of the metrical structure, every strong beat is either spaced two or three beats apart, as in Lerdahl and Jackendoff (1983). This is because there can be meters are subdivision patterns of four or five elements under constraints of maximal evenness (e.g., London, 2012, pp. 144, 155, 161–162), while on the other hand, patterns that do abide by the rule of two or three, are not always maximally even (e.g., London, 2012, p. 164).

In many structures discussed by London (2012), the lowest level (the N cycle) is an isochronous pattern (though subject to potential expressive variation), such that the non-isochronous beats can be understood relative to this isochronous embedding (London, 2012), or *common fast pulse* (Kauffman, 1980; Kvifte, 2007). In fact, these are the structures that he refers to as non-isochronous meters (see London, 2012, p. 124). The focus in London therefore lies on these meters, although he also explores the properties of meters in which *the N cycle is non-isochronous* (see London, 2012, pp. 121–123). London concludes the well-formedness constraints with regularity requirements, of which most ("WFC 4.1.1-3") concern structures with non-isochronous elements on the N cycle (and avoid ambiguity and contradiction and this level). The last two constraints ("WFC 4.2.1-2") formulate a rule on maximal evenness for meters with an isochronous N cycle: a non-isochronous beat cycle must either be maximally even, or the cycle above the beat cycle must be maximally even (London, 2012). London (p. 158) argues that this way, maximal evenness is regarded for the meter as a whole and not on each individual subcycle. One could say that the cycle above a beat cycle can 'stabilize' the beat cycle if it is not maximally even. London also argues that these constraints, together with well-formedness constraint 3.4, ensure that no ambiguities and contradictions arise within metrical structure. We will see in Section 5.1 that the latter claim does not hold for all cases.

On account of *all* well-formedness constraints together, the representation of London (2012) aims to only regard patterns (whether isochronous on all levels or not) as well-formed if they can cognitively serve as a meter. An example of a well-formed non-isochronous meter in London's (2012) cyclical representation is shown in Figure 2.6.

London (2012, pp. 94–97) has additional remarks on the influence of tempo and absolute duration. London calls the description of all cycles that constitute a particular meter its *metrical type*. Metrical types may be further individuated on the basis of the absolute durations of inter-onset intervals within one of the cycles (London, 2012). Through temporal constraints, these factors interact and can thereby define which metrical types are possible at which tempo, and which cycles may serve as the beat cycle at that tempo (London, 2012). In this respect, one metrical type can have different tempo-dependent varieties, which are called *tempo-metrical types* (London, 2012). London's theory does not imply that there is a tempo-metrical type for every single tempo. Rather, there are only a few categorically distinct tempo-metrical types for each metrical type, which, at increasing tempi, change into one another via "transition zones" (London, 2012, p. 95).

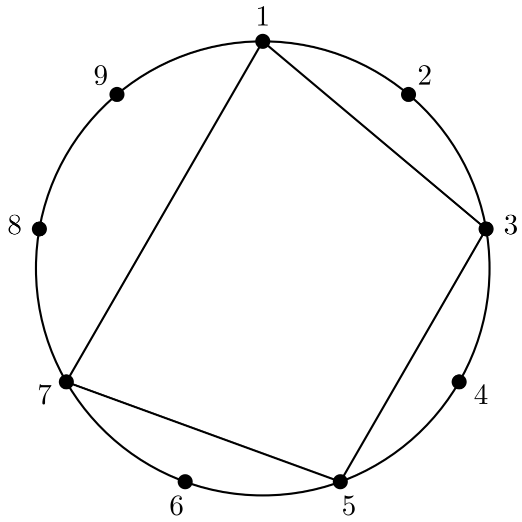


Figure 2.6: Example of London's (2012) representation of a permissible non-isochronous meter with a 2-2-2-3 structure, after London (2012, fig. 8.2, p. 124).

London's (2012) distinction between different tempo-metrical types of the same metrical type is somewhat confusing. London argues that this distinction "does not mean that each tempo-metrical type is defined in terms of the level heard as the tactus" (p. 95). Rather, London seems to imply that distinct tempo-metrical types constitute to categorically distinct patterns of perception, or 'complex gestalts' (p. 96). London seems to argue that now, within the same tempo-metrical type, or pattern of perception, still different levels can be heard as the tactus level. London does not make himself entirely clear in the explanation of this nuance, especially given the fact that tempo itself does influence which level can be perceived as the tactus through the temporal well-formedness constraints. In all, it is clear that tempo influences whether or not a particular meter can be perceived – and if so, that it can be perceived in different particular ways.

In summary, London's (2012) theory of metrical structure meets the requirements of simplicity and perception as presented in Chapter 1 of this thesis. The theory uses ideas of dynamic attending theory and recent findings about neuronal entrainment, the motor area in the brain and the vestibular system. Thereby, London's approach is cross-cultural, as it builds on periodicity rather than strict regularity, accounting for non-isochronous metrical structure as well as isochronous metrical structure. Furthermore, London's theory makes use of perceptual constraints, instead of intuitive rules such as in Lerdahl and Jackendoff (1983) and Longuet-Higgins and Lee (1984). Moreover, London provides many examples of perceivable meters outside of Western classical music, which stresses the cross-cultural aspects of his approach. All the while, London's (2012) metrical representation is discrete, which constitutes to simplicity.

As we have seen in Section 2.3, various studies suggest that strongly hierarchical representations of metrical structure are preferred over weakly hierarchical representations. Furthermore, throughout the current section, we have seen that non-isochronous metrical structure is an important part of metrical entrainment. London's (2012) theory of metrical structure includes both these aspects. Because of this, in combination with being grounded in perceptual studies, it best fits the goals of the current thesis. However, London's theory does not meet the requirement of formality. The next subsection will discuss why formalization of London's theory is necessary.

2.4.4 Formalization

London's (2012) has one important disadvantage: it is not formal. London's metrical representation is not defined in a formal way (for instance, as a cyclical graph) and neither are the well-formedness constraints. Furthermore, the theory leaves multiple details unspecified. This causes problems of underspecification and inconsistency.

The following details are not fully specified in the representation of London (2012). First, the influence of tempo on metrical structure. This is the distinction between different tempo-metrical types of the same metrical type, which can be deemed confusing, as discussed in the previous subsection. Second, presence of a downbeat. This concerns the question whether every meter contains one beat that is perceived as stronger than all others. As we will see in Section 3.3.3, London's (2012) well-formedness constraints do not specify a downbeat and the theory is ambiguous about it. Third, perceivable length of subdivision patterns. London's theory is ambiguous on what is the greatest perceivable length of subdivisions within one beat, and the well-formedness constraints do not specify this either. This point will be discussed in more detail in Section 5.1.1. Fourth, possible organizations of non-tactus subcycles. The well-formedness constraints of London do not make clear whether subcycles may be unevenly organized, as long as the beat cycle is maximally even. This point will be discussed in detail in Section 5.1.3.

There are also several inconsistencies between London's (2012) well-formedness constraints and theory. This happens with regard to the metrical phenomena of ambiguity and contradiction as presented in the previous subsection. London claims that the well-formedness constraints on maximal evenness jointly exclude any structure that involves ambiguity or contradiction, but we will see in Section 5.1.2 that this is not the case. There are related inconsistencies between London's theory and well-formedness constraints with regard to *beat class ratios*, which are the relative proportions of beat classes on the same level. Some beat class ratios (such as two beats having a ratio of 4:7) are not allowed by London's theory through arguments of combinatorics, decomposition and again ambiguity, but do pass the well-formedness constraints. Section 5.1.2 of this thesis contains a more detailed discussion of these inconsistencies and where they are found.

Additionally, London's (2012) theory does not give a clear overview of all perceivable meters; it does provide all perceivable meters for a few periodicities, but a recount makes clear that some structures are missing. Because of the above underspecifications and inconsistencies, it is not clear whether London would include or exclude the missing structures from the representation. Similarly, London's theory is sometimes inconsistent with the provided examples, or the lack thereof. Sometimes, London's theory contains examples that are not allowed by the well-formedness constraints, or does not mention certain structures (within lists of perceivable structures) that should be permitted by the well-formedness constraints. Several instances of these inconsistencies will be discussed in Section 5.1.

In this thesis, we propose a formalization of London's theory. Formalization will specify London's theory and address ambiguities and inconsistencies in the theory, in order to provide a clear and unambiguous overview of all perceivable meters. Finally, formalization may help to implement London's theory and representation in a formal cognitive model of meter perception, such that the theory can be tested empirically. There are already formalization projects on the theory of London in literature. Chapter 3 will discuss these projects and provide an overview of aspects that have been left open by London's proposal. In Chapter 4, we construct the formalization of the current project.

Chapter 3

Related work in formalizing metrical structure

Apart from the current project, other authors have also performed formalization projects on the theory of London (2012). This chapter discusses these projects and provides an overview of aspects that have been left open by London's proposal. Section 3.1 discusses the formalization of Forth (2012); Section 3.2 discusses the formalization of Gotham (2015). The discussions in both sections include the respective advantages and disadvantages of these projects, as well as their relevance for the current project. In Section 3.3, open ends in the theory of London with regard to formalization will be evaluated. This evaluation will include the respective approaches of Forth and Gotham, as well as the current approach. Section 3.4 concludes this chapter with general remarks on the formalizations of Forth and Gotham with regard to the current project.

3.1 Metrical trees in conceptual spaces

Forth (2012) captures London's (2012) theory of meter formally through a tree representation. Forth's trees represent tempo-metrical types and are formally described by a tuple (called a *metrical tree*) that consists of six sets and one integer. Three of the sets in the metrical tree describe the structure of the tree mathematically as a graph with vertices (or nodes), edges and vertex labels (Forth, 2012). Forth presents five definitions that capture that a tree is a connected, undirected graph containing no cycles and seventeen definitions to describe other properties of the tree, such as that it is perfect (every path from root to leaf is of equal length) and ordered (there is a total ordering for the children of each node). Every node represents a beat on a given level and every leaf node represents a beat on the N cycle (Forth, 2012). Three more sets describe the absolute timepoints of every node, the pulse inter-onset interval values of every node and the attentional energy values of every leaf node (Forth, 2012). Lastly, the one integer of the metrical tree denotes which level is the tactus cycle (Forth, 2012). Five more definitions and twelve constraints decide which possible instances of metrical trees are perceptually metrical, comparable to London's well-formedness constraints (Forth, 2012). Note that Forth's constraints are different from those of London: in Forth the maximum duration is 6000 ms (instead of 5000 ms) and all beat intervals on the tactus level should be in between 200 and 2000 ms (instead of 400 to 1200 ms). A permissible meter in Forth's (2012) representation is shown in Figure 3.1.

As Forth's (2012) formalization incorporates absolute timepoints on all tree nodes, it enables unequal subdivision not only at slower levels, but also at the fastest level. This means that there is no common fast pulse in Forth's formalization that corresponds with an isochronous embedding, relative to which non-isochronous beats are understood. This way, Forth's formalization captures how unequal subdivision can be internalized instead of counted, and it "allows for microrhythmic details to be considered as part of metrical structure, rather than either being left

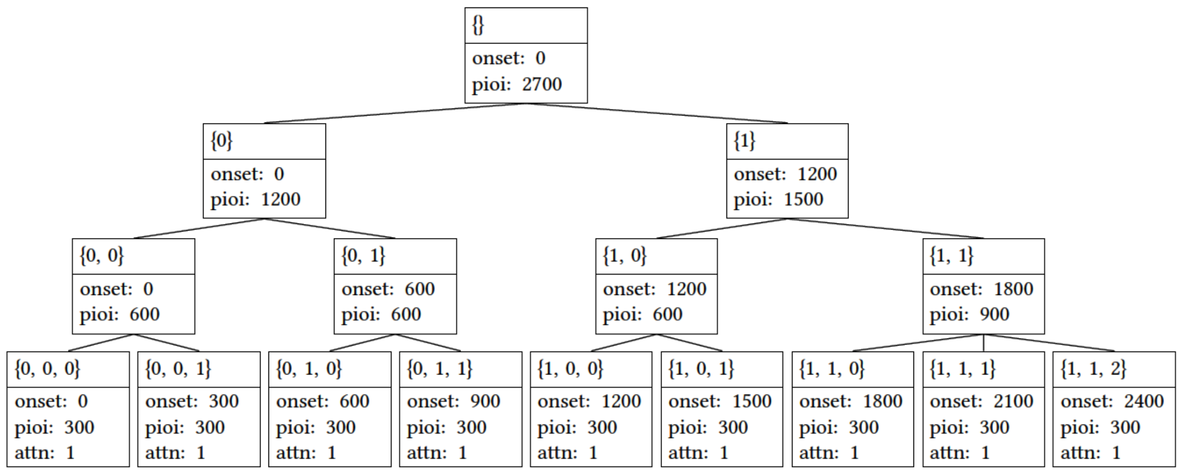


Figure 3.1: Example of Forth's (2012) representation of a permissible meter with a 2-2-2-3 pattern (London's (2012) representation for this meter is shown in Figure 2.6), from Forth (2012, fig. 5.2d, p. 118). Every set is a vertex label of the tree and denote a beat on a given level, 'onset' denotes the absolute point in time where the beat starts, 'pioi' the inter-onset interval between consecutive beats on the same level, and 'attn' the degree of attentional energy of every leaf node (Forth, 2012).

unaccounted for, or potentially misleadingly being considered as random or stylistic deviation from an abstract Western-centric theoretical conceptualisation of metre" (pp. 104–105). Forth here refers to Norwegian *springar* music, in which the beats are systemically unequally subdivided in a way that can not be described in terms of an underlying common fast pulse (Kvifte, 2007).[1] Meters that do *not* contain microrhythmic details still have a place in the formalization, as they are the structures without variation in timing or attentional energy, and they are seen as "prototypical of real-world patterns of entrainment" (Forth, 2012, p. 105). Such a prototypical meter is shown in Figure 3.1: the pioi-values in the second and third level are unequal, but in the lowest level they are equal.

Forth's (2012) project takes place in the framework of *conceptual spaces* of Gärdenfors (2000). Conceptual spaces provide a way of modeling cognitive representations; a conceptual space is defined as a "set of quality dimensions with a geometrical structure" (Gärdenfors, 2000, p. 24). As summarized by Forth, within this framework "concepts can be represented geometrically within perceptually-grounded quality dimensions, and where distance in the space corresponds to similarity" (Forth, 2012, p. 13). Forth explains that incorporation of individual timing relationships and attentional energy patterns requires conceptual spaces of higher dimensionality than incorporation of only prototypical structures. However, Forth argues that the addition of new dimensions is natural within this framework. Hereby, Forth's formalization forms a coherent whole that fits the framework in which it is defined.

3.2 Formalizing relationships

Gotham (2015) has undertaken another formalization project that is relevant to London's (2012) theory. Gotham's main goal is to visualize the different relationships that exist between metrical structures, focusing on relationships between *mixed* metrical structures. Gotham (2015, note 1) argues for the term *mixed* instead of *non-isochronous*, as the latter refers to more kinds of

[1]Arguably, music that *can* be described in terms of a common fast pulse, is not necessarily perceived this way either. Rice (2000) argues that Bulgarian musicians do not count an underlying pulse to keep track of the alternating patterns of two and three pulses, but instead have internalized this two-to-three relationship (see Gotham, 2015, note 36).

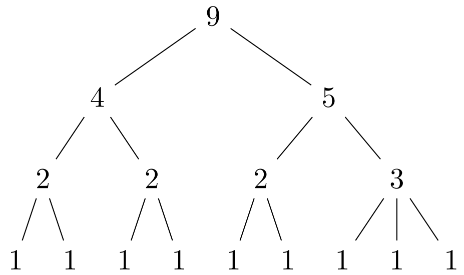


Figure 3.2: Tree structure after Gotham (2015, Figure 1).

non-isochrony than the formally different beat types, such as non-isochronous subdivisions, as we have seen in Forth (2012).

Gotham (2015) is particularly interested in mixed meters as these meters facilitate many compositional possibilities, while an overview of these possibilities has been lacking so far (Gotham, 2015). In order to visualize relationships between metrical structures, Gotham uses formalisms to define the metrical structures themselves. Gotham uses more than one representation, since different relationships apply to different aspects of a meter's structure. Therefore, some relationships can more easily be defined on some representations, while other representations are more suitable for other relationships. The formalisms of Gotham are all based on three initial criteria. The most important and strict criterion is that the majority of temporal events should be metrical (Gotham, 2015). Gotham admits that this excludes the earlier discussed springar repertoire, but argues that "mixed meter is not a part of that musical language anyway (higher levels tend to be isochronous)" (Gotham, 2015, note 9). Note that Gotham here departs from Forth (2012) by not including structural microrhythmic details. This is a deliberate choice, as we have seen that Gotham explicitly only aims to investigate mixed meters. Gotham's other criteria are the presence of a common fast pulse and grouping all pulses into twos and threes (excluding swing rhythms, see Gotham, 2015, note 12). Both these criteria can be relaxed for some relationships (Gotham, 2015).

Gotham (2015) uses four different mathematical representations of metrical structure. The first representation is an ordered succession of twos and threes, separated by nested brackets that "differentiate between different higher-level groupings of the same beat succession" (Gotham, 2015, p. 2). For instance, $((2,2)(2,3))$ denotes that there are three beats of two units (duple beats) and one beat of three units (triple beat). On the next level, the first two duple beats are grouped together while the last duple beat is grouped with the triple beat. The second representation is Forth's (2012) tree formalization, of which Gotham explains that this representation is favored because it is both mathematically formal and yields an intuitive graphical representation. However, Gotham presents a less elaborated version of Forth's trees. An example of Gotham's trees for the same $((2,2)(2,3))$ meter is provided in Figure 3.2. Two other representations are simplifications in which the order of the beats is not important (Gotham, 2015). The first of these is called the *meter vector* and it denotes how many duple beats and triple beats the meter has (Gotham, 2015); for example, $\{3,1\}$ for $((2,2)(2,3))$. The second shows the number of units per level of the meter (Gotham, 2015); for example, $\langle 9\ 4 \rangle$, denoting nine units on the lowest level and four units on the next level – the only way to group four units is to respectively have two units and one unit on the subsequent level. Note that indeed, the examples of the latter two representations also represent meters that are similar to $((2,2)(2,3))$, but in which the beats may be in a different order. Together, the different representations help to formally define relationships between different mixed meters.

3.3 Open ends

In the process of formalizing the theory of London (2012), several open ends are stumbled upon. This section evaluates these open ends. This evaluation also contains a discussion of the choices made by Forth (2012) and Gotham (2015) with respect to the open ends and investigations on whether elements of the formalizations of Forth and Gotham are also useful in the current framework. Where this is not the case, novel proposals for addressing the open ends are provided. This will account for the design of a formalization that best suits the current approach.

3.3.1 Isochrony of the fastest pulse

This open end concerns the question whether to incorporate meters with a non-isochronous fastest pulse. London's (2012) well-formedness constraints 4.1.1, 4.1.2 and 4.1.3 explicitly allow for meters in which the inter-onset intervals on the N cycle are non-isochronous. On pp. 121–123, London substantiates this choice by two arguments: rapid non-isochronous rhythms can be produced without mental subdivision (Repp, London, & Keller, 2005), and subdivision categories in very fast beats cannot be defined as consisting of two or three pulses, as these are too fast to perceive as such (Polak, 2010). This entails that there are meters in which there is no common fast pulse that guides the perception of the (possibly non-isochronous) slower beats that are multiples of this fast pulse. Instead, these meters may have a slower pulse that is isochronous, that guides perception of the (possibly non-isochronous) faster beats that are subdivisions of this *common slow pulse* (Kvifte, 2007). The isochronous pulse may be the tactus, but it may also be an even slower pulse (Polak & London, 2014). Structures without a common fast pulse are included in London's theory (pp. 121–123), but without much further information. Instead, more emphasis is put on structures that are isochronous on the fastest level (London, 2012, pp. 123–170). Microrhythmic details do recur in London's 'Many Meters Hypothesis', but in all, this factor is not central in the theory.

More than London's (2012) theory, Forth's (2012) formalization explicitly includes metrical structure with a non-isochronous fastest pulse, by actively incorporating these structures into the representation and discussing them. The different approaches of London and Forth are reflected in the way they respectively treat the metrical phenomenon of ambiguity (see London, 2012, pp. 130–131, 134–135; Forth, 2012, p. 105). As we have seen in Section 2.4.3, London does not tolerate ambiguities within the representation. In contrast, Forth argues in favor of metrical ambiguity in order to fully account for microrhythmic details and "extreme non-isochronous meters" (p. 105), as Kvifte (2004) argues for their meaningfulness to springar music. Furthermore, in Forth's formalization, there are relatively few restrictions as to which non-isochronous structures are allowed (note that there is no common slow pulse either). Because of this, Forth's formalization will consider more structures metrical than London's theory does; and indeed, Forth's formalization incorporates structures that may not be considered well-formed or metrical at all (Forth, 2012, pp. 105, 147). This is a deliberate choice, as Forth's definition of conceptual spaces of meter "explicitly aims to encompass all theoretically possible metrical structures that can be derived from the first principles of perceptual and physiological limitations" (p. 147), even though large parts of these spaces may not correspond to perceptually relevant structures. In conclusion, Forth puts more emphasis on designing a wider framework in which no possibility is excluded, while London puts more emphasis on specifying a smaller region of well-formed structures.

This difference between Forth (2012) and London (2012) on handling metrical ambiguity is also reflected in their respective consideration of maximal evenness. Forth does not consider maximal evenness at all, while it is central in London's theory, namely to exclude metrical ambiguity. Although Forth does not explicitly argue against maximal evenness itself, principles of maximal evenness are probably deliberately ruled out because of Forth's tolerance to ambiguity.

In order to account for microrhythmic details, Forth's (2012) approach requires conceptual spaces of higher dimensionality. More concretely, Forth defines two sets (consisting of real values representing absolute timepoints of the tree nodes and their pulse inter-onset interval values) with as many elements as there are beats in the representation. This amounts to a large number of possible meters, which limits its applicability in frameworks other than conceptual spaces, and makes Forth's formalization less suitable within the current project. The number of possible meters under a given set of rules or constraints will often be subject of later discussions and will hereafter be referred to as the *metrical space*.

As we have seen, the approach of Gotham (2015) is different with regard to this aspect. Gotham's main focus is on mixed meters and therefore not on microrhythmic details. Indeed, Gotham defines in the 'hard criterion' that the majority of temporal events should be metrical, and in a softer criterion the presence of a common fast pulse. However, in cases in which this latter criterion is relaxed, Gotham's formalization approaches Forth's formalization; in the other cases, it is closer to London's theory. Therefore, the formalisms of Gotham are only suitable as long as the initial criteria are strictly followed.

As London's (2012) main focus lies on meters with an isochronous fastest pulse, only those meters will be incorporated in the current formalization, in the interest of simplicity. Of course, important information is omitted by this simplification, and it will be harder to account for swing rhythms and springar rhythms, but the simplification makes it more feasible to investigate non-isochronous meters that do have a common fast pulse. This includes the metrical structure of many musical cultures, see Polak and London (2014, Table 7.1).[2] Furthermore, this choice restricts the metrical space and also makes the formalization itself simpler mathematically. This means that London's well-formedness constraints 4.1.1-3 are not formalized. Instead, we add a constraint which requires that the fastest pulse in the meter is isochronous.[3]

This entails that the current approach includes only the meters that are called 'mixed meters' in Gotham (2015). However, as this term, like any other term than a neologism, gives rise to confusion (see Gotham, 2015, note 1), this thesis maintains the name 'non-isochronous meters' for this class of meters. As structures that do have a non-isochronous fastest pulse may be the subject of future research, the broader sense of this term may be reestablished later.

3.3.2 Influence of tempo

This open end concerns tempo and its influence on tactus salience (which level is the tactus level) and meter finding. As we have seen in Section 2.4.3, London's (2012) tempo-metrical types account for categorically distinct ways in which ways a particular meter can be perceived and whether it can be perceived at all. In the same section, it was argued that the form of this account is somewhat confusing: London argues that the different tempo-metrical types are not distinguished by which level can be perceived as the tactus, while tempo itself does influence which level can be perceived as the tactus through the temporal well-formedness constraints.

A formal theory requires hard boundaries. In London (2012), not only the "transition zones" between different tempo-metrical types, but also the perceptual thresholds of tactus salience and meter finding are repeatedly addressed as being around or between some values. In order to get an unambiguous formalization, these vague boundaries should be defined as single threshold values, but it is an open question which values should be picked. Moreover, choosing single values has an effect on the aforementioned transition zones, which may have minor theoretical implications. Note that single threshold values need not necessarily be *fixed*, as they can be defined as parameters of the formalization.

Forth (2012) takes tempo into account by defining sets of absolute timepoints and inter-onset interval values and by defining temporal constrains. As was the case for the previous open end,

[2]This table lists different possibilities in terms of which levels are non-isochronous. Both 'type 1' and 'type 3' of this table are represented in the current formalization.

[3]This comes close to the first edition of London (2012), see London (2012, note 5.4, p. 202).

this makes the metrical space very large and therefore less suitable within the current project.

In contrast, Gotham (2015) does not take tempo into account at all, as the focus lies on relations between meters rather than the meters themselves. In Gotham's formalization, there is a notion of relative tempo (via multiples, such that translation of subdivision structures from one metrical level to another can be defined), but not absolute tempo in the sense of an absolute time span between metrical units. This makes Gotham's theory less suitable within the current project, because it does not integrate an important aspect of London's (2012) theory.

The current formalization aims to incorporate the influence of tempo, but in a simple way. Therefore, it disregards the continuous transition zones and instead makes tempo a property of metrical structure. Now, a formalization of London's (2012) theory should have a clear link to continuity of time, so it is important that the tempo is indicated by a real number (otherwise, the notion of tempo becomes 'granular' on an arbitrary level). Like in Forth (2012), such a choice gives rise to an infinite number of possible meters, but in the current formalization, these meters can still be categorized into a finite number of structures. Even though the tempo property is part of the description of a tempo-metrical type in the formalization, one should see it separately. Tempo is not an actual structural property of a specific meter, but a property that constraints which metrical structures can exist within which temporal range.

3.3.3 Presence of a downbeat

The presence of a downbeat in London's (2012) representation is not sufficiently specified. That is, it is unclear whether there is one beat that is perceived as stronger than all others, such that different rotations of the same meter around this point are categorically distinct. On p. 83, London states that the twelve o'clock position of the cycle marks the location of the downbeat. However, it is subsequently stated that there is not an explicit measure of relative strengths or amplitudes of each attentional peak (London, 2012). Though London mentions that the alignment of component cycles does give an implicit weighting, there is never one downbeat; there is always at least one beat that is visually equivalent to the beat at twelve o'clock. Often, there is a subcycle in London's representation that represents the half-measure, such that there is only one beat that is weighted equally to the beat on the twelve o'clock position, namely the other beat of this half-measure cycle. However, in general, every beat on the highest subcycle in London's representation is equally weighted to the beat on the twelve o'clock position, so this might be more than one beat as well. This way, the downbeat is not fully represented visually. Moreover, none of the well-formedness constraints of London states that there is one downbeat that is weighted heavier than any other beat on the cycle.

However, the downbeat has been a point of consideration in the development of London's (2012) well-formedness constraints: well-formedness constraint 3.3 precludes structures that do not coordinate and thereby have more than one downbeat (see London, 2012, p. 94). Furthermore, London, pp. 136–137 states that the orientation of beats relative to the downbeat in non-isochronous meters makes a difference for distinguishing different metrical organizations. London argues that a pattern that consists of three short beats and one long beat ("S-S-S-L") has "four distinct permutations", and "these phase rotations should be regarded as metrically distinct variants of a given beat cycle; they are [non-isochronous] subtypes of the basic S-S-S-L metrical type" (p. 137). There is clearly a sense of downbeat through this argument.

On the other hand, in pp. 137–139, London (2012) agrees with Locke (1998) that some non-isochronous meters are *accentless*: there are different beats, but there are no relative accents, in contrast to the accented metrical hierarchy in Western $\frac{4}{4}$ meter. Consequently, when there is no accent, it can be said that there is no downbeat. However, London argues on the contrary that the rotations are still different, and this gives rise to the listener calling one beat the first and construing a downbeat themselves (London, 2012). It might be that there is no downbeat in the sense of *phenomenal* accent *in the rhythm*, because such an accent is not needed to determine the length of the cycle, as this is clear from the asymmetry of the pattern (London, 2012). However, the listener does place a *metrical* accent somewhere to define where cycle begins. Now, listeners might disagree over which beat is the first, but for every listener, there is a downbeat indeed (London, 2012, pp. 137–142).

In all, London (2012) argues for a downbeat that makes different rotations of non-isochronous meters non-equivalent, but the downbeat is not fully represented visually and not articulated in the well-formedness constraints. It is therefore an open choice whether to define a downbeat in the formalization.

Forth (2012) explicitly defines the downbeat in the representation (which is in every level the nodes that coincide with the *root* of the metrical tree, p. 112). However, in terms of weighting, it is questionable whether this downbeat is really distinct in Forth's formalization. This problem emerges in Forth's discussion on the appointment of attentional energy values to the beats. Forth proposes to appoint each N cycle timepoint with a real number that represents the degree of attentional energy on that point. However, Forth argues that "the degree of attentional energy associated with any pulse may be a combination of many factors" (p. 113) and is therefore beyond the scope of that project. Because of this, Forth chooses to default all attentional energy values to 1. As a possible alternative, Forth proposes to appoint the degrees in terms of alignment such as in Lerdahl and Jackendoff (1983) (which would in fact return the aforementioned problem as in London, 2012), but he argues that "further evidence should be sought before incorporating such an assumption into the model" (Forth, 2012, p. 113). In this sense, the different beats in Forth's formalization or not really distinct. The presence of a downbeat is taken for granted in Gotham (2015), but not explicitly articulated in any of the formalisms, nor visually present.

The current formalization will incorporate a downbeat as this decision follows the theory of London (2012) most closely. This will happen in a similar way as in Forth (2012): the downbeat will be made explicit in the mathematical structure of the formalization: the requirement will be added that there is always one layer that consists of only one beat (namely, the first beat of the meter, equivalent to the twelve o'clock position in London, 2012), making this beat stand out from the others as the downbeat. Distinction of the downbeat in terms of weighting will be incorporated different with respect to Forth, as definition of a set of attentional values would make the metrical space larger again. Instead, we use London's assertion that the alignment of cycles may serve to assign weights to all beats. To this end, a function is incorporated in the formalization that provides a relative weight to the pulses on the N cycle by looking at their alignment. This function, called the saliency function, is completely dependent on the meter, such that it does not increase the metrical space. This way, London's ideas about weighting by alignment may later be incorporated in formal cognitive models of meter finding.

It should be noted that a saliency function entails the additional claim that for every pulse level, the pulses that do not align with pulses on a higher level will be assumed to be equivalent. For instance, in a 3-2-2-2 structure, the second and third beat of the triplet will be assumed to be equivalent to each other, and also to every second beat of the duplets. Empirical evidence in either existing articles (e.g., serial distance effects in Palmer & Krumhansl, 1990, pp. 737–738) or further research may suggest that these beats are not equivalent, or that there should be an absolute weighting of pulses instead of a relative weighting. Because of the claim made by this equivalence assumption, the saliency function is not used as a basis of the other constraints and is only used to appoint a downbeat and a relative weighting that may later be used in a cognitive model. In further research, an advanced version of the representation may be designed, which would contain a more advanced or absolute weighting. If this is the case, the current saliency function can serve to show where and how future empirical evidence for non-equivalence of beats or absolute weighting can be incorporated in the current formalization. For now, the equivalence assumption can safely be made, as the saliency function will stand apart from the metrical structure.

Addition of a downbeat fits best within a different visual representation. Ideas on visual representation will be presented in the next subsection.

3.3.4 Visual representation

The visual representation of metrical structure is not an open end in terms of underspecification or inconsistency in London's (2012) theory, as London unambiguously explains what the visual representation looks like. Rather, this is about visual detail that can be changed when visualizing a formalization. Indeed, London's visual representation has the important disadvantage that it does not make the highly hierarchical sense of meter in London fully clear (only indirectly: different weights can be found in the alignment of the subcycles). Other visual representations may be more fit to make hierarchy explicit, such as that of Lerdahl and Jackendoff (1983) or the tree representation of Forth (2012). Also, with regard to the previous point, selection of a different representation may help to visually define a downbeat. Additionally, Lerdahl (2015) argues that complex metrical structures would be difficult to read in the visual representation of London.

Both Forth (2012) and Gotham (2015) use different visual representations than London (2012). For Gotham, this choice has no consequences on the theoretical understanding of meter in the representation. For Forth, this choice seemingly does influence the theoretical understanding of meter in the representation. However, it is important to note that Forth's (2012) choice for a different visual representation than London should not be confused with the theoretical choice for a different framework. Forth notes that there are important differences between Forth's formalization and London's representation: Forth, pp. 106–108 argues that his formalization is not time-continuous and dynamic in the sense that London's theory is. However, these are not consequences of Forth's choice for a different visual representation. Rather, the visual representation of Forth is a consequence of the theoretical framework of geometric spaces.

As argued above, London's (2012) cyclical visual representation can without problems be changed in the current formalization. This thesis makes use of linear hierarchical dot representation of Lerdahl and Jackendoff (1983) to visualize the structures of the formalization. Such a representation more clearly visualizes the added downbeat and hierarchy in general than London's representation.

By this choice, we also lose an advantage of London's (2012) cyclical representation: it clearly visualizes how meter *perception* is an attentional entrainment process that is stable, recurring and periodical (London, 2012). However, the current choice affects the representation only visually, and we will see in Section 4.1 how the proposed formalization also codes London's cyclical representation. If in future research, this formalization is implemented in a cognitive model that aims to follow London's ideas on meter *perception* (rather than on metrical structure only), it is important that London's ideas on stability and recurrence are incorporated.

3.3.5 Length of subdivision patterns

The theory of London (2012) is a underspecified about the maximum length of subdivision patterns. Although London unambiguously makes clear that the length of permissible subdivision structures is not arbitrarily constrained to 2 or 3, but constrained by the temporal well-formedness constraints, London's theory remains ambiguous about where the maximum lies. This is a problem that would rather appear in the analysis of the formalization than as an open end before formalization,[4] but Forth (2012) and Gotham (2015) already consider this factor in the definitions of their formalizations.

While Forth's (2012) formalization is less restricted than the theory of London (2012) in terms of microrhythmic details and metrical ambiguity, Forth's formalization is more restricted with regard to subdivision structure. Forth (2012) requires that each node of the metrical tree

[4] Indeed, Section 5.1.1 discusses this underspecification in more detail.

can either have zero, two or three child nodes, following one of Lerdahl and Jackendoff's (1983) metrical well-formedness rules. Forth's restriction that the number of child notes cannot be one, agrees with London's well-formedness constraint 3.4 (Forth, 2012), but that the number of child nodes cannot be four or more does not agree with London's theory. Forth deliberately chooses to not allow for beat groups of more than three multiples in order to preserve uniformity within the formalization. This choice can also be defended by relative sparseness of groupings with larger multiples and by the following numerical argument. Every natural number larger than three can be defined as a sum of only twos and threes. Thus, for every structure in which there are groupings of multiples larger than three, we can find a roughly equivalent structure that only has twos and threes – Forth makes a similar argument on p. 115. However, this entails that in Forth, some subdivisions will have a cognitive meaning, while others are artifactual.

In Gotham (2015), all levels are grouped in twos and threes as an initial criterion, but this is again a 'soft criterion'. In cases in which the criterion is relaxed, Gotham's formalization approaches London's (2012) theory; in the other cases, it is closer to Forth (2012). Note that here, the 'relaxed case' is closer to London's theory than to Forth's, while it was the other way around for the open end regarding non-isochrony of the fastest pulse.

Notwithstanding arguments for restricting subdivision patterns to twos and threes only, London (2012) argues that groupings of large multiples are indeed perceivable (pp. 125, 132–133), and discusses multiple instances of structures containing such groupings (pp. 144, 155, 161–162). As the current project attempts to formalize London's theory as a goal on its own, there is therefore no reason to disregard groupings of multiples larger than three by forehand. Therefore, the current formalization maintains London's idea that there are no *a priori* restrictions on length of subdivision patterns. This also means that we cannot directly use the metrical representation of either Forth (2012) or Gotham (2015).[5]

3.4 General remarks and conclusion

As we have seen in the previous section, many aspects of the respective formalizations by Forth (2012) and Gotham (2015) make those formalizations unsuitable for the current project. These aspects jointly illustrate why this thesis proposes a novel formalization instead of taking one of the existing formalization projects as a starting point.

The largest difference between the current formalization on the one hand and Forth (2012) and Gotham (2015) on the other hand, is the perspective. Forth formalizes London's (2012) theory because of practical considerations and not to analyze the theory itself. Related to this difference, Forth formalizes the 'Many Meters Hypothesis' of London, while the current project formalizes *metrical structure* as defined and constrained in London. Gotham's main goal is the understanding of mutual relationships between meters. In contrast, the current approach focuses on the theory of London as a main goal. The choice for this theory is motivated by literature research in Chapter 2 that was aimed to find the optimal cognitive theory of metrical structure as defined in Chapter 1. Therefore, in Section 3.3 we analyzed the aspects in London's theory that are underspecified. In Chapter 4, we will provide a direct formalization of London's well-formedness constraints; in Chapter 5, we will show how this formalization generates meters that are not well-formed according to London's theory. We address the problems of underspecification and inconsistency by proposing additional constraints and analyzing their effect on the set of permissible meters. These factors constitute the main difference between the current formalization and related work.

The difference in perspective between the different approaches is not only reflected in the different choices with regard to the open ends, as discussed in the previous section, but also in

[5]Even though Gotham's (2015) criterion of twos and threes is soft, meters are explicitly defined in twos and threes in all of his *formalizations*. Thereby, Gotham gives no account of structures with subdivision patterns of larger multiples, nor on the constraints that would hold on such structures.

structural differences between the current formalization and those of Forth and Gotham. As these differences can be better appreciated when the current formalization is more clear, they will be discussed at adequate locations in the next chapter.

Chapter 4

Method

This chapter concerns the current proposal for formalization of the well-formedness constraints of London (2012). The chapter is divided into three sections. Section 4.1 contains the definitions and constraints of the current formalization, along with a clarification and several examples. Section 4.2 provides an informal derivation of the way in which London's definitions and well-formedness constraints follow from the current formalization, to show agreement between the current formalization and London's well-formedness constraints. As discussed in Section 2.4.4, London's theory is not described in formal terms, and several details of the theory are not completely specified. Furthermore, there are inconsistencies between London's well-formedness constraints, theory and examples. To address these inconsistencies, we suggest several additional constraints in Section 4.3. Subsequently, Chapter 5 will discuss how these additional constraints affect agreement between London's (2012) theory and well-formedness constraints.

4.1 Definitions and constraints

Definition 1. Let a *meter* be a 3-tuple $M = (L, c, u)$ in which:

- L is a set of *levels* $\ell_0, ..., \ell_k$ in which $k = \#L - 1$. Every level ℓ_x is a set of *beats*, such that every beat is represented by a natural number n. We say that x is the *level index* of ℓ_x, and that ℓ_x is the x-th level of L. We call L the *structure* of the meter.
- $c \in \mathbb{N}$ is a constant that denotes the *tactus level index*, such that ℓ_c is the *tactus level.*
- $u \in \mathbb{R}$ is a constant that denotes the *temporal unit*, measured in milliseconds.

In a meter M, the structure L represents the levels ℓ_x of beats n and their coordination. The tactus level index c denotes which level ℓ_x is the tactus level ℓ_c, and temporal unit u is the duration of any beat interval on the zeroth level ℓ_0, giving a measure of tempo. An instance of a meter can be compared to an instance of a *tempo-metrical type* (at a given tempo) in London (2012): The levels of a meter correspond to *cycles* in London, its beats correspond to *attentional peaks* and the structure as a whole to the *metrical type* (the alignment of the cycles). The tactus level is the *beat cycle* of London and the value of the temporal unit corresponds to one *inter-onset interval* on the *N cycle* (and therefore the shortest inter-onset interval of the meter). Note that the two constants c and u formally have the same status, although semantically, they are of a different order: the tactus level index is a *structural* property of the meter, while the temporal unit is a *temporal* property that can be used to tell within which temporal range a structure L (with a given tactus level c) is perceivable as a meter; see also Section 3.3.2. We call this range the *temporal frame* of the meter.

An example of the visual representation of meter is provided in Figure 4.1. In this example, structure L consists of three levels, corresponding to rows in the dot notation. The bottom

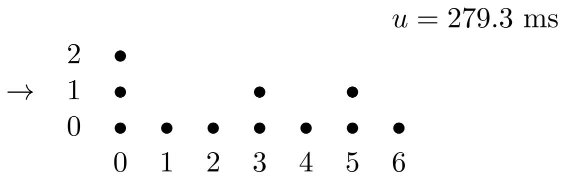


Figure 4.1: Dot representation for the meter $M = (\{\{0, 1, 2, 3, 4, 5, 6\}, \{0, 3, 5\}, \{0\}\}, 1, 279.3)$.

<table>
<tr><th>L in formal representation</th><th>Convenient notation</th></tr>
<tr><td>$\{\{0, 1\}, \{0\}\}$</td><td>2</td></tr>
<tr><td>$\{\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}, \{0, 4, 7, 10\}, \{0\}\}$</td><td>4-3-3-3</td></tr>
<tr><td>$\{\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}, \{0, 4, 7, 10\}, \{0, 7\}, \{0\}\}$</td><td>4-3|3-3</td></tr>
<tr><td>$\{\{0, 1, 2, ..., 15\}, \{0, 2, 4, 6, 8, 10, 12, 14\}, \{0, 4, 8, 12\}, \{0, 8\}, \{0\}\}$</td><td>2-2|2-2||2-2|2-2</td></tr>
</table>

Table 4.1: Examples of the convenient notation.

row corresponds to $\ell_0 = \{0, 1, 2, 3, 4, 5, 6\}$, the middle row to $\ell_1 = \{0, 3, 5\}$ and the top row to $\ell_2 = \{0\}$. Reading the columns instead of rows, we see which beat exists on which level(s). The arrow before row 1 denotes that the tactus level is the first level, ℓ_1, and the statement '$u = 279.3$ ms' denotes the value of constant u.

As a formally defined instance of a structure L does not always give a clear overview of its structural organization, we also adopt London's (2012) convenient notation of numbers separated by punctuation marks to refer to different structures (e.g., 3-2-2 for $\{\{0, 1, 2, 3, 4, 5, 6\}, \{0, 3, 5\}, \{0\}\}$ in Figure 4.1). In this notation, natural numbers represent the length of the beat intervals on the first level ℓ_1 (in terms of units of the zeroth level ℓ_0). Arrangements on the second and higher levels are shown by bars. For completeness, we add another convention that beat arrangements on levels higher than the second level are represented by multiple bars. Every possible structure L can unambiguously be represented by the convenient notation; for a few examples, see Table 4.1. For London's (2012) visual representation of these structures, see Figure 4.2.

On the sets in the meter definition, we define two orders:

Definition 2. On L, we define the strict total order $<$, such that for all $\ell_x, \ell_y \in L$, it holds that $\ell_x < \ell_y \iff \ell_x \supsetneq \ell_y$. We say that level ℓ_y is *higher than* (or *above*) level ℓ_x if $\ell_x < \ell_y$.

The levels ℓ_x of a meter are ordered, and every next level ℓ_{x+1} is strictly contained in the level below, ℓ_x. This ensures that we can speak unambiguously about the height of a level and say that one level is above or below another level. The definition restricts the possible number of levels of a meter and makes sure that the number of beats monotonically decreases by increasing level. Because of the strictness of the order, no two levels can be equal. Furthermore, a beat can only be present in a given level if it is also a beat in *all* lower levels.

Definition 3. On every level ℓ_x, we define the strict total order $<$ of the natural numbers on the beats of that level.

The beats n on every level ℓ_x are ordered in the same way as the natural numbers. That is, beat 0 comes before beat 1, beat 1 comes before beat 4, etc. This ensures that we can unambiguously refer to the different beats in a level.

Through the above definitions, the current formalization is structurally different from Forth's (2012) formalization. While Forth's trees are familiar mathematical objects for those with a background in graph theory, the current formalization offers a more compact structure by only defining a set and multiple subsets that jointly constitute a hierarchy. Because of this, the current formalization is less similar to familiar mathematical objects, but also requires

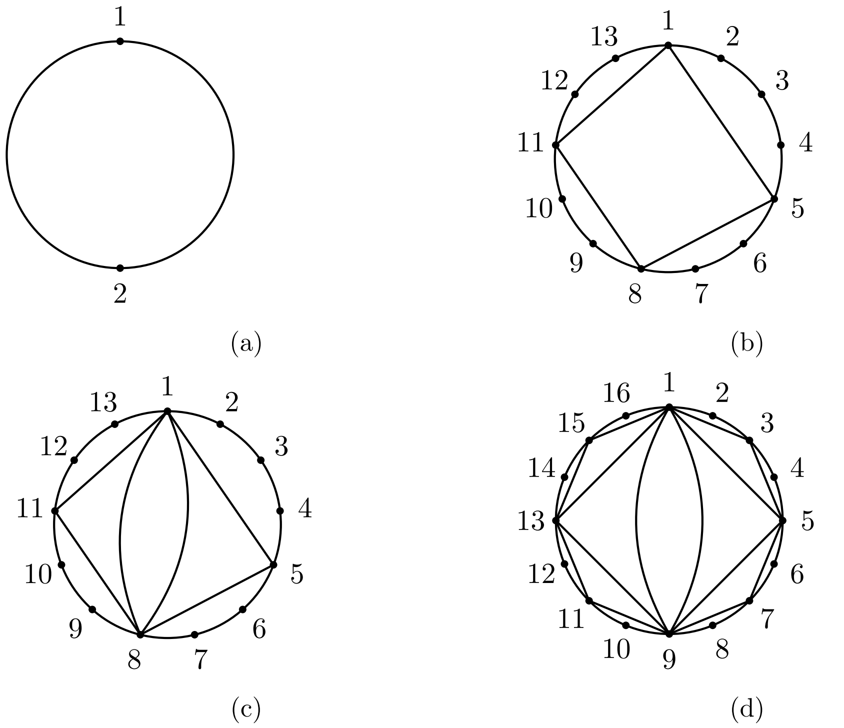


Figure 4.2: London's (2012) representation for the four meters from Table 4.1.

less definitions than Forth's formalization. Furthermore, the current approach does not allow microrhythmic details into the formalization, in contrast to Forth. This restricts metrical space.

Now that we have defined a meter, let us turn to well-formedness of meters. Well-formedness is defined by meeting eight *constraints*, **C1** to **C8**. Basic structural constraints **C1**, **C2** and **C3** ensure proper hierarchical organization of the meter. Temporal constraints **C4**, **C5** and **C6** place perceptual boundaries on the temporal unit u. Advanced structural constraints **C7** and **C8** restrict the possible relations between beats, both within and between levels. We say that a meter is *well-formed* if and only if the constraints **C1** to **C8** all hold:

C1 Existence of a lowest level.
$\exists \ell_0 \in L \; \exists p \in \mathbb{N} \; [L_0 = \{n \in \mathbb{N} \mid n < p\}]$

The meter M has a *lowest level*, which is the zeroth level ℓ_0, that contains a given number p of beats n, such that the beats of this level are represented by consecutive natural numbers $0, 1, 2, ..., p-1$. We call p the *period* of the meter and it denotes the total number of beats in the meter (analogous to the *cardinality* in London, 2012). From this constraint, it directly follows that $p = \#\ell_0$. The lowest level ℓ_0 can be compared to the *N cycle* in London (2012) and is essentially the common fast pulse of the model, as it is the fastest level and it is defined as a series of isochronous units. Through this, constraint **C1** fills the open end of isochrony of the fastest pulse, as discussed in Section 3.3.1.

By requiring the existence of a lowest level, **C1** ensures that no structure is empty in terms of levels. By the requirement that this lowest level is the zeroth level, ℓ_0, there is by definition 2 no level that contains any beats which are not in the lowest level ℓ_0. Hence, we can unambiguously refer to ℓ_0 as the lowest level.

C2 The highest level contains only the first beat (the number 0).
$\ell_k = \{0\}$

The *highest level* ℓ_k (with $k = \#L - 1$ by definition 1) contains exactly one beat, which is the natural number 0. By this constraint in combination with definition 2, it holds that 0 is a beat on all levels and it is the only beat that exists on all levels. Furthermore, by definition 3, it is the zeroth beat of all levels. Because of these reasons, we call the beat represented by the natural number 0 the *downbeat* of the meter.

Defining a downbeat in the formalization is a divergence from London's (2012) well-formedness constraints. The motivation for this addition has been discussed in Section 3.3.3. Definition of the downbeat (or rather, the constraint that the highest level contains only the downbeat) has no implications on other constraints, except for **C3** (for this implication, see the explanation at **C3**). However, it does have structural implications. In combination with definition 2, **C2** ensures that the downbeat is in all levels, and therefore all levels coordinate. Also, if a level has only one beat, then by **C2** with definition 2, there is no level above this level. This ensures hierarchical organization of meters.

C3 Existence of the tactus level.
$c < \#L - 1$

The tactus level, ℓ_c, is one of the levels in the structure L and it is not the highest level ℓ_k. This latter aspect is not found explicitly in London's (2012) well-formedness constraints, due to the fact that in London (2012), it is not clear whether there is a highest level that only contains a downbeat. As we have just seen, there is such a level in the current formalization by **C2**. Given this fact, London's (2012) well-formedness constraint 2.2 ("the beat cycle must involve at least two beats", p. 92) becomes equivalent with the statement that the tactus level is not the highest level. This is by the following reasoning: If the tactus level has at least two beats, it cannot be the highest level, as the highest level has only one beat by **C2**. Conversely, the highest level is strictly included all levels below by definition 2, so if the tactus level is not the highest level, it is below the highest level and must therefore have strictly more than one beat.

Similarly, as the structure has a lowest level by **C1** (which is by definition 2 either the tactus level or a strict superset of it), the lowest level contains at least two beats, which entails that for every structure L, $p \geq 2$. This also entails that every structure has at least two levels, since if it were the case that $\#L \leq 1$, then $\#L - 1 \leq 0$ such that $c < 0$ while $c \in \mathbb{N}$. This property of the structure is not defined in London (2012) either, but follows again from the requirement that the highest level the contains only the downbeat.

Before we formulate constraints **C4** to **C8**, we introduce some more definitions that represent different notions of distance between the beats of a meter. In order to more straightforwardly define one of these notions, we first define the notion of *metrical position.*

Definition 4. Let $n \in \ell_x$ be a beat such that $\ell_x \in L$ is level with index x. For these n, x, we define the *metrical position* of beat n on level ℓ_x as follows:
$\mathrm{mp}(n, x) = \#\{m \mid m \in \ell_x \wedge m < n\}$

The metrical position of a beat n at a given level ℓ_x is equal to the number of beats that preceded it at that same level. The metrical position indicates the order number (zeroth, first, second, etc.) of the beat on that level.

An example of the indication of metrical position is shown in Figure 4.3. In this structure, beat 5 (indicated by the arrow) is preceded by five beats on level ℓ_0, and therefore its metrical position on level ℓ_0 is 5; it is the fifth beat (counting from zero) in ℓ_0. In contrast, on ℓ_1, beat 5 is preceded by only two beats, hence its metrical position on level ℓ_1 is 2.

Now, we can define different notions of distance. Henceforth, the term *distance* in general refers to the difference between any two beats (or, more correctly, beat *onsets*) on a given level. This difference can be measured in three ways.

First, it can be measured as the absolute time that elapses between the two beats, in milliseconds. This time span is expressed by a real number and is independent of the level on which we are looking. We call this time span the *absolute distance* and it will recur in definition 9.

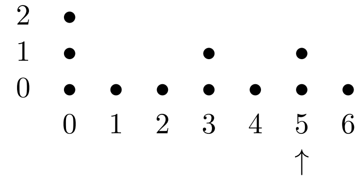


Figure 4.3: Beat 5 highlighted in a meter with structure 3-2-2; $M = (\{\{0, 1, 2, 3, 4, 5, 6\}, \{0, 3, 5\}, \{0\}\}, c, u)$, with c and u omitted for readability.

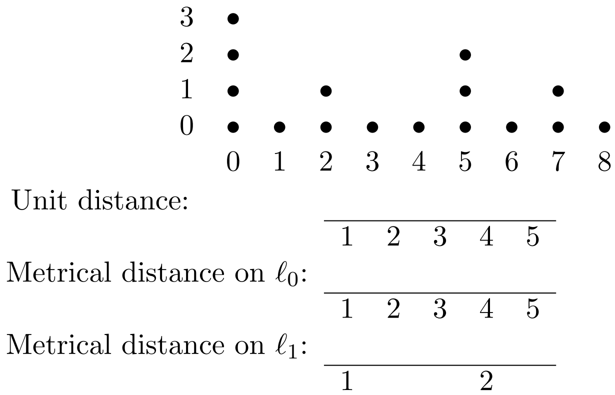


Figure 4.4: Unit distance and metrical distance between beats 2 and 7 in a 2-3|2-2 meter; $M = (\{\{0, 1, 2, 3, 4, 5, 6, 7, 8\}, \{0, 2, 5, 7\}, \{0, 5\}, \{0\}\}, c, u)$, with c and u omitted for readability.

Second, the difference can be measured as the number of *unit intervals*, or simply *units*, between the two beats. In this notion, one unit is the duration between two adjacent beats on the lowest level (note that one unit is equal to u, the temporal unit). As we are looking at the *number of* units, it does not matter what the value of the temporal unit is. Hence, this difference is expressed by a natural number that represents a multiple of the temporal unit. We call this difference the *unit distance*, and like absolute distance, unit distance is independent of level. Unit distance will be formally defined in definition 5.

Third, the difference can be measured as the number of *beat intervals* between the two beats on the level in question. In this notion, one beat interval is the duration between two beats that are adjacent on the level in question. As we are looking at the *number of* beat intervals, it does not matter what the durations of these intervals are in terms of unit distance or absolute distance. Instead, we are looking at the difference in the metrical position of the beats on the level in question, expressed relatively to the other beats on that same level. We call this difference the *metrical distance*. This distance will be formally defined in definition 6. The distinction between unit distance and metrical distance is the formal version of London's (2012) distinction between specific and generic interval size, respectively, as discussed in Section 2.4.3. Figure 4.4 contains an example of this distinction, by showing the metrical distance between beats 2 and 7 on both ℓ_0 and ℓ_1 in a meter with structure 2-3|2-2, as well as the unit distance between these beats. The following two formal definitions will explain how exactly these distances are calculated.

Note that in general, the current proposal differentiates distance and interval as follows: *distance* denotes the difference between any two beats, while *interval* refers to the difference between two beats that are adjacent with regard to the definition in question.

Now, we introduce formal definition of unit distance and metrical distance:

Definition 5. Let $m, n \in \ell_x$ be beats such that $\ell_x \in L$ is a level with index x. For these m, n, we define the *circular unit distance* between m and n as follows:
$\mathrm{UD}(m, n) = (n - m) \bmod p$

The circular unit distance between beat m and n is the difference (in units) between m and n under modulo of the period of the meter. We define the unit distance in a *circular* way, since meter is periodic and circular in London's (2012) theory and visual representation. Circularity is ensured by the modulo operator. As we do not define a non-circular unit distance, we will hereafter refer to the *circular unit distance* simply as the *unit distance*.

An example of calculating unit distance is shown in Figure 4.4. The indicated unit distance (that is, the unit distance between beats 2 and 7) is 5 units: $\mathrm{UD}(2, 7) = (7 - 2) \bmod 9 = 5$. In contrast, the unit distance between beats 7 and 2 is 4 units: $\mathrm{UD}(7, 2) = (2 - 7) \bmod 9 = (-5) \bmod 9 = 4$.

Note that 'mod' in this definition refers the binary operator that finds the remainder of *Euclidean division* or *floored division*, such that, for instance, $\mathrm{UD}(7, 2)$ with period 9 is equal to 4 and not to -5, as is the case for *truncated division*. As argued above, unit distance is independent of level, and indeed, it is not measured with respect to any level. The unit distance between two beats m and n is 0 if and only if m and n are the same beat. One might intuitively think that the distance between two the same beats is a full period p, but by the modulo operator in the definition, this is 0. This has potentially counterintuitive consequences for definition 7, but we will see that these do not give rise to semantic problems.

Definition 6. Let $m, n \in \ell_x$ be beats such that $\ell_x \in L$ is a level with level index x. For these m, n, x, we define the *circular metrical distance* between m and n on level ℓ_x as follows:
$\mathrm{MD}_x(m, n) = (\mathrm{mp}(n, x) - \mathrm{mp}(m, x)) \bmod \#\ell_x$

The circular metrical distance between beats m and n on a given level ℓ_x is the difference between the metrical positions of m and n on that level, under modulo of the total number of beats in that level. Like unit distance, metrical distance is defined in a circular way, which is again shown by the modulo operator. As is the case of unit distance, hereafter we will refer to the circular metrical distance simply as the *metrical distance*.

An example of calculating metrical distance is shown in Figure 4.4. The indicated metrical distance on level ℓ_1 (that is, the metrical distance between beats 2 and 7 on ℓ_1) is 2 beat intervals, as the metrical position of beat 7 on ℓ_1 is 3, and the metrical position of beat 2 on ℓ_1 is 1. In contrast, the indicated metrical distance on ℓ_0 (that is, the metrical distance between beats 2 and 7 on ℓ_0) is 5 beat intervals: as the metrical positions of beats 2 and 7 on ℓ_0 are 2 and 7, respectively.

In contrast to unit distance, metrical distance is dependent on the level: it looks at the *position* (or order number) of the beats in that level relative to the other beats. Note that the metrical distance between two beats on a given level is 1 if and only if the beats are adjacent on that level. Furthermore, the metrical distance between two beats is 0 on all levels if and only if the two beats in question are the same. Again, this has potentially counterintuitive consequences for definition 7, which again do not give rise to problems. Finally, the metrical distance between two beats on level ℓ_0 is equal to their unit distance (see also Figure 4.4).

The next definition is necessary for the formal definition of maximal evenness (definition 11). It will come in handy as well for other constraints, such **C6** (temporal limitation of the tactus level) and later additional constraints in Section 4.3, such as explicit exclusion of metrical ambiguity, and restriction of the number of beat classes.

Definition 7. For two variables x, d such that $\ell_x \in L$ is a level and $d \in \mathbb{N}$, we define the *unit distance set* $\mathcal{Q}(x, d)$ as follows:
$q \in \mathcal{Q}(x, d) \iff \exists m, n \in \ell_x \text{ s.t. } [\mathrm{MD}_x(m, n) = d \wedge \mathrm{UD}(m, n) = q]$

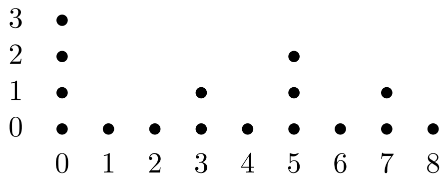


Figure 4.5: Meter with structure 3-2|2-2; $M = (\{\{0,1,2,3,4,5,6,7,8\},\{0,3,5,7\},\{0,5\},\{0\}\},c,u)$, with c and u omitted for readability.

The unit distance set contains the *unit distances* between all beat pairs m, n on level ℓ_x that are separated by *metrical distance* d. In other words, $\mathcal{Q}(x, d)$ is the set of the unit distances q between every pair $m, n \in \ell_x$ between which the metrical distance is equal to predefined variable d. For example, the unit distance set of level ℓ_1 in a meter with a 3-2|2-2 structure (see Figure 4.5) for metrical distance $d = 1$ is the set that contains the unit distances between the beat pairs on level ℓ_1 of this structure between which the metrical distance is 1 (that is, they are adjacent on level ℓ_1). These are the pairs $(0, 3)$, $(3, 5)$, $(5, 7)$ and $(7, 0)$, between which the unit distances are respectively 3, 2, 2 and 2. Therefore, $\mathcal{Q}(1, 1) = \{2, 3\}$. Likewise, the unit distance set for level ℓ_1 with metrical distance $d = 2$ contains the unit distances between the pairs $(0, 5)$, $(3, 7)$, $(5, 0)$ and $(7, 3)$, which are respectively 5, 4, 4 and 5. Therefore, $\mathcal{Q}(1, 2) = \{4, 5\}$.

The definition of the unit distance set is inspired by the work of Clough and Douthett (1991, p. 94–97) on maximal evenness in scale theory. It is defined for every level ℓ_x in the structure and every $d \in \mathbb{N}$, but is non-empty only for $d < \#\ell_x$, as the metrical distance between two beats on a level never exceeds the total number of beats of that level. In fact, given that ℓ_x is a level in the structure, the unit distance set is *always* non-empty for every $d < \#\ell_x$, because one can always find a pair of beats m, n on level ℓ_x such that the metrical distance between m and n on ℓ_x is equal to any number between 0 and $\#\ell_x$.

Note that for all levels ℓ_x, the unit distance set for $d = 0$ is $\{0\}$, because in this case, every $m = n$, so the unit distance between all beats m and n is 0. Also, for all levels ℓ_x, the unit distance set for $d = 1$ is the set of all unit intervals between beats that are adjacent on that level, which is used in definition 9 (and thereby in constraint **C6**). Furthermore, the unit distance set for the lowest level ℓ_0 and $d = 1$ is $\{1\}$, as on the lowest level, there is one unit between each pair of adjacent beats. In general, for all $d \leq \#\ell_0 - 1$ it holds that the unit distance set on level ℓ_0 is $\{d\}$, as on the lowest level, there are d units between each pair of beats between which the metrical distance is d.

For all d, the unit distance set for the highest level ℓ_k is either $\{0\}$ or the empty set, since the highest level contains only one beat, and both unit distance and metrical distance between the same beat are 0 by definition 5 and 6, respectively. One might intuitively think that the unit distance set is $\{p\}$ for this case, by assigning the case $d = 1$ to the idea m and n being two 'adjacent downbeats'. However, as metrical distance is circular, we cannot speak of 'adjacent downbeats'. Therefore, there is no case $d = 1$, hence the unit distance set is empty for this metrical distance. Similarly, for all levels ℓ_x, the unit distance set is empty when $d \geq \#\ell_x$.

The unit distance set provides a measure of beat distribution for a given level ℓ_x, as it lists all the unit distances that exist between given pairs of beats (the choice of pairs depends on d). For all d, the set has one element when the beats are evenly distributed on that level (i.e., all beats are isochronous on that level), while it has multiple values when the beats are unevenly distributed (except for $d = 0$ as above). The actual number of elements in the set and their mutual differences indicate the extent to which the beats on ℓ_x are distributed unevenly.

The next two definitions are important with regard to absolute durations in milliseconds instead of units or position in the level.

Definition 8. We define the *absolute period* a of a meter M as $a := p * u$

We get the absolute period by multiplying the period, p, which is measured in units, by the value of the temporal unit, u, measured in milliseconds. It follows that $a \in \mathbb{R}$ and is measured in milliseconds.

Definition 9. For every variable x such that $\ell_x \in L$ is a level, we define the *absolute beat interval set* as follows:
$\mathcal{I}(x) = \{i \mid i = q * u \wedge q \in \mathcal{Q}(x, 1)\}$

The absolute beat interval set contains the values of the unit distance set for level ℓ_x with $d = 1$, multiplied by the value of the temporal unit, u. Thereby, it returns the set of *absolute distances* (in ms) between the pairs of beats that are adjacent on that level, that is, it returns the *absolute beat intervals* on that level. For an example, see Figure 4.6. The absolute beat interval set $\mathcal{I}(x)$ of level ℓ_1 for this meter is $\mathcal{I}(1) = \{440 \text{ ms}, 660 \text{ ms}\}$.

Note that in particular, the absolute beat interval set for the lowest level is $\{u \text{ ms}\}$, as the lowest level is isochronous and all absolute beat intervals have a duration of u ms. The absolute beat interval set of the highest level is $\{0 \text{ ms}\}$, as there is only one beat on the highest level, and both metrical distance and unit distance are circular.

Before we can define the temporal constraints **C4**, **C5** and **C6**, we will define four parameters. As discussed in Section 2.4.3, London (2012) puts perceptual temporal constraints on the tempo-metrical types. These constraints are obtained from empirical research, in which high degrees of task-dependency and inter-subject variation were found (London, 2012, p. 27). Because of this, we define these boundaries as parameters of the formalization, which can be tweaked to suit individual (or cross-cultural) differences.

Definition 10. We define the following four *perceptual limit parameters*:

- The minimum unit interval, u_{min}, in ms;
- The maximum absolute period, a_{max}, in ms;
- The minimum tactus beat interval, i_{min}, in ms;
- The maximum tactus beat interval i_{max}, in ms.

For the current project, we choose the values as defined in London's (2012) well-formedness constraints, which are the following: $u_{\text{min}} = 100$ ms, $a_{\text{max}} = 5000$ ms, $i_{\text{min}} = 400$ ms, $i_{\text{max}} = 1200$ ms.

Definition of perceptual limits as parameters is another divergence from the formalization of Forth (2012). Through this parameter definition, the perceptual limits can take any value in the current formalization, while in Forth, these perceptual limits are fixed in the definitions (pp. 110, 112, 119).

The temporal constraints **C4**, **C5** and **C6** are derived from the temporal constraints in London (2012):

C4 Unit interval limitation.
$u \geq u_{\text{min}}$

The value of the temporal unit u is greater than the minimum unit interval u_{min}. This constraint defines the duration of the metric floor, as the shortest beat intervals cannot be perceived as such when the tempo is too high.

C5 Absolute period limitation.
$a \leq a_{\text{max}}$

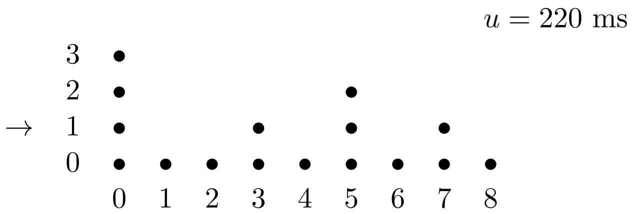


Figure 4.6: Dot representation for meter with structure 3-2|2-2; $M = (\{\{0,1,2,3,4,5,6,7,8\},\{0,3,5,7\},\{0,5\},\{0\}\},1,220)$.

The absolute period a of the meter is less than the maximum absolute period a_{max}. This constraint defines the duration of the metric ceiling, as metrical entrainment is limited to periods shorter than a_{max}.

If we have $a_{\text{max}} = 5000$ ms and $u_{\text{min}} = 100$ ms, then $p \leq 50$, since $p = \frac{a}{u}$ by definition 8. We can conclude that for these values of a_{max} and u_{min}, p is constrained by structural constraint **C3** and temporal constraint **C5** to be $2 \leq p \leq 50$.

C6 Temporal tactus level limitation.
$\forall i \in \mathcal{I}(c)\ [i_{\text{min}} \leq i \leq i_{\text{max}}]$

All absolute beat intervals on the tactus level are in between the minimum and maximum tactus beat interval i_{min} and i_{max}. This way, no beat in the tactus level is too fast or too slow to be perceived as a tactus. As the tactus level may be non-isochronous, there might be multiple beat intervals on the tactus level, which all need to be taken into account.

For an example, see Figure 4.6. The absolute beat interval set $\mathcal{I}(x)$ of the tactus level ℓ_c for this meter is $\mathcal{I}(c) = \{440 \text{ ms}, 660 \text{ ms}\}$, which is within the required range, hence this meter meets constraint **C6**. Note that not every level ℓ_x in which the absolute beat intervals fall within the range required by **C6**, is 'a tactus level'; by definition 1, there is only one tactus level ℓ_c, as the tactus level is determined by the constant c. We can see this for the current example (Figure 4.6): $\mathcal{I}(2) = \{880 \text{ ms}, 1100 \text{ ms}\}$, which falls within the required range, but $c \neq 2$. What this does mean, is that **C6** does not only permit the meter with structure 3-2|2-2, tactus level index $c = 1$ and temporal unit $u = 220$ ms, but also permits a different meter with the same structure and temporal unit, but with ℓ_2 as the tactus level (notwithstanding that this meter might violate one of the other constraints).

Dependent on the particular structure L of a meter, temporal constraints **C4**, **C5** and **C6** restrict the possible values of u such that the meter is still well-formed. For example, a meter 3-2|2-2 with $c = 1$ (Figure 4.6) is ill-formed for $u > 400$ ms, since then, one of the elements of the absolute beat interval set, $\mathcal{I}(c)$, exceeds $i_{\text{max}} = 1200$ ms. A meter with structure 3-2|2-2 is also ill-formed for $u < 200$ ms, as this makes one of the elements of the absolute beat interval set less than $i_{\text{min}} = 400$ ms. Often, the i_{min} and i_{max} are the most restricting parameters, but for meters with lengthy subdivision patterns or large periods, u_{min} and a_{max} respectively come into play as well.

As mentioned earlier this section, the last two constraints, **C7** and **C8**, restrict the possible relations between beats, both within and between levels:

C7 Principle of non-adjacency.
$\forall \ell_x \in L\ \forall m, n \in \ell_x\ [\text{MD}_x(m,n) = 1 \rightarrow (m \notin \ell_{x+1} \vee n \notin \ell_{x+1})]$

Every two adjacent beats m and n on a given level ℓ_x are not both in the level directly above ℓ_x (recall that the metrical distance between two beats on a given level is 1 if and only if the beats are adjacent on that level). This constraint ensures that any two beats on a given level are not adjacent on the level below, as motivated by London's (2012) well-formedness constraint 3.4. London argued that this constraint ensures that beats and subdivisions on the same cycle are not mixed up.

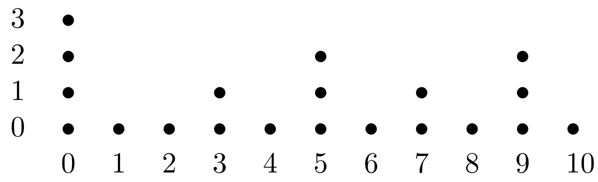


Figure 4.7: Dot representation for an ill-formed meter with structure 3-2|2-2|2; $M = (\{\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}, \{0, 3, 5, 7, 9\}, \{0, 5, 9\}, \{0\}\}, c, u)$, with c and u omitted for readability.

For an example, see Figure 4.7. This is an ill-formed meter, since we can find the following violation to **C7**: there is a level ℓ_x (namely, ℓ_1) such that there is a beat m and a beat n in ℓ_x (namely, the beats 9 and 0) that are adjacent on level ℓ_1 (i.e., $\text{MD}_1(9, 0) = 1$) and are both also a beat in ℓ_2.

Note that **C7** is always met on the highest level ℓ_k, as there is no ℓ_{k+1}, so both arguments of the disjunction in **C7** are true, which makes the conclusion of the implication true.

In combination with definition 2, **C7** requires that every level contains at most half of the beats of the level directly below (as two adjacent beats are never allowed to both be in the next level). This puts constraints on the maximum number of levels in a structure. For $u_{\text{min}} = 100$ ms and $a_{\text{max}} = 5000$ ms, we had $p \leq 50$. Therefore, $\#\ell_0 \leq 50 \Rightarrow \#\ell_1 \leq 25 \Rightarrow \#\ell_2 \leq 12 \Rightarrow \#\ell_3 \leq 6 \Rightarrow \#\ell_4 \leq 3 \Rightarrow \#\ell_5 \leq 1$. But if a level has only one beat, then by **C2** with definition 2, there is no level above this level. So there cannot be a level ℓ_6 under the current perceptual limit parameter settings, and therefore, all structures have six (ℓ_0 to ℓ_5) or less levels. When the perceptual limit parameters are changed, there may be another maximum value for the period p. In this case, there may be another maximum number of levels as well; these values are flexible. Note that this again diverges from the formalization of Forth (2012). Not only the perceptual limits are fixed in the formalization of Forth, as discussed above, but also the maximum number of levels (see Forth, 2012, p. 114).

Before we can define the last constraint **C8** of maximal evenness, we first define maximal evenness as a property of a level.

Definition 11. A level ℓ_x is *maximally even* iff $\#\ell_x \geq 2$ and $\forall d \in \mathbb{N}\ \forall q, r \in \mathcal{Q}(x, d)\ [|q - r| \leq 1]$

A level is *maximally even* if and only if it has at least two beats, and for every metrical distance d, the difference between all elements of its unit distance set is at most 1. For an example, see Figure 4.8. Level ℓ_1 of this structure is not maximally even, because there is a metrical distance d (namely, 2), such that there are unit distances q and r in the unit distance set of level ℓ_1 between which the difference is greater than 1 (namely, 4 and 6, as $\mathcal{Q}(1, 2) = \{4, 5, 6\}$). Note that ℓ_2 however is maximally even, since it has two beats, the unit distance set is for $d = 0$ is $\{0\}$, the unit distance set for $d = 1$ is $\{6\}$ and the unit distance set for all $d \geq 2$ is empty.

Definition 11 entails that the unit distance set for any level is not allowed to have more than two elements, since if it does, there are unit distances q and r in the unit distance set such that $|q - r| > 1$. Note that the set *is* allowed to have one element, since then $q = r$, so the difference between all unit distances q and r is 0.

The latter part of this definition is motivated by the formal definition of maximal evenness in Clough and Douthett (1991, p. 96). In the current formalization, there is an additional requirement that a level needs to have at least two beats, since maximal evenness is a measure of beat distribution, and it makes no sense for a level with only one beat to call its 'beat distribution' maximally even. Also, such a level cannot serve to stabilize a tactus level that is not maximally even, so this requirement is also important with regard to the next constraint. Note that proving maximal evenness for a given level is not needed for all $d \in \mathbb{N}$, but only for

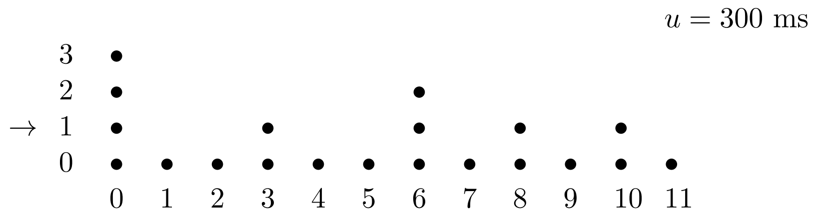


Figure 4.8: Dot representation for meter with structure 3-3|2-2-2; $M = (\{\{0,1,2,3,4,5,6,7,8,9,10,11\},\{0,3,6,8,10\},\{0,6\},\{0\}\},1,300)$.

$1 \leq d \leq \#\ell_x - 1$. This is because of two things. First, recall that the unit distance set for $d = 0$ is $\{0\}$ for all levels ℓ_x, so the difference between all unit distances q and r in the set is 0. Second, for all ℓ_x and $d \geq \#\ell_x$, the unit distance set is empty, so trivially the difference between all unit distances q and r in the unit distance set is at most one, as there are no such q and r. Finally, the lowest level, ℓ_0, is always maximally even, as its unit distance set is $\{d\}$ for all metrical distances $d \leq \#\ell_0 - 1$.

C8 Principle of maximal evenness.
ℓ_c is maximally even or ℓ_{c+1} is maximally even.

Either the tactus level, ℓ_c, or the level directly above the tactus level, ℓ_{c+1}, is maximally even, or both are. Level ℓ_{c+1} exists by **C2**, which requires that the tactus level ℓ_c is not the highest level. Since ℓ_{c+1} often recurs in later discussion, we will hereafter refer to it as the super-tactus level.

If a structure abide by the principle of maximal evenness, we call this structure as a whole maximally even, as it is maximally even in the global sense used by London (2012). For example, the meter in Figure 4.8 is maximally even, since its super-tactus level is maximally even.

Because every maximally even level has at least two beats by definition 11, **C8** indirectly puts an additional restriction on the the tactus level, namely, it has at least two beats. However, this is not a problem, as this is also required in London's (2012) well-formedness constraint 2.2, and we have seen that it already followed from **C2** as well. Note that this does not make **C2** unnecessary, as it also has other structural functions, such as ensuring that the tactus level ℓ_c is one of the levels in the structure and ensuring that the downbeat is a beat on all levels.

Finally, we define *metrical saliency.*

Definition 12. Over ℓ_0, we define the *metrical saliency* of every beat $n \in \ell_0$ as follows:
$S(n) = \#\{x \in \mathbb{N} \mid \ell_x \in L \wedge n \in \ell_x\}$

The metrical saliency of a beat is the number of levels in which it appears. In this definition, the downbeat is the most salient beat of the meter, as by **C2**, it is the only beat that exists on all levels of a meter. This definition is motivated by Section 3.3.3 and resembles the definitions of *strength* in Lerdahl and Jackendoff (1983) and *weight* in Longuet-Higgins and Lee (1984), as both defined for isochronous meters.

4.2 Derivation of London's well-formedness constraints from the formalization

In this section, we will construct proofs of London's (2012) well-formedness constraints, taking the definitions and constraints of the current formalization as premises. Note that the proofs cannot be fully formal, as London's theory is not formal. The definitions and constraints are reproduced from London, pp. 191–192.

“*Definition.* A meter is a coordinated set of periodic temporal cycles of sensorimotor attention.”
As can be seen, the definition of meter in the current formalization is different. However, the current formalization can still be used to describe a coordinated set of periodic temporal cycles of sensorimotor attention. The levels in the current representation are the periodic temporal cycles in London (2012) and they are temporal through temporal unit u. The structure in the current representation represents the full coordinated set of cycles in London (2012) and the tactus level is the beat cycle. Furthermore, by the temporal unit, different tempo-metrical types of the same metrical type can be differentiated.

“*Definition.* The lowest level of the metrical cycle involves the shortest IOI and includes all of the attentional peaks in the cycle; it is referred to as the *N cycle*, as N = number of attentional peaks present.”
The N cycle of London (2012) is the zeroth level in the current formalization. Its existence is ensured by **C1**. By definition 2, this level contains all attentional peaks of the structure.

The fact that ℓ_0 involves the shortest inter-onset intervals can be drawn by computing the absolute beat interval set (definition 9) of all levels. As the unit distances between adjacent beats on any level ℓ_x such that $x \neq 0$ are multiples of unit distances between adjacent beats on ℓ_0 through definition 2, the absolute inter-onset intervals of beats on higher levels are also multiples of the absolute intervals on ℓ_0. Therefore, the elements of absolute beat interval set of ℓ_0 are smaller than all elements of the absolute beat interval set for any other level ℓ_x.

“*Definition.* Higher levels, involving larger IOIs than those present in the N cycle, are referred to as subcycles.”
Exactly all levels ℓ_x such that $x \neq 0$ are subcycles in this definition and by the same argumentation as above, they all involve larger inter-onset intervals.

“WFC 1.1: IOIs between attentional peaks on the N cycle must be greater than $\approx$ 100 ms.”
This follows directly from **C4**. The exact value is made into a single threshold value, but is flexible through its definition as a parameter in definition 11.

“WFC 1.2: The *beat cycle* involves those attentional peaks whose IOIs fall between $\approx$ 400 ms and $\approx$ 1200 ms.”
This follows directly from **C6**. The remarks at the previous well-formedness constraint apply here as well.

“WFC 1.3: A meter may have only one beat cycle.”
For every meter, its tactus level is defined by constant c which can have only one value at a time. But as all ℓ_x within a meter are unique by definition 2, there can be only one ℓ_x such that $x = c$ and therefore only one tactus level for every meter. As explained at **C6**, there can be multiple levels that *may* serve as tactus level for a given structure and given temporal unit, but any meter can have only one tactus at a time.

“WFC 1.4: The maximum duration for any or all cycles is $\approx$ 5000 ms.”
This follows directly from **C5**. The remarks for well-formedness constraint 1.1 apply here as well. Note that all levels should have the same duration for this constraint to hold, but this follows from the fact that all levels coordinate, as 0 is a beat on all levels by the combination of **C2** and definition 2.

“WFC 2.1: A meter must have a beat cycle.”

This follows directly from **C3**, which ensures that the tactus level constant c denotes a level in the structure.

"WFC 2.2: The beat cycle must involve at least two beats."
This follows from **C3** as well, as argued below its definition: By definition 2, the highest level is strictly included in all levels below, and by **C3**, the tactus level is not the highest level. Therefore, the tactus level is below the highest level, so the highest level (containing exactly one beat by **C2**) is strictly included in the tactus level, which entails that the tactus level must have strictly more than one beat, hence at least two beats.

"WFC 2.3: The N cycle may serve as the beat cycle."
This is a possibility rather than a restriction and there is no rule in the formalization that forbids this. This is the case $c = 0$ and we can create meters for which this holds, as long as that is permissible through the temporal constraints **C4**, **C5** and **C6**. In practice, $c = 0$ can only hold when $p \leq 12$, as for $p = 13$, **C5** requires that $u \leq 384.6$, while since $c = 0$, $u \geq 400$ by **C6**.

"WFC 3.1: All cycles must have the same total period/duration."
This follows from the fact that all levels coordinate, as 0 is a beat on all levels by combination of **C2** and definition 2.

"WFC 3.2: All cycles must be continuous."
As definition 5 and 6 are defined on the full period and circular, there are no discontinuities.

"WFC 3.3: The N cycle and all subcycles must begin and end at the same temporal location: that is, they must all be in phase."
This follows from the fact that 0 is a beat on all levels by combination of **C2** and definition 2.

"WFC 3.4: Each subcycle must connect nonadjacent time points on the next lowest cycle."
This follows directly from **C7**.

"WFCs 4.1.1-3" (omitted - these rules are trivially met as the interonset intervals on the lowest level are always isochronous in this formalization: interonset intervals are absolute distances between adjacent units, all units are consecutive by **C1** and there is only one u per meter by definition 1)

"WFC 4.2.1: If the IOIs on the N cycle are isochronous, then the beat cycle need not be."
Like WFC 2.3, this is a possibility rather than a restriction. As we have seen directly above, the beats on the zeroth level are always isochronous. Indeed, the tactus level need not be isochronous as long as the other constraints are met, and we have seen that we can create meters for which this holds (see for instance the meter in Figure 4.8 that passes **C8**).

"WFC 4.2.2: If the beat cycle is NI, then either (1) it is maximally even or (2) the cycle above the beat cycle, in most cases the half-measure cycle, must be maximally even."
This follows directly from **C8** (with definition 11).

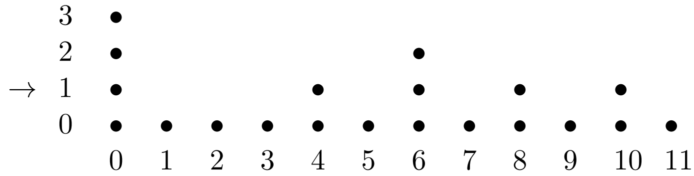


Figure 4.9: Meter with structure 4-2|2-2-2, that is ill-formed with respect to **C9**; $M = (\{\{0,1,2,3,4,5,6,7,8,9,10,11\},\{0,4,6,8,10\},\{0,6\},\{0\}\},1,u)$, with u omitted for readability.

4.3 Additional constraints

As discussed in Section 2.4.4, there are several inconsistencies between London's (2012) well-formedness constraints and theory (see also Section 5.1). Because of this, the current section introduces additional constraints that are not explicitly present in London's (2012) well-formedness constraints. Some of the additional constraints are motivated through personal communication with Justin London (October 30, 2016) about inconsistencies in London (2012). Where this is the case, it will be indicated.

In most cases, the current formalization is weaker than London's (2012) theory: it includes structures that are not allowed by London (2012). In Section 4.3.1, four strengthening constraints are presented that strengthen the formalization in different ways and by different motivations. There are also cases in which the formalization is stronger than London's (2012) theory. For these cases, the principle of maximal evenness (**C8**) is redefined in Section 4.3.2. In future research, some of the additional constraints may be useful for implementation of the formalization in a cognitive model. To this end, effects of the additional constraints will be analyzed per constraint in Chapter 5.3. The analysis will focus on general effects as well as effects on individual structures. Additionally, the analysis will visualize the disagreements between the theory and well-formedness constraints of London (2012).

4.3.1 Strengthening constraints

As we have seen in Section 2.4.3, London's (2012) theory argues for avoiding the metrical phenomena of ambiguity and contradiction within metrical structure (this is confirmed by London, personal communication, October 30, 2016). London's (2012) well-formedness constraints 3.4, 4.1.2, 4.1.3 and 4.2.2 are meant to rule out these phenomena (p. 131), and they partially do (though indirectly), but not on all levels. Therefore, the first additional constraint concerns explicit exclusion of structures in which ambiguity or contradiction occurs on any level:

C9 Exclusion of ambiguity and contradiction.
$\forall \ell_x \in L \; [\max \; \mathcal{Q}(x,1) < 2 * \min \; \mathcal{Q}(x,1)]$

In every level, the longest beat interval is not equal to or longer than twice the shortest beat interval. This constraint compares the minimal and maximal elements of the unit distance set of the adjacent beats (hence $d = 1$) on all levels ℓ_x. We can speak of a minimal and maximal element of the unit distance set, since this set is finite for all levels ℓ_x (every level has a finite number of beats since p is finite by **C5**).

An example of a meter that does not meet this constraint is shown in Figure 4.9. In this meter, the unit distance set of level ℓ_1 is $\{2,4\}$. Note that this meter does meet the structural well-formedness constraints, including **C8**, since the super-tactus level is maximally even.

As stated above, avoiding ambiguity and contradiction on all levels is not directly represented in the well-formedness constraints of London (2012). However, part of it is represented in London's (2012) well-formedness constraint 4.1.2. This constraint avoids ambiguity and contradiction on the lowest level when this level is non-isochronous, by requiring that the length of

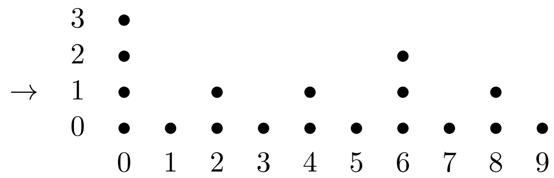


Figure 4.10: Meter with structure 2-2-2|2-2, that is ill-formed with respect to **C10**; $M = (\{\{0,1,2,3,4,5,6,7,8,9\}, \{0,2,4,6,8\}, \{0,6\}, \{0\}\}, 1, u)$, with u omitted for readability.

the shortest beat on this level is more than half the length of the longest beat (London, 2012). However, this well-formedness constraint was not included in the current formalization, as it only holds for cases in which the lowest level is non-isochronous, which is never the case in the current formalization. The constraint therefore has no effect on the current formalization. Furthermore, it only holds for the lowest level, while we have seen that London (2012; personal communication, October 30, 2016) argues for avoiding this restriction on all levels of metrical structure. Therefore, addition of **C9** may be a good complement to the current formalization such that it better reflects London's (2012) theory. The effect of **C9** is discussed in Section 5.3.1.

The second additional constraint was also designed to increase regularity of the possible structures:

C10 Half/third-measure rule.
$\forall \ell_x \in L\ [2 \leq \#\ell_x \leq 3 \rightarrow \forall q, r \in \mathcal{Q}(x,1) : |q - r| \leq 1]$

Any level that has two or three beats (that is, it can serve as a half-measure or third-measure, see Section 2.4.3) should be maximally even. An example of a meter that does not meet this constraint is given in Figure 4.10. In this meter, level ℓ_2 has two beats, but $\mathcal{Q}(2,1) = \{4,6\}$. Note that this meter does meet all other structural well-formedness constraints, including **C9**.

Note that **C10** makes a difference only when the level in question is not the lowest level ℓ_0 or the tactus level ℓ_c: the lowest level ℓ_0 is always maximally even as argued before, and by **C8**, the tactus level is always maximally even when it has two or three beats (as in this case, there cannot be a super-tactus level with more than one beat by **C7**).

C10 is less grounded within the theory of London (2012) and more artificially designed to remove structures that are intuitively wrong, but on which London's (2012) theory does not specify whether they are allowed. The constraint is designed for cases in which the tactus level is maximally even, so by **C8**, the super-tactus level (potentially a half- or third-measure) need not be maximally even. **C10** requires that in this case, any half- or third-measure should still be maximally even. As we will see in Section 5.3.2, **C10** does address problematic cases, but the idea behind it is not as clearly a goal in London's (2012) theory as avoiding ambiguity is. Therefore, **C9** is a more logical extension to London's (2012) well-formedness constraints than **C10**. Note also that the 2-2-2|2-2 meter of Figure 4.10 intuitively is not necessarily problematic, as it is in fact a 3-2 meter with duple subdivision.

The third additional constraint is motivated by personal communication with London (October 30, 2016):

C11 Maximum of two beat classes.
$\forall \ell_x \in L\ [\#\mathcal{Q}(x,1) \leq 2]$

For every level, the unit distance set for metrical distance $d = 1$ has at most two elements. That is, there are maximally two different beat intervals (beat classes) on each level.

An example of a meter that does not meet **C11** is given in Figure 4.11. We can find the following violation for ℓ_1: $\mathcal{Q}(1,1) = \{3,4,5\}$.

C11 follows from London's (personal communication, October 30, 2016) proposal for a "covering law" that there cannot be more than two beat classes. In the current formalization,

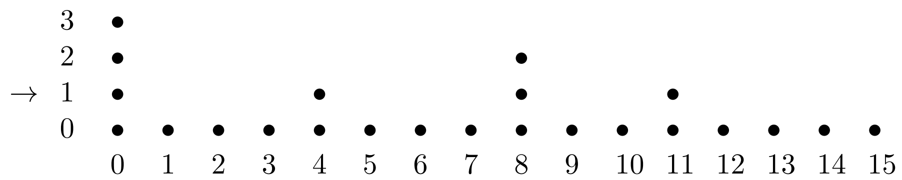


Figure 4.11: Meter with structure 4-4|3-5, that is ill-formed with respect to **C11**; $M = (\{\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}, \{0, 4, 8, 11\}, \{0, 8\}, \{0\}\}, 1, u)$, with u omitted for readability.

we take the strong form of the proposal: there cannot be more than two beat classes *at any level*. An alternative is to apply this rule to the tactus level only. The effect of **C11** is discussed in Section 5.3.3.

Arguably, it is not entirely natural to eliminate structures with three beat classes *a priori*. Therefore, we propose one last additional constraint that eliminates many uncommon structures through a more general consideration:

C12 Meta-rule of temporal frame.
For all well-formed meters $M = (L, c, u)$, there is a meter $M' = (L', c', u')$ and a meter $M'' = (L'', c'', u'')$ such that the following four statements hold:

1. $L = L' = L''$
2. $c = c' = c''$
3. $u' > u'' + 100$ ms
4. M' and M'' are both well-formed.

For every well-formed meter, there are two other well-formed meters that are the same as the meter in question, but their temporal unit values are mutually 100 ms apart. In other words, every well-formed meter must have a temporal frame of 100 ms; there must be a range of 100 ms within which the temporal unit u can move without violating another constraint of the formalization.

This constraint is aimed at all points in which London's (2012) well-formedness constraints are less strict than the theory. It is motivated by the fact that many meters that are irregular in some way (or have large subdivision patterns), can only perceived as that meter within a small temporal frame. The constraint is designed to eliminate these cases by excluding all meters that are only well-formed within a u-range that is smaller than 100 ms (that is, meters that can only exist at very specific tempos). Indeed, a meter that (for a given c) is only allowed when $100 \leq u \leq 102$ arguably does not have a meaning, as the grounds for its perceivability are very limited. This way, constraint **C12** also discards the occurrence of *limit cases*: meters that can only be perceived when u is confined to one specific value. Such limit cases are an artifact of formalization (they are the result of defining hard boundaries) and are therefore undesirable, though useful to visualize extremities.

Intuitively, it is straightforward to eliminate structures with a small temporal range, but mathematically, this has to be defined on a meta-level. In this sense, **C12** is different from the other constraints. Note that the fixed range of 100 ms could be changed into a parameter as well (Section 5.3.4 discusses the choice for the 100 ms range).

Alternative to introducing the extra constraint **C12**, it is also an option to tweak the values of the perceptual limit parameters such that they include less meters. However, in that case the undesirable presence of limit cases and other extreme cases (structures that are only

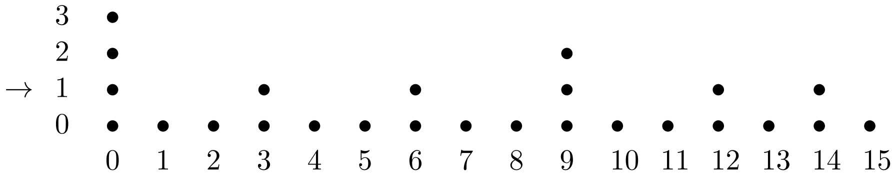


Figure 4.12: Meter with structure 3-3-3|3-2-2, that is ill-formed with respect to **C8**, but well-formed with respect to **C13**; $M = (\{\{0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15\},\{0,3,6,9,12,14\},\{0,9\},\{0\}\},1,u)$, with u omitted for readability.

allowed within a very specific temporal frame) will be maintained. **C12** avoids this problem. Furthermore, in Section 5.3.4, we will see that **C12** has a large effect on metrical space.

4.3.2 A weakening constraint

The theory of London (2012) sometimes allows for meters that are not maximally even, but in which there is a half-measure that abides by another principle, the *principle of rhythmic oddity*, as a compensation. This principle refers to the finding of Arom (1991) that central African musical practice often includes odd half-measures for even N cycles (London, 2012). More specifically, these half-measures follow the rule that their beat intervals arrange as follows: $N/2+1$ versus $N/2-1$. Some meters for which this holds, are permitted by London's (2012) well-formedness constraints (and thereby also by the formalization), because they also have a maximally even beat cycle. However, London's (2012) well-formedness constraints do not permit meters that abide by the principle of rhythmic oddity but do not have a maximally even beat cycle (because these meters are not maximally even by well-formedness constraint 4.2.2). Therefore, we define one weakening constraint that relaxes the principle of maximal evenness and makes use of the principle of rhythmic oddity. This constraint is specifically designed to include meters that do not have a maximally even tactus level, but do abide by the principle of rhythmic oddity. For this, we first formally define rhythmic oddity.

Definition 13. A level ℓ_x is *rhythmically odd* iff $\#\ell_x = 2$ and $\forall q,r \in \mathcal{Q}(x,1)\ [|q-r| \leq 2]$.

A level ℓ_x is *rhythmically odd* if it is a half-measure (i.e., has exactly two beats) on which the difference between the unit distances is at most 2. An example of a meter with a level that is rhythmically odd is shown in Figure 4.12. In the structure of this meter, level ℓ_2 has two beats and the unit distances on this level are 9 and 7.

Now, we can define constraint **C13** as an adjustment of **C8** that relaxes it to include the concerned meters.

C13 Principle of maximal evenness with rhythmic oddity.
At least one of the following statements holds:

- ℓ_c is maximally even
- ℓ_{c+1} is maximally even
- ℓ_c is rhythmically odd
- ℓ_{c+1} is rhythmically odd

This constraint is a weakened adjustment of **C8** and can therefore replace **C8**. For an example of a meter that does not meet **C8** but does meet **C13**, see again Figure 4.12. This meter fails **C8**, as both the tactus level ℓ_c and super-tactus level ℓ_{c+1} are not maximally even ($\mathcal{Q}(1,2) = \{4,5,6\}$

and $\mathcal{Q}(2,1) = \{7,9\}$). However, this meter meets **C13** as the super-tactus ℓ_{c+1} is rhythmically odd as shown above.

If this constraint is combined with **C10**, the latter constraint restricts exactly the cases in which **C13** relaxes **C8**, so **C10** and **C13** cannot combine meaningfully. The effect of **C13** is discussed in Section 5.3.5.

Chapter 5

Analysis

This chapter contains an analysis of the current formalization of the theory and representation of London (2012). As discussed in Section 2.4.4, London's (2012) theory is not described in formal terms, and several details of the theory are not completely specified. Section 2.4.4 also showed that there are inconsistencies between London's (2012) well-formedness constraints, theory and examples. Section 5.1 will show for which structures and groups of structures the current formalization reveals inconsistencies between London's (2012) well-formedness constraints and theory. In Section 5.2, a large testset of structures is constructed and analyzed, which gives insight into the metrical space associated with the formalization (recall from Section 3.3.1 that metrical space refers to the number of possible meters under a given set of rules or constraints). In Section 5.3, the effect of additional constraints will be investigated on both individual structures and the general testset of Section 5.2. Finally, Section 5.4 provides a summary and recommendations with regard to the additional constraints.

5.1 Individual points of difference

This section discusses both individual structures and groups of structures for which the formalization has revealed that London's (2012) well-formedness constraints are inconsistent with the theory. The inconsistencies are categorized into four main points. Additionally, through personal communication, the current view of Justin London on some of the inconsistencies is taken into account. This discussion includes London's (personal communication, October 30, 2016) remarks whenever they diverge from the original theory.

5.1.1 Length of subdivision patterns

London (2012) states that subdivision patterns are at most seven to nine beats long (p. 132). London (2012) argues that this is due to timing constraints, cues in the musical surface for interpolation of levels and interpolation because of *subjective rhythmization* (the "propensity to impose a sense of accent or grouping on a series of identical tones"; London, 2012, p. 13). As the timing constraints are vague, London (2012, p. 134) sets the "practical limit" of subdivision pattern length to six, but does not strongly commit to this limit. In the discussion of possible meters, London (2012) argues that a 5-7 pattern in a 12-cycle and an 8-8 pattern in a 16-cycle *might* be possible (pp. 155, 161). On the other hand, London (personal communication, October 30, 2016) argues that all structures with subdivision patterns of 6 beats or more are problematic, referring to a study by Repp (2007b). In this numerosity judgment task (judging the number of tones in rapid tone sequences) participants generally enumerated a sequence of up to 5 tones accurately, but their judgments became inaccurate for sequences with 6 tones or more (Repp, 2007b).

The current formalization is less strict with regard to length of subdivision patterns than London's (2012) theory. Through interplay of **C4**, **C5** and **C6**, the formalization allows meters that have subdivision structures of up to twelve elements (e.g., a 11-11 pattern with $c = 1$ and $100 \leq u \leq 109.1$). Even though these structures are only possible within a very restricted temporal frame, they are formally permitted within the formalization. The limit cases are a 12-12 structure for $c = 1$, in which u must be *exactly* 100, and a 12 structure for $c = 0$, in which $400 \leq u \leq 416.7$ (recall that $c = 0$ is not possible with $p \geq 13$).

5.1.2 Number and proportions of beat classes

As discussed in Section 2.4.3, London (2012) argues that some proportions between beat classes cannot be perceived as such, because this would give rise to the metrical phenomena of ambiguity or contradiction, and thereby confusion of levels. Apart from ambiguity and contradiction, other factors interfere as well. London (2012, p. 134) argues that beat class ratios of 2:3 or 3:4 are not problematic, but other ratios of the form $J : J + 1$ (with J a positive integer) tend to approximate the 1:1-ratio to such extent that these proportions will sound like a 1:1 proportion played with rubato. For ratios of the form $J : J + N$ (with $1 < N < J$), London (2012) argues that the long beat tends to decompose into two short beats; 7-4 becomes 3-4-4 or 4-3-4. Ratios of the form $J : J + N$ (with $N \geq J$) give ambiguity or contradiction, which is resolved as follows: in the case $J = N$, we get an isochronous meter (3-6 turns into 3-3-3) and in the case $J < N$ we get a proportion of $J : N - J$ (2-5 turns into 2-2-3) (London, 2012). London (2012) argues that through arithmetic on beat classes (with 8 taken as the maximum beat class length by the argument in the previous subsection), every structure will either be isochronous or have a beat class ratio of the form $J : J + 1$ (pp. 134–135). This way, only 2:3 and 3:4 (apart from 1:1, of course) do not suffer from these perceptual and combinatorial effects (London, 2012, p. 135).

Furthermore, London (2012) argues for a maximum of two different beat classes on every level, as the existence of three beat classes will often give rise to ambiguity because of their mutual proportions: when we have beat classes with mutual ratios 2:3:4, the smallest beat class, 2, could be seen as a subdivision unit of the biggest beat class, 4 (being 2+2). In the case 4:5:6, the 6 easily decomposes into 3+3 such that we get 3:4:5 (London, 2012). In turn, 3:4:5 is problematic as well, but not impossible, and London (2012, pp. 161–162) does give instances of structures with such beat classes. In general, London (2012) sometimes argues for specific instances of meters in which beat class ratios do not abide by the arguments he gives on pp. 134–145.

A governing rule about number and proportions of beat classes can be found in London's (2012) well-formedness constraints, namely in those that concern maximal evenness. These constraints put explicit borders on the permissible proportions of different beats on the same level. By the definition of maximal evenness, the beat cycle is only maximally even if there are at most two different beat classes, which are related as $J : J + 1$ with J a positive integer.

However, the constraints are not complete. As mentioned in Section 2.4.3, London's (2012) well-formedness constraint 4.2.2 is formulated in such a way that it captures that the meter should be 'globally' maximally even, and that a maximally even cycle above the beat cycle stabilizes a beat cycle that is not maximally even. However, this formulation of London's (2012) well-formedness constraint 4.2.2 also entails that ambiguities and contradictions may occur. In the case that the cycle above the beat cycle is maximally even, anything goes for the beat cycle by well-formedness constraint 4.2.2. Vice versa, when the beat cycle *is* maximally even, anything goes for all levels above or below the beat cycle, by the same constraint. This phenomenon is reflected in the formalization as well.

In terms of $J : J + 1$-ratios, the formalization permits meters that have ratios of up to 11:12 (limit case is the meter with structure 11-12, $c = 1$ and $u = 100$). Restrictions are mainly brought up by temporal constraints **C4**, **C5** and **C6**, as in the previous point. This, of course, is consistent with London's (2012) theory. There are also inconsistencies on the side of small integers: London (2012) sometimes abstains from listing structures with 3:4-ratios in the overviews (such as the structures 4-3 or 4-4-3). However, London (personal communication, October 30, 2016) argues that London's (2012) theory is inconsistent here and that all those meters should be allowed, though ratios with higher numbers than 5 might be problematic (as motivated through Repp, 2007b). Through this argument, the maximally permissible ratio of the form $J : J + 1$ is 4:5.

The formalization also permits meters that have $J : J + N$ (with $1 < N < J$) ratios on the tactus level, as long as the super-tactus level is maximally even. Herewith, the formalization allows for instance a 3-5|3-5 structure with $c = 1$ and $133.3 \leq u \leq 240$, but also 5-5-5|5-9 with $c = 1$ and $100 \leq u \leq 133.3$. The temporal frame decreases by increasing ratio, but many of these structures are possible within the formalization. In contrast, London (2012) does not allow for either structure, but in personal communication, London (October 30, 2016) adds that the first structure is allowed.[1] On the other hand, London (2012) argues that 3-5, 4-6-6 and 5-7 are perceivable as well, while neither are maximally even, but in personal communication, London (October 30, 2016) argues against the latter two structures, remaining ambivalent on the 3-5 structure.

Apart from this, the formalization also supports meters with more extreme beat class ratios. The structure 2-4|2-4 with $c = 1$ and $200 \leq u \leq 300$ is allowed for its maximally even super-tactus level, which is of course not allowed in London (2012), as it clearly gives rise to the metrical phenomenon of ambiguity. Limit cases are structures in which the longest beat class is three times the shortest beat class (e.g., 2-2-2-2|2-6 with $c = 1$ and $u = 200$), but these are only limit cases by the fact that **C6** puts restrictions on the absolute beat intervals on the tactus level (and not by considerations on maximal evenness). In all, the formalization enables many structures that are metrically ambiguous or contradictory.

Likewise, the formalization permits meters that have three beat classes, such as 4-4|3-5 with $c = 1$, which London (2012) doubtingly allows. In contrast, London (2012) more firmly allows for a 4-3|4-5 structure, while this structure is not maximally even, as neither the tactus level nor the super-tactus level are maximally even, such that this structure is discarded by **C8** of the formalization. However, in personal communication, London (October 30, 2016) argues against this structure. Instead, London (personal communication, October 30, 2016) proposes a covering law that forbids any structure with three beat classes in general. This covering law was the motivation for **C11** of the current formalization.

In the extreme, the formalization allows for meters that have more than three beat classes, as the following meters are all allowed (although temporal frames decrease): 5-2|4-3 with $c = 1$ and $200 \leq u \leq 240$, 3-4-5|6-7 with $c = 1$ and $133.3 \leq u \leq 171.4$, 8-3-5|4-6-7 with $c = 1$ and $133.3 \leq u \leq 150$, and finally 6-9-10|5-8-4-7 with $c = 1$ and $100 \leq u \leq 102.0$. This proves that London's (2012) well-formedness constraints are not specified sufficiently with regard to number and proportions of beat classes.

5.1.3 Existence and organization of intermediate levels

The theory of London (2012) leaves unspecified whether there are levels that are neither the lowest level nor the tactus level (except for the super-tactus level if the tactus level is not maximally even), henceforth *intermediate levels*. This means that as long as there is a tactus level that is maximally even, it is not relevant whether there is any other level. For instance, London (2012) theory explicitly allows for a 2-2-2-2-2-2 meter, but does not mention the option of a 2-2|2-2|2-2 meter, which is also present in the formalization (for $c = 0$ such that $400 \leq u \leq 416.7$, $c = 1$ such that $200 \leq u \leq 416.7$ and $c = 2$ such that $100 \leq u \leq 300$) and distinct from any similar meter, such as 4-4-4 with $c = 1$ with $100 \leq u \leq 300$.

[1] That is, provided that further subdivision structure does not give rise to ambiguity. However, the idea of subdivision structure with respect to a given beat does not apply to the current formalization, as it refers to approaches in which there is no common fast pulse.

Likewise, London's (2012) theory does not specify how possible intermediate levels are organized. Again, this relates to the way in which London's (2012) well-formedness constraint 4.2.2 is defined. As long as there is a maximally even tactus level, intermediate levels may be organized in any possible way. This is reflected in the formalization, that allows for structures like 2-2-2-2-2-2|2-2 (for $c = 1$ and $200 \leq u \leq 312.5$), even though this meter is metrically ambiguous and does not occur in the list of possible structures with cardinality 16 in London (2012). The same holds for intermediate levels that are not above the tactus level, but in between the lowest level and tactus level, such as the structure 4-2|4-2|4-2 with $c = 2$ and $100 \leq u \leq 200$ is allowed in the formalization). It is not clear whether the subdivisions in this structure give rise to metrical ambiguity.[2]

5.1.4 Relaxed maximal evenness

As we have seen in the previous subsections, London's (2012) well-formedness constraints about maximal evenness are meant to avoid the metrical phenomena of ambiguity and contradiction, but cannot always preclude ambiguity or contradiction, because some maximally even structures do have ambiguities or contradictions. Vice versa, there are structures that are not maximally even, but do not give rise to ambiguity or contradiction. This is for instance the case for the aforementioned 3-5 structure (London, 2012, p. 153). In fact, this structure is rhythmically odd as defined in Section 4.3.2.

Other structures for which this holds are 4-6-6, 5-7 and 4-3|4-5. These structures are all deemed possible by London (2012), even though 4-6-6 does not abide by the principle of rhythmic oddity either. Now, in personal communication, London argues against these structures, but does so because of general reasons (length of subdivision and existence of three beat classes) and not because of considerations on maximal evenness. Indeed, London (personal communication, October 30, 2016) argues that meters may not have to be as maximally even as proposed in London (2012). For instance, London (personal communication, October 30, 2016) proposes that ideas of maximal evenness may be relaxed to include structures like 3-3-3|3-2-2, which is not maximally even by London's (2012) well-formedness constraints. On the other hand, London (2012, p. 164) convincingly argues that the 3-3-3-3-2-2 pattern (without the half-measure) acts a surface rhythm rather than a mode of metric entrainment. The underlying meter of this surface rhythm either has an isochronous structure (e.g., 2-2|2-2|2-2|2-2) or the non-isochronous but maximally even structure 3-3-3-3-4 (London, 2012). The relation between non-isochronous metrical structure and non-isochronous surface rhythm is subtle, as surface rhythm is not only distracting, but also an important cue for metric entrainment. Mental maintenance of especially non-isochronous meters depends on the surface rhythm, as the surface rhythm articulates the beat cycle (London, 2012, pp. 150–151).[3]

5.2 Metrical space

This section presents the construction of a set of meters, the *testset*, that is used to visualize a significant part of the metrical space of the current formalization. Discussion of this testset includes how the number of possible structures behaves by increasing period, and the role of **C8** (the principle of maximal evenness) in this. In the next section, this testset will be used to analyze the effect of the additional constraints from Section 4.3.

[2] In personal communication, London (October 30, 2016) suggests that subdivision ratios of 1:2 are allowed for levels lower than the tactus. But again, this refers to approaches in which there is no common fast pulse, and therefore does not apply to the current formalization.

[3] However, in other musical cultures, enculturated association may favor a non-isochronous metrical interpretation in other situations as well. London (2012, p. 150) hints to this, but for a more elaborate discussion of this, see van der Weij, Pearce, and Honing (2016)

The testset is formulated such that it contains meters that have period 2 to 18 and are well-formed with regard to **C1** to **C8**, with the default perceptual limit parameter settings ($u_{\min} = 100$ ms, $a_{\max} = 5000$ ms, $i_{\min} = 400$ ms, $i_{\max} = 1200$ms). Like in the previous section, the temporal unit u remain undefined, and instead the temporal frame (as constrained by **C4**, **C5** and **C6**) is calculated, such that it becomes clear which structures are possible at all. A meter will be regarded as well-formed when there is at least one value of u such that the meter is well-formed. In this definition, limit cases are included.

For feasibility reasons, the testset does not include meters that have a period of 19 or greater. This is because number and complexity of possible structures increases exponentially per period until the maximum $p = 50$, while the temporal frame in which these structures can be perceived as meters decreases (and therefore the relative relevance of these structures decreases as well). For similar reasons, the testset only includes meters in which the tactus level is level ℓ_1 ($c = 1$), plus all structures that have only two levels, such that the tactus level is necessarily ℓ_0 ($c = 0$). These are the meters with structures 2, 3, etc., up to and including 12 (for a meter with structure 13, **C5** requires that $u \leq 384.6$, while **C6** requires that $u \geq 400$). The testset includes these meters for the reason that the minimal period would otherwise be 4, which gives a substantially incomplete view of metrical space. Other meters with $c = 0$ are not included, but adding this class would not increase metrical space dramatically, as it is smaller than the class of meters with $c = 1$ and the maximal period of meters with $c = 0$ is 12 through **C6**. Excluding meters in which $c > 1$ does not imply a dramatic change either. There is a significant amount of meters with $c = 2$, but meters with $c = 3$ can only exist within small temporal frames ($100 \leq u \leq 150$) because of **C6**, while by the same constraint, $c = 4$ is not possible at all. In future research, the constraints of the current formalization may be defined algorithmically, such that metrical space can be computed algorithmically and the full metrical space can be taken into account, that is, *all* possible meters under the current formalization. Appendix A provides the set of rules (and some examples) under which the testset was systematically constructed, as well as construction of the testset without **C8**.

A full overview of the number of structures per period in the testset is provided in Table 5.1. This table also shows the effect of the principle of maximal evenness (**C8**), by providing the number of structures that can be generated without this constraint (for periods up to 16, because of feasibility). Figure 5.1 shows a graphical plot of Table 5.1. This figure visualizes how the number of structures increases exponentially with increasing period. Especially in the case that **C8** is not taken into account, the exponential relationship between period and possible structures is evident. Trend line $y = 0.3556e^{0.5482x}$ has a determination coefficient of $R^2 = 0.995$ and extrapolation to maximum period $p = 50$ gives $2.9 * 10^{11}$ structures (notwithstanding effects that would arise for large periods).[4] Indeed, **C8** is an important constraint with regard to the size of the metrical space. For the period range 2-16, only 24% of the structures that is allowed under **C1-7** is allowed by **C8**; for $p = 16$ only, this percentage is 18% and for extrapolation to $p = 50$ (with trend line $y = 0.295e^{0.4281x}$, $R^2 = 0.985$), this is 0.20%.

Addition of **C8** to **C1-7** also makes the number of structures dependent on whether the period is an even or an odd number; there are less structures with even periods than structures with odd periods. This is probably due to an effect for meters that are not maximally even on the tactus level, such that the super-tactus level must be maximally even. Often, this level consists of two beats. For even meters, a level of two beats is only maximally even when both intervals are equal (e.g., for $p = 16$: two beat intervals of 8 units), while for odd meters, there are two ways to get a maximally even level of two beats (e.g., for $p = 15$: there is one beat interval of 7 units and one beat interval of 8 units). Therefore, under **C8** there are more options for odd meters.

In general, it is remarkable that the metrical space associated with the formalization is much larger than the numbers presented by London (2012, Table 9.6, p. 169) in an overview.

[4] Trend lines were calculated using Microsoft Excel 2010.

p	No. of structures	No. without **C8**
2	1	1
3	1	1
4	2	2
5	3	3
6	3	5
7	6	8
8	7	14
9	12	23
10	12	42
11	37	80
12	39	144
13	100	266
14	99	490
15	242	895
16	287	1563
17	660	-
18	550	-
Total 1-16	851	3537
Total 1-18	2061	n/a

Table 5.1: Number of possible structures in the testset per period. The third column represents the number of possible structures without the principle of maximal evenness, **C8**, for periods up to 16.

This is partially due to the fact that London's (2012) overview does not intend to give a complete overview of metrical space (it only counts non-isochronous rotations) and the fact that the numbers in London's (2012) overview are incorrectly counted in some cases (London, personal communication, October 30, 2016). However, the difference mainly originates from the inconsistencies between London's (2012) theory and well-formedness constraints as discussed in Section 5.1. Indeed, London (personal communication, October 30, 2016) confirmed that *Hearing in Time* does not list all structures that are possible under the well-formedness constraints, as to avoid structures that have three beat classes, involve ambiguity or contradiction, or are likely to decompose (such as 3-5). This is another reason to analyze the effect of additional constraints on metrical space, as the constraints may restrict metrical space to such extent that it more resembles Table 9.6 in London (2012), as well as the ideas of London as formulated through personal communication (October 30, 2016). Because of this reason, we will regard restriction of the testset generally as a virtue.

5.3 Effect of additional constraints

This section provides an overview of the effect of additional constraints by discussing individual examples (along the lines of Section 5.1) as well as the effect of the constraints on the testset.

5.3.1 Exclusion of ambiguity and contradiction

The effect of addition of **C9** (exclusion of ambiguity and contradiction) is shown in the respective column of Table 5.2 and Figure 5.2 (see also Appendix A). The effect is notable: **C9** permits only 768 of the 2061 testset meters (37 %). Furthermore, the percentage of structures permitted by **C9**, relative to **C1-8** only, decreases by increasing period (see Figure 5.3). However, there

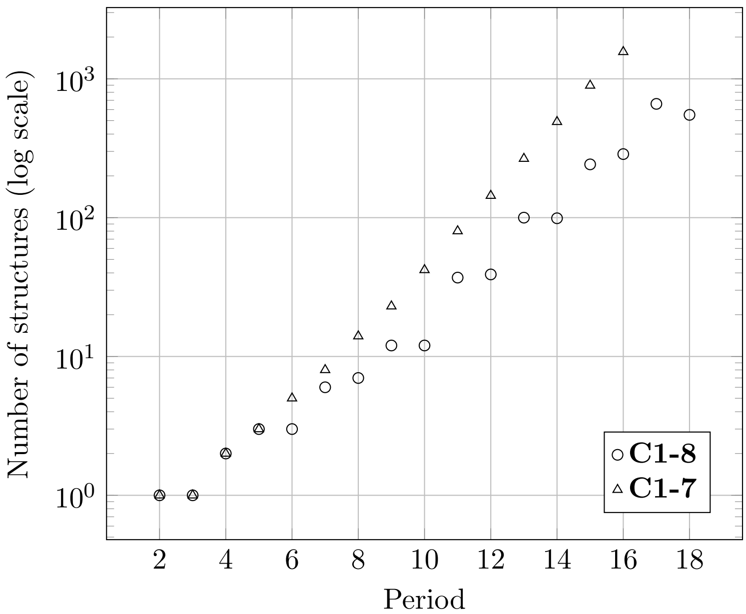


Figure 5.1: Number of possible structures in the testset, plotted on log scale to the period.

is only an effect for $p \geq 11$.

Constraint **C9** excludes a notable part of the structures discussed in Section 5.1.2 (structures that have multiple different beat classes, or beat classes with large ratios). **C9** excludes all beat class ratios of the form $J : J + N$ with $N \geq J$, as well as many structures in which there are three beat classes. However, the constraint does permit structures with three beat classes as long as the longest beat is not twice as long as the shortest beat. This is for instance the case for the structure 4-4|3-5 with $c = 1$ and $133.3 \leq u \leq 240$, but also the more exotic structure 7-7|5-8 with $c = 1$ and $100 \leq u \leq 150$ is permitted by **C9**. The maximum number of beat classes becomes five instead of seven (e.g., the structure 5-6-7|8-9 with $c = 1$ and $100 \leq u \leq 133.3$ is still possible). As an additional advantage, **C9** restricts possibilities in the organization of intermediate levels (Section 5.1.3), as it prohibits any metrical ambiguities there as well (e.g., the structure 2-2-2-2-2-2|2-2). The large drop of possible structures by **C9** is due to these two effects.

On the other hand, **C9** does not prohibit any structure which was allowed by both the initial formalization and London (2012) or London (personal communication, October 30, 2016). In other words, the constraint does not create disagreements where there were none initially; it does not make the formalization profusely stricter. Therefore, **C9** is a suitable supplement to the formalization.

5.3.2 Half/third-measure rule

Constraint **C10** (the half/third-measure rule) was specifically designed to eliminate structures with an uneven organization on intermediate levels (see Section 5.1.3). Indeed, the constraint prohibits these structures and thereby decreases metrical space; see the respective column of Table 5.2 and Figure 5.2 (see also Appendix A). This effect is much lower than for **C9**; **C10** permits 1635 of 2061 structures in the testset (79%). Also, this percentage does not decrease by increasing period and instead stabilizes around 80% (see Figure 5.3). As is the case for **C9**, **C10** only affects the testset for $p \geq 11$.

As expected, excluded structures fall in the category of uneven intermediate levels. Here, **C10** is slightly stricter than **C9** (e.g., while **C9** does not exclude structure 2-2-2|2-2, **C10** does). In general, **C10** eliminates structures of which the tactus level is maximally even, but the higher organization is not. In this, it is stricter than **C9**, as **C10** also excludes structures that are moderately uneven, while **C9** only excludes structures that explicitly give rise to metrical ambiguity or contradiction. Whether or not this is desirable, depends on where one prefers to put the boundaries of well-formedness.

In another sense, **C10** is less strict than **C9**, as it permits structures that are very uneven on a level *below* the maximally even level (which is by **C8** either the tactus level or the super-tactus level). Therefore, **C10** alone does not sufficiently rule out structures that are metrically ambiguous; as long as the half- or third-measure is maximally even, structures with large beat class ratios are not excluded by this constraint. However, **C10** has another advantage: for every structure that it eliminates from the testset, there is another structure within the testset that has the same organization (and same tactus level) as the eliminated structure, except that it lacks the intermediate level that was the ground for elimination by **C10**. For example, structure 2-2-2|2-2 with $c = 1$ is not allowed by **C10**, but 2-2-2-2-2 with $c = 1$ remains in the testset (for $200 \leq u \leq 500$). This is because **C10** only excludes structures in which there is a level that is not needed to maintain maximal evenness of the full structure, as otherwise this level would have been maximally even itself by **C8**.

5.3.3 Beat class maximum

By eliminating structures with three or more beat classes, **C11** restricts metrical space as well, see the respective column of Table 5.2 and Figure 5.2 (see also Appendix A). The effect is somewhat less strong than for **C9**, but stronger than for **C10**: **C11** permits 907 out of 2061 structures of the testset (44 %). Furthermore, the percentage of structures permitted by **C11**, relative to **C1-8** only, decreases by increasing period (see Figure 5.3). The elimination effect is strongest at structures with odd periods, lowering their prevalence over the structures with even periods. As was the case for the previous two constraints, there is only an effect for $p \geq 11$.

Just like **C10**, **C11** has some overlap with **C9**, as it mainly works on beat classes (see Section 5.1.2). Many structures that are eliminated by **C9** because of their beat class ratios are also eliminated by **C11** on grounds of number of beat classes and vice versa; when there are multiple beat classes, there is a higher chance that some of their mutual ratios are problematic. Of course, there are also differences. **C11** is stricter than **C9** with regard to *number* of beat classes: we have seen that there can still be five beat classes when only **C9** is applied (such as 5-6-7|8-9 with $c = 1$ and $100 \leq u \leq 133.3$), while **C11** eliminates these structures. In turn, **C11** is less strict with regard to beat class *ratios* than **C9**, as there can still be large ratios between beat classes when only **C11** is applied (such as 2-2-2-2|2-6 with $c = 1$ and $u = 200$), while these structures are excluded by **C9**.

5.3.4 Meta-rule of temporal frame

The effect of addition of **C12** (the meta-rule of temporal frame) is shown in the respective column of Table 5.2 and Figure 5.2 (see also Appendix A). The effect is the largest so far: **C12** permits only 538 of the 2061 testset meters (26 %). Furthermore, the percentage of structures permitted by **C12**, relative to **C1-8** only, decreases by increasing period, more than for any other additional constraint (see Figure 5.3). The effect starts at $p = 10$, where the constraint eliminates structure 10 with $c = 0$ as this structure is only allowed for $400 \leq u \leq 500$ (through **C5** and **C6**), such that for all u' and u'', $u' \not> u'' + 100$ ms. For $p \geq 11$, the percentage of allowed structures quickly drops further. This is because **C12** reinforces the boundaries placed by temporal constraints **C4**, **C5** and **C6**. For instance, all subdivision patterns of 6 or longer are eliminated by **C12** (except when $c = 0$), because of the following: By **C4**, no beat is shorter than 100 ms. This means that the absolute distance between beats that are adjacent on the tactus level must be more than 200 ms in order to abide by **C12**. But since **C6** requires that no beat on the tactus level is longer than 1200 ms, the unit distance between beats that are adjacent on the tactus level may not be larger than 5 (since when this is 6, the absolute distance drops beneath 200 ms). Through this, **C12** is the only additional constraint that attacks the problem of disagreement with regard to length of subdivision patterns (see Section 5.1.1).

For $p \geq 17$, the effect of adding **C12** becomes so large that it stops the exponential growth of possible meters with increasing period. This is also due to reinforcement of the temporal constraints. For meters in which $c = 1$ and a tactus level that has 2s, the minimum value of u is 200 ms by **C6** (since if $u \leq 200$, the absolute beat interval set has an element that is smaller than 400). However, when $p \geq 17$, the maximum value of u is 294.1 by **C5**, which makes the temporal frame of u smaller than 100 ms. Therefore, all structures with $p \geq 17$ which have any 2s are eliminated under **C12** (e.g., 3-3-3|3-3-2). For $p \geq 22$, a similar problem arises for 3s, such that structures may only have 4s and 5s under **C12** (as we have already seen that 6s are not allowed under **C12**). Furthermore, no meter with $p \geq 25$ is allowed at all by **C12**, as then by **C5**, $u \leq 200$, while $u \geq 100$ by **C4**. Because of this ceiling, the testset can be extended to $p = 50$ for the case of **C12** with minimal effort. A full overview of the number of structures that are allowed by **C12** per period in this extended testset is provided in Figure 5.4 (see also Appendix A). Note that we still only regard meters with $c = 1$ (plus the exceptions for $c = 0$). The total number of extended testset structures that are permitted by **C12** is 757.

The temporal unit frame range of 100 ms in the definition of **C12** is rather strict. The range was chosen arbitrarily, but not at random: it is exactly the range that eliminates structures with subdivision patterns larger than 5 (as favored by London, 2012; and London, personal communication, October 30, 2016), ratios of the form $J : J + 1$ in which $J \geq 5$ (as favored by London, 2012) and many more ratios on which there is disagreement between London's (2012) theory and well-formedness constraints. This way, **C12** addresses these disagreements (Section 5.1.1–3) all at once and is stricter than any of the other three additional constraints. In a sense, it is too strict, as the 100 ms range eliminates any meter with $p \geq 25$. Future research may provide a more suitable range.

On the other hand, in some cases **C12** is less strict than the other additional constraints. For instance, **C12** permits some structures that have three beat classes, such as the structure 4-4|3-5 with $c = 1$ and $133.3 \leq u \leq 240$. This disagrees with London (personal communication, October 30, 2016), but on the other hand, it seems intuitive to not *a priori* disregard structures that have three beat classes. Rather, these structures should be disregarded only when they are unlikely (i.e., only possible within a small temporal frame). **C12** is also less strict than **C9** in the case of structures that have an uneven organization of intermediate levels. This is effect is most notable for $12 \leq p \leq 16$, where the period is large enough for these structures to emerge, but not large enough for **C12** to eliminate them by excluding 2s for $p \geq 17$. Because of this, we also regard the effect of **C9** and **C12** combined. As can be seen in the corresponding column of Table 5.2 and Figure 5.2, only 473 of the 2061 testset structures are permitted by **C9+C12** (23 %); see also Figure 5.3. In the extended testset (i.e. the testset including $p > 18$), the total number of permitted structures becomes 676. A full overview of the structures permitted by **C12** (per period, with and without **C9**) is provided in Figure 5.4.

Constraint **C12** incorporates a form of perceptual stability into the formalization. The percept of a meter that is only perceivable within a very specific time frame, is probably not very stable; that is, it is likely to transition into a similar percept that is easier to maintain. In a possible future implementation of the formalization in a cognitive model, stability can be assigned in more natural ways than a hard constraint like **C12**. Chapter 6 will discuss some speculations of how stability could be implemented in such a model.

5.3.5 Principle of maximal evenness with rhythmic oddity

As analyzing the effect of **C13** on the testset requires tedious construction of another testset, we will only look at its effects on individual meters. This will give an impression of the advantages and problems of **C13**. However, we can predict two general effects. First, the number of

p	**C1-8**	**C9**	**C10**	**C11**	**C12**	**C9+C12**
2	1	1	1	1	1	1
3	1	1	1	1	1	1
4	2	2	2	2	2	2
5	3	3	3	3	3	3
6	3	3	3	3	3	3
7	6	6	6	6	6	6
8	7	7	7	7	7	7
9	12	12	12	12	12	12
10	12	12	12	12	11	11
11	37	29	31	29	26	26
12	39	23	27	35	25	21
13	100	55	88	64	58	53
14	99	56	76	67	62	55
15	242	92	193	97	108	90
16	287	121	212	153	146	117
17	660	193	531	203	37	37
18	550	152	430	212	30	28
Total	2061	768	1635	907	538	473

Table 5.2: Number of possible structures in the testset per period. Every column represents one constraint that is added to the formalization with only constraints **C1-8** (the last column represents addition of both **C9** and **C12** to **C1-8**).

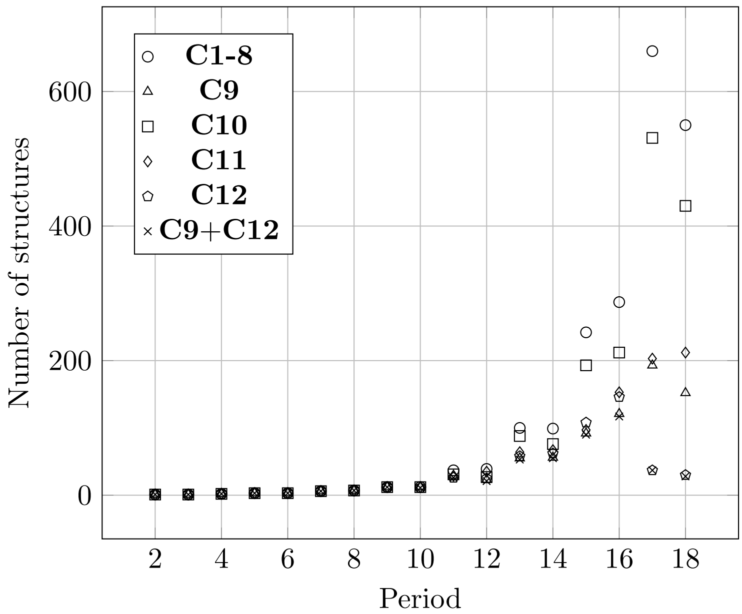


Figure 5.2: Number of possible structures in the testset per period under the constraints **C1-8**, **C1-9**, **C1-8+C10**, **C1-8+C11**, **C1-8+C12** and **C1-8+C9+C12**.

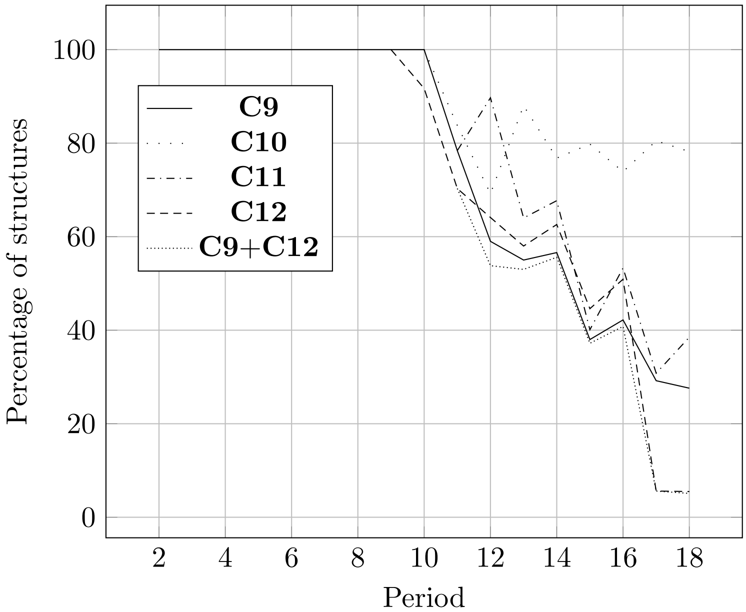


Figure 5.3: Percentage of testset structures permitted by constraints **C1-9**, **C1-8+C10**, **C1-8+C11**, **C1-8+C12** and **C1-8+C9+C12**, with respect to constraints **C1-8**, per period.

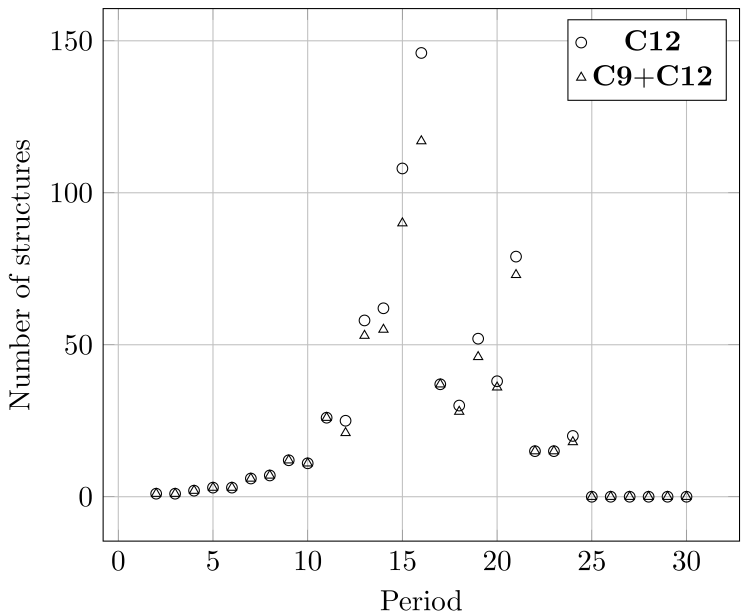


Figure 5.4: Number of possible structures in the extended testset per period under constraints **C1-8+C12**, and **C1-9+C12**.

permissible meters under **C1-7+C13** will be more than under **C1-8** (as **C13** adds a disjunction to **C8**), but less than under **C1-7** (as **C13** does define boundaries that are not there for **C1-7**). Second, **C13** reverses the relative prevalence of structures with odd-numbered period which was discussed in Section 5.2. This is because for even periods, there are now three ways to get a permissible level of two beats (e.g., for $p = 16$: 8|8, 7|9 and 9|7, which was only 8|8 with **C8**), while for odd periods, there are still only two ways to get a permissible level of two beats (e.g., for $p = 15$, 7|8 and 8|7).

As discussed in Section 4.3.2, **C13** was designed to relax the principle of maximal evenness (strictness of this principle was identified as a problem in Section 5.1.4). **C13** is a weakened version of **C8** that relaxes the principle of maximal evenness in order to include meters that do not have a maximally even tactus level or super-tactus level, but do abide by the principle of rhythmic oddity.

By its addition of a disjunction with respect to **C8**, **C13** includes structures to the formalization that were not permitted by **C8**. As we have seen in Section 4.3.2, one of these structures is 3-3-3|3-2-2. This structure was not allowed by London's (2012) theory nor the well-formedness constraints, but London (personal communication, October 30, 2016) proposes that this structure may be permissible.

The converse holds for the structures 3-5, 5-7 and 4-3|4-5, as discussed in Section 5.1.2 and 5.1.4; that is, London (2012) defends these structures, while London (personal communication, October 30, 2016) retracts them (remaining ambivalent on 3-5), although he argues that this is because of general reasons (length of subdivision and existence of three beat classes) and not because of considerations on maximal evenness. Like for 3-3-3|3-2-2 in the previous paragraph, London (personal communication, October 30, 2016) in general argues that meters may not have to be as maximally even as proposed in London (2012). Note that a rhythmically odd level always has exactly two beats, so structure 4-6-6 is still impermissible under **C13**.

As the previous two paragraphs show, **C13** has consequences for cases of doubt, but it also has more severe implications. For instance, it allows for the structures 4-6 (with $c = 1$ and $100 \leq u \leq 200$) and 2-4 (with $c = 1$ and $200 \leq u \leq 300$). To prevent this, **C13** could be formulated differently such that the tactus level is not allowed to be the level of rhythmic oddity, but this eliminates the structure 3-5. Furthermore, **C13** in general aggravates the problem of **C8** that in some cases, all proportions and number of beat classes are allowed on the tactus level (as discussed in Section 5.1.2): for **C13**, not only a maximally even super-tactus level, but also a rhythmically odd super-tactus level makes any distribution on the tactus level possible. This latter problem can only be addressed through intervention by one of the other additional rules, such as **C9** or **C12**. **C9** would eliminate 2-4 but not 4-6, while **C12** would eliminate both structures. However, it is doubtful whether the principle of rhythmic oddity is a general solution for the problem that maximal evenness is too strict (as posed in Section 5.1.4), and whether it has a cognitive ground.

5.4 Summary and recommendations

In this chapter, we examined the effect of the additional constraints from Section 4.3 on the formalization. Of the additional constraints, **C9** (exclusion of ambiguity and contradiction) and **C12** (meta-rule of temporal frame) exclude the most structures.

C9 fits well in the theory of London (2012), as it is a generalization of London's (2012) well-formedness constraint 4.1.2, and we have seen in Section 4.3.1 that London (2012; personal communication, October 30, 2016) argues against ambiguity and contradiction. Furthermore, **C9** does not eliminate structures that are explicitly allowed by London (2012) or London (personal communication, October 30, 2016). Therefore, it is a suitable supplement to the formalization. As **C9** also shows that a constraint for maximal evenness alone does not sufficiently restrict the possibilities for a well-formed meter, future research may provide a general alternative to the combination of **C8** (principle of maximal evenness) and **C9** that is more restricting. Ideally, such a rule would be grounded in perceptual studies and also address the problem that the principle of maximal evenness is too strict, which is not satisfactorily addressed by **C13**.

C12 also fits well within the theory of London (2012), as it addresses inconsistencies with regard to length of subdivision, number and proportions of beat classes and existence and organization of intermediate levels. Furthermore, it incorporates a form of perceptual stability into the formalization and combines well with **C9** (the constraints complement each other). As it is currently too strict, another temporal frame may be chosen, or the idea of this constraint could be incorporated in another way through implementation in a cognitive model.

In contrast to **C9** and **C12**, **C10** (the half/third-measure rule) only eliminates a limited part of metrical space. **C10** fits less natural within the theory of London (2012), and although it does eliminate structures with uneven intermediate levels, the constraint does not sufficiently rule out structures that are metrically ambiguous. On the other hand, **C10** eliminates some moderately uneven structures that are not excluded by **C9**. Furthermore, for all structures that it eliminates, there is another structure within the testset that has a similar organization. In all, **C10** is less recommendable than **C9** or **C12**. The same holds for **C11** (the beat class maximum), which fits the governing rule of London (personal communication, October 30, 2016), but eliminates structures with three beat classes *a priori* instead of eliminating them through considerations that are more general.

In summary, it is recommendable to add **C9** and **C12** to the formalization. However, it might be better if further research does not unthinkingly add these constraints to the formalization, but rather focuses on development of a general, perceptual rule of maximal evenness, and/or development of a cognitive model, such that temporal effects can be implemented in more natural ways.

Chapter 6

Conclusion and future directions

The aim of this thesis was to obtain a formalization of non-isochronous metrical structure. As argued in Chapter 1, a cognitive theory of metrical structure should be: (1) as simple as possible, while capturing all relevant details; (2) formal, that is, completely and unambiguously specified; (3) grounded in empirical perceptual studies and cross-cultural (and therefore incorporating non-isochronous metrical structure).

Chapter 2 has shown that most existing formal theories do not meet requirement (3) by disregarding non-isochronous metrical structure. The theory of London (2012) does meet requirements (1) and (3), but is not formal and thereby does not meet requirement (2). In this thesis, we have therefore formalized the theory of London. In this formalization (Chapter 4), every *meter* is presented by a tuple that contains its *structure* (a set of *levels*, which are in turn sets of *beats*), a *tactus level index* and a *temporal unit*. The temporal unit functions to define a *temporal frame* within which a structure can be perceived as a meter. The formalization contains eight constraints, of which three basic structural constraints, three temporal constraints and two advanced structural constraints. Together, the constraints agree with the well-formedness constraints in London. Apart from these constraints, additional constraints were formulated to account for inconsistencies within the theory of London.

The current formalization took a perspective that differs from related work (Chapter 3) by aiming to specify the theory of London (2012) and analyzing its ambiguities and inconsistencies. This analysis (Chapter 5) identified several inconsistencies between London's theory and well-formedness constraints. The most important inconsistencies were the following: London's theory is stricter than the well-formedness constraints with regard to perceivable length of subdivision patterns, permissible number and proportions of beat classes and possible organizations of intermediate levels than the well-formedness constraints. London's theory is less strict than the well-formedness constraints with regard to the principle of maximal evenness. Chapter 5 also analyzed the number of possible meters, or *metrical space*, under the current formalization, as well as effects of the additional constraints on the metrical space. The chapter also regarded how the additional constraints did or did not rule out specific inconsistencies. Future research may provide a general, perceptual rule as an alternative to the principle of maximal evenness, as this principle does not yet give a fully satisfactory account of regularity within meter – it is both too strict and not strict enough.

Apart from providing a general regularity rule, future research may extend the formalization to include meters with a non-isochronous lowest level. For instance, this could be done by not taking a single temporal unit u, but a set of constants representing every beat interval on the lowest level. In such a definition, there is no common fast pulse. Indeed, in such a definition, there may not be any isochronous level at all, so no common slow pulse either (see Section 3.3.1). In order to get a common slow pulse model, another constraint should be added that requires that one of the levels in the meter is (almost) isochronous. In such a model, temporal unit u could be also be replaced by a temporal constant on the tactus level, full period or 'the level that is isochronous' as above, potentially differing per meter. In all, such extensions provide an alternative interpretation for the open end discussed in Section 3.3.1.

A general idea for future research is implementation of the current formalization in a cognitive model of meter perception. Such a model could for instance be the probabilistic model of Temperley (2007) or the model of van der Weij et al. (2016) that focuses more on simulating enculturation. In such a model, data sets of non-isochronous meters in different cultures could assign typical onset patterns for rhythms associated with every possible meter, as well as a prior probability to every meter. By training the model to different data sets, cultural differences with regard to meter perception may be modeled as well. After training, the model may be tested empirically to investigate its cognitive validity. Of course, it is important to decide which (if any) additional constraints from Chapter 5 are used for implementation. Besides its usefulness for implementation, algorithmic description of the constraints may also provide a more complete image of the metrical space under the formalization.

A future probabilistic model may take perceptual factors into account. The most pressing factor to investigate is the influence of tempo. Rather than by a single *value*, prior probability of every meter could be represented by a probability *distribution* (a bell-shaped curve) as a function of tempo. Now, the requirement could be added that probability functions of meters with large temporal frames have a bigger area-under-the-curve than probability functions of meters with smaller temporal frames. Alternatively, a prior probability function could be assigned to tactus intervals themselves (with the optimum of the probability function on 600 ms, the preference region for the tactus; see London, 2012, pp. 30–31) which yields an optimal value for the temporal unit u per meter (depending on the unit distances in the tactus level). In turn, the probability of a meter for a given value of u could be derived from this. Through such probability functions, extreme meters may be perceivable in some cases, but probability of another, less extreme meter may be higher for that particular case and value of u, such that these unstable percepts are likely to transition into similar, more stable meters. This would be a more natural way of modeling stability than is provided by the additional constraint **C12** of the current formalization. Furthermore, this way tempo can be fully separated again from the structural properties of different meters. In turn, this may also bring transition zones between tempo-metrical types (as discussed in Section 2.4.3) back into focus, by mapping the way in which probability curves of similar meters overlap at changing tempos (representing overlapping percepts for the same structure or similar structures).

An advanced version of the model may implement the severity of metric changes as well. In musical practice, the meter of a piece is often subject to slight changes, such as the addition or deletion of certain metrical levels (London, 2012; Gotham, 2015). These changes do not drastically change how the meter of the piece is perceived, because its periodicity and global structure are preserved, even though the meter *after* the change is theoretically distinct from the meter *before* the change (e.g., a 2-2 structure with $c = 1$ and $u = 300$ ms versus a 2-2|2-2 structure with $c = 2$ and $u = 150$ ms). To distinguish between these subtle changes and more drastic changes, operations (along the lines of the relationships of Gotham, 2015) could be defined on meters, such that metrical changes are form-preserving if and only if the new meter is a result of such an operation applied to the old meter. These operations could be formally defined (for instance as relations), or again data sets may be used to incorporate this factor probabilistically, such as defining the probability of different subtle changes in a given meter. This way, the model may implement London's (2012) ideas on metric entrainment as a dynamic process, including ideas of stability and recurrence of meter.

Apart from redefining the above properties of meter, there are many more possibilities for implementation of the formalization in a probabilistic model. For instance, a similar Gaussian approach could be applied to the location of attentional peaks under temporal spread as discussed in Section 2.4.3. Furthermore, a probabilistic model could incorporate individual factors, such as individually differing metric floors or optimal tactus intervals. Moreover, advanced versions of the proposed cognitive model may reinvoke the possibility of a non-isochronous N-cycle, as was dropped in the current formalization. An alternative to a probabilistic model would be a cognitive model that describes the process of meter construction through interpolation (as described, for instance, in London, 2012, pp. 146–151).

In summary, this thesis may arouse the development of a formal cognitive model of meter perception that is based on the theory of London (2012). Such a cognitive model may provide insight into the cognitive processes behind meter perception within a cross-cultural paradigm.

References

Arom, S. (1991). *African polyphony and polyrhythm* (M. Thom, B. Tuckett, & R. Boyd, Trans.). Cambridge, United Kingdom: Cambridge University Press.

Bregman, M. R., Iversen, J. R., Lichman, D., Reinhart, M., & Patel, A. D. (2013). A method for testing synchronization to a musical beat in domestic horses (Equus ferus caballus). *Empirical Musicology Review*, *7*(3-4), 144–156.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, *18*(12), 2844–2854.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, *2*(3), 113–124.

Chomsky, N. (1957). *Syntactic structures.* The Hague, Netherlands: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 473–500). San Diego, CA: Academic Press.

Clayton, M. (2008). *Time in Indian music: rhythm, metre, and form in North Indian rāg performance.* Oxford, United Kingdom: Oxford University Press.

Clough, J., & Douthett, J. (1991). Maximally even sets. *Journal of Music Theory*, *35*(1/2), 93–173.

Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music.* Chicago, IL: University of Chicago Press.

Dahl, S., Huron, D., Brod, G., & Altenmüller, E. (2014). Preferred dance tempo: does sex or body morphology influence how we groove? *Journal of New Music Research*, *43*(2), 214–223.

Desain, P. (1992). A (de)composable theory of rhythm perception. *Music Perception: An Interdisciplinary Journal*, *9*(4), 439–454.

Desain, P., & Honing, H. (1999). Computational models of beat induction: The rule-based approach. *Journal of new music research*, *28*(1), 29–42.

Essens, P. J. (1995). Structuring temporal sequences: Comparison of models and factors of complexity. *Perception & Psychophysics*, *57*(4), 519–532.

Essens, P. J., & Povel, D.-J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception & Psychophysics*, *37*(1), 1–7.

Fitch, W. T., & Rosenfeld, A. J. (2007). Perception and production of syncopated rhythms. *Music Perception: An Interdisciplinary Journal*, *25*(1), 43–58.

Forth, J. C. (2012). *Cognitively-motivated geometric methods of pattern discovery and models of similarity in music* (Unpublished doctoral dissertation). Goldsmiths, University of London.

Fracile, N. (2003). The Aksak rhythm, a distinctive feature of the Balkan folklore. *Studia Musicologica Academiae Scientiarum Hungaricae*, *44*(1-2), 191–204.

Gärdenfors, P. (2000). *Conceptual spaces: the geometry of thought.* Cambridge, MA: MIT Press.

Gotham, M. (2015). Meter metrics: Characterizing relationships among (mixed) metrical structures. *Music Theory Online*, *21*(2), 1–21.

Hannon, E. E., Soley, G., & Ullal, S. (2012). Familiarity overrides complexity in rhythm perception: A cross-cultural comparison of American and Turkish listeners. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 543–548.

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science*, *16*(1), 48–55.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(35), 12639–12643.

Haugen, M. R. (2015). Asymmetrical meter in Scandinavian folk music and dance: A case study of Norwegian Telespringar. In *Proceedings of the ninth triennial conference of the european society for the cognitive sciences of music* (pp. 1–5). European Society for the Cognitive Sciences of Music.

Honing, H. (2012). Without it no music: Beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences*, *1252*(1), 85–91.

Honing, H. (2013). Structure and interpretation of rhythm in music. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 369–404). San Diego, CA: Elsevier.

Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation.* Cambridge, MA: MIT Press.

Imbrie, A. W. (1973). "Extra" measures and metrical ambiguity in Beethoven. In A. Tyson (Ed.), *Beethoven studies* (1st ed., pp. 45–66). New York, NY: Norton.

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences*, *1169*(1), 58–73.

Iyer, V. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception: An Interdisciplinary Journal*, *19*(3), 387–414.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*(3), 459–491.

Kalender, B., Trehub, S. E., & Schellenberg, E. G. (2013). Cross-cultural differences in meter perception. *Psychological Research*, *77*(2), 196–203.

Kauffman, R. (1980). African rhythm: A reassessment. *Ethnomusicology*, *24*(3), 393–415.

Keith, M. (1991). *From polychords to Pólya: Adventures in musical combinatorics.* Philadelphia, PA: Vinculum Press.

Komar, A. J. (1971). *Theory of suspensions.* Princeton, NJ: Princeton University Press.

Kvifte, T. (2004). Description of grooves and syntax/process dialectics. *Studia Musicologica Norvegica*, *30*, 54–77.

Kvifte, T. (2007). Categories and timing: On the perception of meter. *Ethnomusicology*, *51*(1), 64–84.

Ladinig, O., Honing, H., Háden, G., & Winkler, I. (2009). Probing attentive and preattentive emergent meter in adult listeners without extensive music training. *Music Perception: An Interdisciplinary Journal*, *26*(4), 377–386.

Ladinig, O., Honing, H., Háden, G., & Winkler, I. (2011). Erratum to probing attentive and preattentive emergent meter in adult listeners without extensive music training (vol 26, pg 377, 2009). *Music Perception: An Interdisciplinary Journal*, *28*(4), 444.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159.

Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, *6*(2-3), 177–208.

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*(1), 1–37.

Lee, K. M., Barrett, K. C., Kim, Y., Lim, Y., & Lee, K. (2015). Dance and music in "Gangnam Style": How dance observation affects meter perception. *PloS ONE*, *10*(8), e0134725.

Leman, M. (2007). *Embodied music cognition and mediation technology.* Cambridge, MA: MIT Press.

Lempel, A., & Ziv, J. (1976). On the complexity of finite sequences. *IEEE Transactions on Information Theory*, *22*(1), 75–81.

Lerdahl, F. (2015). Concepts and representations of musical hierarchies. *Music Perception: An Interdisciplinary Journal*, *33*(1), 83–95.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Lindblom, B., & Sundberg, J. (1969). Towards a generative theory of melody. *STL-QPSR*, *10*(4), 53–86.

Locke, D. (1982). Principles of offbeat timing and cross-rhythm in southern Eve dance drumming. *Ethnomusicology*, *26*(2), 217–246.

Locke, D. (1998). *Drum Gahu: An introduction to African rhythm.* Tempe, AZ: White Cliffs Media.

London, J. (2012). *Hearing in time: Psychological aspects of musical meter* (2nd ed.). Oxford, United Kingdom: Oxford University Press.

Longuet-Higgins, H. C. (1973). Comments on the Lighthill Report and the Sutherland Reply. In *Artificial intelligence: a paper symposium* (pp. 35–37).

Longuet-Higgins, H. C. (1976). Perception of melodies. *Nature*, *263*, 646–653.

Longuet-Higgins, H. C. (1978). The perception of music. *Interdisciplinary Science Reviews*, *3*(2), 148–156.

Longuet-Higgins, H. C., & Lee, C. S. (1982). The perception of musical rhythms. *Perception*, *11*(2), 115–128.

Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception: An Interdisciplinary Journal*, *1*(4), 424–441.

Longuet-Higgins, H. C., & Steedman, M. J. (1971). On interpreting Bach. In B. Meltzer & D. Michie (Eds.), *Machine intelligence* (Vol. 6, pp. 221–241). Edinburgh, United Kingdom: Edinburgh University Press.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, *79*(6), 487–509.

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*(3), 348–367.

Mirka, D. (2009). *Metric manipulations in Haydn and Mozart: Chamber music for strings, 1787-1791.* Oxford, United Kingdom: Oxford University Press.

Mithen, S. J. (2007). *The singing Neanderthals: The origins of music, language, mind, and body.* Cambridge, MA: Harvard University Press.

Morris, R. (1998). Sets, scales, and rhythmic cycles: A classification of talas in Indian music. In *21st annual meeting of the society for music theory* (pp. 1–46).

Naveda, L., & Leman, M. (2011). Hypotheses on the choreographic roots of the musical meter: a case study on Afro-Brazilian dance and music. In A. P. Ghiena, P. Jacquier, M. Valles, & M. Martínez (Eds.), *Musicalidad humana: Debates actuales en evolución, desarrollo y cognición e implicancias socio-culturales* (pp. 477–495). Actas del X Encuentro de Ciencias Cognitivas de la Música.

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *The Journal of Neuroscience*, *32*(49), 17572–17581.

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

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

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(4), 728–741.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception: An Interdisciplinary Journal*, *11*(4), 409–464.

Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception: An Interdisciplinary Journal*, *24*(1), 99–104.

Patel, A. D. (2008). *Music, language, and the brain.* Oxford, United Kingdom: Oxford University Press.

Patel, A. D. (2014). The evolutionary biology of musical rhythm: Was Darwin wrong? *PLoS Biology*, *12*(3), e1001821.

Pearce, M., & Rohrmeier, M. (2012). Music cognition and the cognitive sciences. *Topics in cognitive science*, *4*(4), 468–484.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*(5727), 1430.

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, *105*(3), 533–546.

Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, *67*(1), 94–102.

Polak, R. (2010). Rhythmic feel as meter: Non-isochronous beat subdivision in jembe music from Mali. *Music Theory Online*, *16*(4).

Polak, R., & London, J. (2014). Timing and meter in Mande drumming from Mali. *Music Theory Online*, *20*(1).

Povel, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(1), 3–18.

Povel, D.-J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception: An Interdisciplinary Journal*, *2*(4), 411–440.

Povel, D.-J., & Okkerman, H. (1981). Accents in equitone sequences. *Perception & Psychophysics*, *30*(6), 565–572.

Proca-Ciortea, V. (1969). On rhythm in Rumanian folk dance. *Yearbook of the International Folk Music Council*, *1*, 176–199.

Rahn, J. (1991). Coordination of interval sizes in seven-tone collections. *Journal of Music Theory*, *35*(1/2), 33–60.

Repp, B. H. (2007a). Embodied rhythm: Commentary on "The contribution of anthropometric factors to individual differences in the perception of rhythm" by Neil P. McAngus Todd, Rosanna Cousins, and Christopher S. Lee. *Empirical Musicology Review*, *2*(1), 14–16.

Repp, B. H. (2007b). Perceiving the numerosity of rapidly occurring auditory events in metrical and nonmetrical contexts. *Perception & Psychophysics*, *69*(4), 529–543.

Repp, B. H., London, J., & Keller, P. E. (2005). Production and synchronization of uneven rhythms at fast tempi. *Music Perception: An Interdisciplinary Journal*, *23*(1), 61–78.

Rice, T. (1994). *May it fill your soul: Experiencing Bulgarian music.* Chicago, IL: University of Chicago Press.

Rice, T. (2000). Béla Bartók and bulgarian rhythm. In E. Antokoletz, V. Fischer, & B. Suchoff (Eds.), *Bartók perspectives: Man, composer, & ethnomusicologist* (pp. 196–212). Oxford, United Kingdom: Oxford University Press.

Schaefer, R. S., Vlek, R. J., & Desain, P. (2011). Decomposing rhythm processing: Electroencephalography of perceived and self-imposed rhythmic patterns. *Psychological research*, *75*(2), 95–106.

Sethares, W. A., & Staley, T. W. (2001). Meter and periodicity in musical performance. *Journal of New Music Research*, *30*(2), 149–158.

Shmulevich, I., & Povel, D.-J. (2000). Measures of temporal pattern complexity. *Journal of New Music Research*, *29*(1), 61–69.
Simon, H. A. (1968). Perception du pattern musical par "auditeur". *Science de l'Art*, *5*(2), 28–34.
Simon, H. A., & Sumner, R. K. (1968). Pattern in music. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 219–250). New York, NY: Wiley.
Singer, A. (1974). The metrical structure of Macedonian dance. *Ethnomusicology*, *18*(3), 379–404.
Slawson, A. W. (1968). [Review of the book *Computer applications in music*, by G. Lefkoff]. *Journal of Music Theory*, *12*(1), 105–11.
Smith, L. M., & Honing, H. (2006). Evaluating and extending computational models of rhythmic syncopation in music. In *Proceedings of the international computer music conference* (pp. 688–691). San Francisco, CA: International Computer Music Association.
Styns, F., van Noorden, L., Moelants, D., & Leman, M. (2007). Walking on music. *Human Movement Science*, *26*(5), 769–785.
Tanguiane, A. (1993). *Artificial perception and music recognition.* Berlin, Germany: Springer-Verlag.
Temperley, D. (2007). *Music and probability.* Cambridge, MA: MIT Press.
Temperley, D. (2013). Computational models of music cognition. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 327–368). San Diego, CA: Elsevier.
Thul, E., & Toussaint, G. T. (2008). Rhythm complexity measures: A comparison of mathematical models of human perception and performance. In *Proceedings of the ninth international symposium on music information retrieval* (pp. 663–668). International Society for Music Information Retrieval.
Todd, N. P. M. (1994). The auditory "primal sketch": A multiscale model of rhythmic grouping. *Journal of New Music Research*, *23*(1), 25–70.
Todd, N. P. M., Cousins, R., & Lee, C. S. (2007). The contribution of anthropometric factors to individual differences in the perception of rhythm. *Empirical Musicology Review*, *2*(1), 1–13.
Todd, N. P. M., & Lee, C. S. (2015). The sensory-motor theory of rhythm and beat induction 20 years on: a new synthesis and future perspectives. *Frontiers in Human Neuroscience*, *9*(444), 1–25.
Todd, N. P. M., O'Boyle, D. J., & Lee, C. S. (1999). A sensory-motor theory of rhythm, time perception and beat induction. *Journal of New Music Research*, *28*(1), 5–28.
Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: hierarchical eigenmodes in music-induced movement. *Music Perception: An Interdisciplinary Journal*, *28*(1), 59–70.
Touma, H. H. (1996). *The music of the Arabs.* Portland, OR: Amadeus Press.
Toussaint, G. T. (2002). A mathematical analysis of African, Brazilian, and Cuban clave rhythms. In *Proceedings of BRIDGES: Mathematical connections in art, music and science* (pp. 157–168). The Bridges organization.
Trainor, L. J., Gao, X., Lei, J.-J., Lehtovaara, K., & Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *Cortex*, *45*(1), 35–43.
van der Weij, B., Pearce, M. T., & Honing, H. (2016). *A probabilistic model of meter perception: Simulating enculturation.* Manuscript submitted for publication.
Vos, P. G. M. M. (1973). *Waarneming van metrische toonreeksen* (Unpublished doctoral dissertation). Radboud University Nijmegen. (Retrieved from Radboud Repository; http://hdl.handle.net/2066/148457)
Vuust, P., & Witek, M. A. (2014). Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, *5*(1111), 1–14.

Wang, L.-C., & Tsai, C.-G. (2009). Rock to the music: Theory exploration and electroencephalography (EEG) investigation. In *Proceedings of the 2nd international conference of students of systematic musicology* (pp. 59–64). Institute for Psychoacoustics and Electronic Music, Ghent University.

Yeston, M. (1976). *The stratification of musical rhythm.* New Haven, CT: Yale University Press.

Appendix A

Systematic construction of meters for analysis

This appendix provides the set of rules under which the testset of Section 5.2 was systematically constructed. Note again that the testset only includes meters with $c = 1$, plus all structures that have only two levels, such that necessarily $c = 0$.

A.1 Testset construction for C1-8

Systematic construction of the testset has been performed under the following rules. The testset was constructed per period. For every period, first the maximum possible number of beats on level ℓ_1 was determined, henceforth $b_{\max}$ (with b the number of beats on a level ℓ_1 with a given organization). Because of **C7** (the principle of non-adjacency), the minimum subdivision length is 2, which entails that $b_{\max} = \lfloor \frac{p}{2} \rfloor$ (if b is higher than this, there will be forbidden beat intervals of 1 unit). Recall that p is the period of the meter, or the total number of beats on the lowest level ℓ_0. Now, for every $b < b_{\max}$, all possible distributions of beats were determined. Then, for every distribution of beats, all allowed permutations and higher-level organizations were determined. Finally, all possibilities were summed, which gave the total number of structures for the relevant period.

With regard to meeting the constraints (including the temporal constraints **C4**, **C5** and **C6**), the following considerations were taken into account for determining all possible distributions of beats and all their permutations:

- For $b = 1$, there is only one distribution with one permutation and therefore one structure, which is the structure p. Note that $b = 1$ is only possible for $p \leq 12$. This is because with $b = 1$, level ℓ_1 has only one beat, so ℓ_1 cannot be the tactus level by **C3**, hence $c = 0$. But this entails that $u \geq 400$ by **C6**, while for structures with $p \geq 13$, **C5** requires that $u \leq 384.6$.

- For $b = 2$, there is only one distribution as well, which is the distribution that is as close as possible to dividing the structure exactly in half. If p is even, this distribution has one permutation (e.g., 6-6 for $p = 12$); if p is odd, there are two permutations (e.g., 6-7 and 7-6 for $p = 13$). Note that any other distribution of 2 beats over any p is not maximally even.

- For $b = 3$, there is only one distribution as well, which is as close as possible to dividing the structure exactly in three. If p is divisible by 3, there is one permutation (e.g., 5-5-5 for $p = 15$); if not, there are three permutations (e.g., 5-5-6, 5-6-5 and 6-5-5 for $p = 16$).

- For all $b \geq 4$, the possibility of higher order structuring makes structures possible in which tactus level ℓ_1 is not maximally even, but super-tactus level ℓ_2 is (this is not a possibility for $b < 4$ as this would violate **C7** on level ℓ_2). The general procedure for counting the total number of possibilities for $b \geq 4$ is therefore as follows. First, determine all possible maximally even structures; then, determine the structures in which ℓ_1 is not maximally even, but ℓ_2 is. These can be found by dividing p in two parts that are as close as possible to two exact halves. Within these two parts, every combination of beat classes is possible, as long as two requirements are met. First, there is no beat class of length 1 because of **C7**. Second, the longest beat interval is not more than three times as long as the shortest beat interval. This is because on the tactus level (which is ℓ_1 in this testset), the shortest beat interval is at least 400 ms by **C6**, such that three times the shortest beat interval will exceed 1200 ms, violating **C6**. When all possible distributions are given, determine all possible permutations for these possible distributions. For the structures of which ℓ_1 is maximally even, only cyclic permutations need to be taken into account, but keep in mind that these are allowed without and with additional organization on higher levels. Some (but not all) of these organizations lead to double-counting for being maximally even on ℓ_2 as well.

  We present the following example for $p = 15$ and $b = 4$. The only structure that is maximally even on ℓ_1 is 4-4-4-3, which has four permutations, and its higher-organized version 4-4|4-3, which also has four permutations. For the structures that are not maximally even on ℓ_1, the distribution in ℓ_2 is 8|7 (the case 7|8 is taken into account through including all permutations on ℓ_1). As $b = 4$, both beats on ℓ_2 have to be cut into two to get the distribution on ℓ_1. This way, 7 becomes either 5-2 or 4-3 and 8 becomes either 4-4, 5-3 or 6-2. One of the possibilities is 6-2|5-2, in which the following permutations are possible: swapping 6-2 (two options), swapping 5-2 (two options) and swapping the two groups 6-2 and 5-2 (two options). Together, this amounts to $2 * 2 * 2 = 8$ permutations. The other possibilities are as follows (P denotes the number of permutations): 5-3|5-2 ($2*2*2 = 8$P), 4-4|5-2 ($2 * 1 * 2 = 4$P, as 4-4 cannot be swapped), 6-2|4-3 (8P) and 5-3|4-3 (8P). 4-4|4-3 is not taken into account, as this structure was already counted above because it is maximally even on ℓ_1 as well.

The following considerations were borne in mind for higher numbers of b:

- For $b = 5$, higher level organization of structures in which ℓ_1 is maximally even may be either x-x|x-x-x or x-x-x|x-x. Also, the number of permutations for the case in which ℓ_1 is not maximally even increases, as there are different ways to organize super-tactus level ℓ_2. For example, for $p = 15$, we either have that 7 becomes 3-2-2 while 8 becomes 4-4, 5-3 or 6-2, or we have that 7 becomes 5-2 or 4-3, while 8 becomes 3-3-2 or 4-2-2.

- For $b \geq 6$, structures in which ℓ_1 is maximally even can be organized on higher levels in an increasing number of ways, and there may be more than two beats on ℓ_2. For example, for $b = 8$, we have the following. One beat: x-x-x-x-x-x-x-x; two beats: x-x|x-x-x-x-x-x, x-x-x|x-x-x-x-x, 4x|4x, 5x|3x, 6x|2x; three beats: 2x|2x|4x, 2x|3x|3x , 2x|4x|2x, 3x|2x|3x, 3x|3x|2x, 4x|2x|2x; four beats: 2x|2x|2x|2x. For $b \geq 8$, there may also be more than one beat on ℓ_3 (e.g., x-x|x-x||x-x|x-x for $b = 8$). As long as **C7** is met, any organization on higher levels is possible.

- For $b \geq 6$, structures in which ℓ_1 is not maximally even, not only a level ℓ_2 with two beats may stabilize this, but also a level ℓ_2 with three beats. For example, for $p = 16$ and $b = 6$, the distribution in ℓ_2 is either 8|8 or 6|5|5. For 8|8, we either have that one 8 becomes 2-2-2-2 while the other 8 becomes 4-4, 5-3 or 6-2, or one 8 becomes 3-3-2 or 4-2-2 while the other 8 becomes 3-3-2 or 4-2-2 as well. For 6|5|5, we either have that 6 becomes 3-3

b	Possible distributions and permutations
1	$11^{\S}$ (with $c = 0$; 1P, $0^{\S}$)
2	6-$5^{\S}$ (2P, $0^{\S}$)
3	4-4-3 (3P)
4	3-3-3-2 (4P) 3-3|3-2 (4P) 4-3-2-2; not maximally even, but: 4-2|3-$2^{*,\ddagger,\S}$, 4-2|2-$3^{*,\ddagger,\S}$, 2-4|3-$2^{*,\ddagger,\S}$, 2-4|2-$3^{*,\ddagger,\S}$, 3-2|4-$2^{*,\ddagger,\S}$, 3-2|2-$4^{*,\ddagger,\S}$, 2-3|3-$2^{*,\ddagger,\S}$, 2-3|2-$4^{*,\ddagger,\S}$ (8P, $0^{*,\ddagger,\S}$)
5	3-2-2-2-2 (5P) 3-2|2-2-2 (5P, $2^{\dagger}$) 3-2-2|2-2 (5P, $2^{\dagger}$)

Table A.1: All possible structures for meters with period 11 and $c = 1$ (including the structure 11 with $c = 0$). The P-number represents the number of permutations for every distribution. Structures that are not allowed by additional constraints are marked with the following symbols: **C9***, **C10**†, **C11**‡, **C12**§. The number of permissible permutations under the constraints is shown through the same symbols. The total number of structures for $p = 11$ under **C1-8** is 37, and under the additional constraints respectively $29^{*,\ddagger}$, $31^{\dagger}$ and $26^{\S}$.

or 4-2 (the 5s always become 3-2). The number of permutations increases as well: for instance, 3-3-2|4-2-2 has $3 * 3 * 2 = 18$ permutations, and 4-2|3-2|3-2 as $2 * 2 * 2 * 3 = 24$ permutations. Keep in mind that a maximally even ℓ_2 may divide the number of beats on ℓ_1 in an uneven way (e.g., 4-4|2-2-2-2).

- For $b \geq 8$ with structures in which ℓ_1 is not maximally even, also a level ℓ_2 with four beats may compensate for this. This is theoretically possible from $p = 16$, but matters first at $p = 18$ (as for all structures such that $b = 8$ and $p = 16$ or $p = 17$, ℓ_1 is maximally even). Keep in mind that ℓ_2 with four beats means that there might be an ℓ_3 with two beats, so for every permissible structure x-x|x-x|x-x|x-x there is also a permissible structure x-x|x-x||x-x|x-x.

- Some structures are symmetric and thereby have less cyclic permutations than the value of b suggests. For example, the maximally even structure 3-3-2|3-3-2 has only three distinct cyclic permutations, even though $b = 6$. Note however that structures like 3-3-2-3-3-2, 3-3|2-3-3-2 and 3-3|2-3|3-2 are all unique structures (which all have three cyclical permutations).

- Beware of double-counting structures in which ℓ_1 is maximally even that may emerge from the systematic construction of structures in which ℓ_2 is maximally even. Some examples of this are more obvious than others. An example of a less obvious duplicate is 3-3|2-3|3-2 (maximally even on ℓ_1) as one of the permutations of structure 3-3|3-2|3-2 (not maximally even on ℓ_1).

An example of construction of the testset for period 11 is shown in table A.1.

A.2 Testset construction without C8

The testset version for **C1** to **C7** (without **C8**) was constructed in a more combinatorial way, as considerations of maximal evenness did not play a role. Like above, the testset was constructed for all possible values of b given p. Then, the different distributions were determined by dividing the beats from the extreme uneven to an increasingly more even way. It was taken into consideration that the longest beat interval of the tactus level cannot be more than three times the shortest beat interval of that level by **C6**. It was also taken into consideration that no beat interval has a length of 1 unit by **C7**. All possible permutations were taken into account, which was done by combinatorics (e.g., there are 12 permutations of the sequence x-y-z-z and 24 permutations of x-y-z-a). Summation of the possible permutations per distribution delivered the total of all ℓ_1-possibilities for given p and b. Subsequently, this number was multiplied by the number of possibilities for higher level organization. This number depends on b: for $b = 4$, there are 2 possibilities (with or without an extra beat on ℓ_2); for $b = 5$, there are 3 (a beat such that x-x|x-x-x, a beat such that x-x-x|x-x or no beat, hence x-x-x-x-x); for $b = 6$, there are 5 (x-x|x-x-x-x, x-x-x|x-x-x, x-x-x-x|x-x, x-x|x-x|x-x and x-x-x-x-x-x). For higher $b = 7, 8, 9$ there are respectively 8, 14 and 25 possibilities of higher level organization.

For instance, for $p = 15$ with $b = 4$, the possibilities are as follows: 6-5-2-2 (12P), 6-4-3-2 (24P), 6-3-3-3 (4P), 5-5-3-2 (12P), 5-4-4-2 (12P), 5-4-3-3 (12P) and 4-4-4-3 (4P), giving a total of 80. Multiplication by 2 because $b = 4$ gives 160 structures for $p = 15$ with $b = 4$.

An example of the testset for $p = 11$ for all b is given in Table A.2.

A.3 Testset construction with additional constraints

The behavior of the testset under the additional constraints was analyzed by taking the testset for **C1** to **C8**, selectively marking the structures that are not allowed by the additional constraint in question and counting the number of structures that remain. Often, full groups of structures can be eliminated this way, but in some cases, only some permutations are eliminated. The guidelines for elimination are listed below.

- **C9**: Eliminate all structures in which there is a beat class ratio of 2:1 or higher on any level. On ℓ_1, these are all the structures in which 4s or larger are combined with 2s, 6s or larger with 3s, etc., such that **C9** eliminates all permutations for that distribution. Apart from that, structures in which the higher level grouping gives rise to ambiguity or contradiction are also eliminated (e.g., 3-2-3|2-2, as this gives 8|4 on ℓ_2, or 2-2-2-2|2-2 for the same reason). For some structures in which ℓ_1 is maximally even, this only holds for some cyclic permutations (e.g., 3-2-2-3|2-2 is eliminated while 2-3-2-2|3-2 is not).

- **C10**: Eliminate all structures in which the highest level has two or three elements but is not maximally even. For instance, this eliminates three of the cyclical permutations of 2-2|3-2-2, but not the two other cyclic permutations 3-2|2-2-2 and 2-3|2-2-2.

- **C11**: Eliminate all structures in which there are three or more beat classes. On ℓ_1, this always eliminates entire groups of distributions, but on higher levels, this may depend on cyclic permutations. The latter situation only happens for structures with $p \geq 15$, since then, ℓ_1 may be maximally even such that ℓ_2 may have a distribution with three different beat intervals (e.g., 3-2|2-2-2|2-2, but not its cyclic permutation 2-2|3-2-2|2-2).

- **C12**: This constraint acts as a meta-rule on temporal frame, but an equivalent object-level rule system can be derived through analyzing temporal frames for groups of structures. This rule system is listed below, along with its respective motivations.

  - Eliminate the structures 10, 11 and 12 from the testset.

    Recall that for $c = 0$, $\mathcal{I}(c) = \{u\}$. Now, by **C6**, $u \geq 400$. But for $p \geq 10$, $u \leq 5000/10 = 500$ by **C5**. Therefore, $400 \leq u \leq 500$, such that **C12** is not met, as there are no u' and u'' such that $400 \leq u'' \leq u' \leq 500$ while $u' > u'' + 100$.

  - Eliminate all structures that have a 6 on ℓ_1.

With a 6 on the tactus level, it must be that $u \leq 200$, as otherwise there is an i in $\mathcal{I}(c)$ such that $i > 1200$, which fails **C6**. But by **C4**, $u \geq 100$, and therefore $100 \leq u \leq 200$ such that **C12** is not met.

- Eliminate all structures in which there is a beat class ratio of 2:1 or higher on ℓ_1 (the beat class ratio on higher levels is less restricted).

  Suppose that there is a beat class ratio on ℓ_1 that is at least 2:1. Suppose now that the shortest beat at the tactus level is j temporal units, so $j * u$. Now, the longest beat is at least $2j * u$. By **C6**, for all $i \in \mathcal{I}(c)$ it must hold that $400 \leq i \leq 1200$. In particular, $j * u \geq 400$ while $2j * u \leq 1200$. Now, $u \geq 400/j$, but $u \leq 1200/2j$, so $400/j \leq u \leq 600/j$. For any $j \geq 2$ ($j \neq 1$ by **C7**), u must be between two values that are less than 100 ms apart. Hence, **C12** is not met.

- For $p \geq 17$, eliminate all structures that have a 2 on ℓ_1.

  Since $p \geq 17$, $u \leq 5000/17 = 294.1$ by **C5**, while with a 2 on the tactus level, it must be that $u \geq 200$ (otherwise there is an i in $\mathcal{I}(c)$ such that $i < 400$, which fails **C6**). Therefore, $200 \leq u \leq 294.1$ such that **C12** is not met.

- For $p \geq 22$, eliminate all structures that have a 3 on ℓ_1.

  Since $p \geq 22$, $u \leq 5000/22 = 227.3$ by **C5**, while with a 3 on the tactus level, it must be that $u \geq 133.3$ (otherwise there is an i in $\mathcal{I}(c)$ such that $i < 400$, which fails **C6**). Therefore, $133.3 \leq u \leq 227.3$ such that **C12** is not met.

- For $p \geq 25$, eliminate all structures.

  Since $p \geq 25$, $u \leq 5000/25 = 200$ by **C5**, while $u \geq 100$ by **C4**. Therefore, $100 \leq u \leq 200$ such that **C12** is not met.

Because of the strictness of the rule system, the testset under **C12** has been constructed as well for $p > 18$, as for these structures, **C12** only allows combinations of 5s, 4s and 3s.

- **C9** + **C12**: Eliminate all structures that are not allowed by **C9** or **C12** (or both). For $p > 18$, construct the testset under **C12** as above and eliminate all structures that are not allowed by **C9**.

All possible structures for $p = 11$ under the different additional constraints are given through the symbols in Table A.1.

b	Possible distributions	No. P	$*$	Total
1	11 (with $c = 0$)	1	1	1
2	8-3	2	1	2
	7-4	2	1	2
	6-5	2	1	2
3	6-3-2	6	1	6
	5-4-2	6	1	6
	5-3-3	3	1	3
	4-4-3	3	1	3
4	5-2-2-2	4	2	8
	4-3-2-2	12	2	24
	3-3-3-2	4	2	8
5	3-2-2-2-2	5	3	15

Table A.2: All possible structures under **C1-7** for meters with period 11 and $c = 1$ (including the structure 11 with $c = 0$). 'No. P' denotes the number of permutations, '$*$' the higher level multiplier. The total number of structures for $p = 11$ is 80.