# Neurocognitive entrainment to meter influences syntactic comprehension in music and language: An individual-differences approach

DISSERTATION
Submitted to the Faculty of Human Sciences
of the University of Potsdam
for the completion of the degree
Doctor of Philosophy (PhD)

by
Eleanor Elizabeth Harding, M.Sc.

2016

supervised by
Prof. Dr. Sonja A. Kotz and Dr. Daniela Sammler

# Acknowledgements

Many people have contributed to the production of this work. Foremost I would like to thank my primary advisor, Prof. Dr. Sonja Kotz, for her creative vision and personal faith in me as a researcher. To Dr. Daniela Sammler, my second advisor, I owe a quality standard. Without her dedicated feedback and insightful criticism this work would be less than what it is. To both of them I also owe the opportunity and financial resources to have conducted this research. I am also grateful to the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig, where the core of this work took place, particularly for the resources of the Neuropsychology department, headed by Prof. Dr. Angela Friederici. I am also grateful to Peter Keller for being the external reviewer, as well as some insightful comments he made in the stage of designing the stimuli. I also extend thanks to the "Europe Brain and Music" network for your opportunities and feedback at various stages of this project.

To Dr. Molly Henry I give thanks, for her expertise and help with the EEG frequency analysis, as well as her humor and friendship. I thank Prof. Dr. Carrie Jackson for her feedback at a sensitive and critical turning point in this work. For their daily comradery, friendship and professional support I also would like to thank Drs. Iris Knierim, Julia Merrill, Paula Roncaglia-Denissen, Marieke van der Steen, and Laura Verga. Thanks also to the other members of "Subcortical Contributions to Comprehension" for daily support and humor. I'd also like to generally thank the research groups "Music Cognition and Action" and "Auditory Cognition," for their openness to the discussion of thematic and empirical issues that cropped up along the way.

I would like to thank my interns for their hours helping with stimulus preparation, data acquisition, and data analysis: Mareike Augsburger, Christina Engelmann, Anja Ruisinger, and Elisabeth Wente. I am also grateful to the technical assistants Heike Böthel and Kristiane Klein, for their help in organizing participants and in the EEG laboratory.

Finally, I thank my family for supporting my person during this endeavor. My four parents, the late Ruth Borkes, Steve Borkes, Leonard Harding, and Sandra Harding, as well as my parents in-law, Jan Schuurmans and Tineke Scherer, have provided me with emotional and practical support — so have countless immediate and extended family on both sides of the Atlantic. And thanks (a word that isn't sufficient) to Edwin and Lillian, my husband and daughter, for your most important love.

# Contents

# Chapter 1

# Introduction

Entrainment is an everyday phenomenon, such as having ones attention captured by the pounding gavel of a judge in a noisy courtroom, or noticing the flashing turn signal of a vehicle on the road. This dissertation looks at the underlying neurocognitive mechanisms of entrainment, and how they can be harnessed to aid comprehension-related processes. Here, existing research that motivated the current approach is presented, showing how the focus of the investigation was specified to meter and syntax in music and language. The current approach is then briefly introduced, and finally an overview of the remaining chapters is provided.

## 1.1 Motivation

### 1.1.1 Music meter remediates language syntactic comprehension in Parkinson's Disease patients

Patients with Parkinson's disease (PD) are reported to have language syntactic comprehension deficits (e.g., Lieberman, Friedman, & Feldman, 1990). For example, when sentences with ill-formed syntax are presented to PD patients, an ERP component that is typically elicited in healthy adults, the P600 (e.g., Osterhout and Holcomb 1992; see Section 3.1.2), is pathologically missing (e.g., Friederici, Kotz, Werheid, Hein, & von Cramon, 2003). Two studies, however, have shown that a P600 may be reinstated when, before hearing syntactically incorrect sentences, PD patients first listen to a musical excerpt of a march (Kotz & Gunter, 2015; Kotz, Gunter, & Wonneberger, 2005). More precisely, when German PD patients were cued with a musical march, they showed a P600 effect time-locked to a categorical syntax error in spoken German sentences (such as "gegessen/eaten" in "Das Eis wurde im__gegessen/ The ice cream was in__ eaten").

The musical meter of a march contains groups of two beats in a strong-weak dynamic

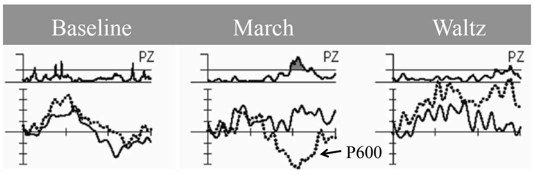



*Figure 1.1: A key phenomenon that inspired the current dissertation topic is that listening to music can remediate language syntactic comprehension in PD patients. Only the march, which shared a metrical strong-weak pattern with presented German sentences, reinstated the P600. Upper axes show effect difference wave and significance threshold, lower axes show amplitude per condition. Figure adapted from Kotz and Gunter (2015).*

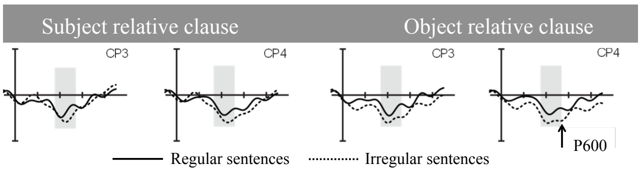



*Figure 1.2: A second key phenomenon to motivate this work is that in sentences, regularized language meter can facilitate syntactic integration (significantly reduced P600 in regular sentences, indicated by arrow). Figure adapted from Roncaglia-Denissen et al. (2013).*

relationship and is symmetrical to the meter of spoken German (Eisenberg, 1991; Féry, 1997). While results of both the PD patient group study (Kotz et al., 2005) and subsequent PD case study (Kotz & Gunter, 2015) found a march to reinstate the P600, the case study reported that a waltz, whose meter is strong-weak-weak, did not have the same reinstating effect (and neither did a baseline condition, with no musical cue; Figure 1.1). Thus, when a temporally predictable auditory cue such as a march excerpt contains a metrical structure that maps onto linguistic meter, the PD patient may be able to realign otherwise impaired internal temporal processing with external linguistic events, such that temporal expectation of strong beats allocates attention to the occurrence of linguistically salient, strong syllables. This latter concept is central to the dissertation and will be further explored in the coming sections.

### 1.1.2 Language meter facilitates language syntactic integration

Not only musical meter preceding a march, but also the regularity of language meter can impact syntactic comprehension. A recent EEG study (Roncaglia-Denissen, Schmidt-Kassow, & Kotz, 2013) showed that regular meter can facilitate syntactic integration more than irregular meter. Healthy non-musician adults were presented with sentences, which established a regular or irregular metrical context before introducing a local syntactic ambiguity (2 x 2 design; Table 1.1).

As reported elsewhere (Frisch, Schlesewsky, Saddy, & Alpermann, 2002), a larger P600

*Table 1.1: Stimuli from Roncaglia-Denissen et al. (2013). Apostrophes mark strong syllables. The metrical contexts were constructed with either regular or irregular positioning of strong and weak syllables. The local syntactic ambiguity consisted of an ambiguous "die/who" that referred alternatively to "they" (subject) or "them" (object) at the onset of a relative clause. The ambiguity was resolved by a verb conjugation, which showed whether "die/who" was the agent (subject) or recipient (object) of the action.*

| | |
|---|---|
| **Regular meter** | |
| 'Ro-land trifft die 'Die-ner, die An-'to-nio mal ge-'stört ha-ben, im 'Park | **Subject first** |
| *Roland meets the helpers, who once bothered Antonio, in the park.* | |
| … die An-'to-nio mal ge-'stört hat, im Ge-'schäft. | **Object first** |
| *…, whom Antonio once bothered, in the store.* | |
| **Irregular meter** | |
| 'Bern-hard trifft die Ge-'hil-fen, die Ni-'cole mal ge-'stört haben, im 'Park. | **Subject first** |
| *Bernhard meets the assistants, who once bothered Nicole, in the park.* | |
| ..., die Ni-'cole mal ge-'stört hat, im Ge-'schäft. | **Object first** |
| *…, whom Nicole once bothered, in the store.* | |

effect was elicited when the local ambiguity was disambiguated as an object relative clause. Importantly here, the P600 effect was significantly reduced when the meter was comprised of predictably grouped syllables (Figure 1.2). This was interpreted such that the regular meter provided a segmentation cue that helped the mental syntactic parsing, easing the processing cost of the ambiguity.

The finding that a regular metrical context reduced an ERP effect (compared to a less predictable metrical context) has also previously been reported in paradigms investigating language phoneme perception (Cason & Schön, 2012) and language semantic perception (Rothermich, Schmidt-Kassow, & Kotz, 2012). In those previous studies, the ERP effect reduction (a P300[1] and an N400, respectively) was interpreted such that a regular metrical context allowed for accurate temporal expectation of salient linguistic features, thus easing processing difficulty, since attention was appropriately allocated to task-relevant information. Though Roncaglia-Denissen et al. (2013) explained their processing ease such that regular metrical grouping of syllables created small, predictably sized units optimally "guiding" the syntactic parser through large amounts of information, equally applicable is the explanation from Cason and Schön (2012), Rothermich et al. (2012), and Kotz and Gunter (2015), that the regular meter optimally primed temporal allocation of attention for receipt of salient linguistic information.

[1]But see Schwartze, Rothermich, Schmidt-Kassow, and Kotz (2011) for the opposite reported impact of temporal regularity, an increased P300 effect, in an auditory oddball paradigm.

### 1.1.3 A hypothesis for shared music and language syntactic integration resources

Music and language are hypothesized to have common neurocognitive resources that integrate syntactic constituents. An influential theory in music-language cross-domain research is the Shared Syntactic Integration Resource Hypothesis (SSIRH; Patel, 2003, see Section 3.1.3). Based on accumulated theoretical and empirical research in respective music and language cognition fields, the SSIRH proposed that overlapping neural resources access separately stored musical and linguistic syntactic representations, and that these resources are at play in the online integration of syntactic constituents in both domains. Main observations supporting the online shared resource hypothesis were theoretical similarities in music and language syntax, empirical evidence that online processing of syntactic violations activated similar anatomical structures, comparable ERP components elicited, and induced interference effects when presented simultaneously. The separate storage of syntactic representation is consistent with reports of aphasia (language deficits) and amusia (music deficits) where the other domain's syntax processing remained intact.

Language syntactic comprehension was shown above to be influenced by both musical and linguistic meter (Sections 1.1.1 and 1.1.2). If music and language rely on shared syntactic resources as the SSIRH claims, and if meter influences language syntactic comprehension, then by extension meter should similarly influence music syntactic comprehension. The same explanation also applies, i.e., that meter directs resources to salient, temporally predictable events when integrating music syntactic constituents. While seemingly common sense, this particular cross-domain paradigmatic angle has up to this point never been specifically tested, and is pioneered in this dissertation.

### 1.1.4 Individual differences in auditory perception affect meter perception and syntactic comprehension

One last phenomenon in the literature motivates the approach taken in this dissertation: individual differences in various aspects of auditory perception are reported to impact both meter perception and syntactic comprehension. Particularly musical expertise, temporal perception, and working memory are reported to play a role in either or both meter perception and syntactic comprehension.

Meter perception is influenced by individual differences in musical expertise and temporal perception ability. Regarding musical expertise, musicians (compared to non-musicians) have better awareness of metrical hierarchy in music and better perception of small metrical units in both music and language (Geiser, Sandmann, Jäncke, & Meyer, 2010; Magne et al., 2007; Marie, Magne, & Besson, 2011; Palmer & Krumhansl, 1990). Regarding temporal perception, ERP studies have reported temporal encoding of syllables to drive meter perception in language (Magne et al., 2007; Marie et al., 2011); if temporal encoding is necessary to perceive linguistic meter, then it follows that participants with better timing ability can consequently perceive meter better than their counterparts with worse timing ability. Though meter is reported to impact temporal perception in music paradigms (Ellis & Jones, 2010; Grube & Griffiths, 2009), it remains to be seen whether temporal perception modulates meter perception in music. There is a paucity of literature directly assessing whether or how working memory is linked to meter perception in music or language; the current investigation may provide evidence for a role of working memory in meter perception across domains.

Syntactic comprehension is impacted by individual differences in musical expertise, temporal perception, and working memory. Musical expertise is associated with better early violation detection and late integration of syntax in music (Besson & Faïta, 1995; Koelsch, Schmidt, & Kansok, 2002), and better early violation detection and learning of hierarchical structures in language (Brod & Opitz, 2012; Fitzroy & Sanders, 2013). Temporal perception impacts syntactic event detection in music and language (Tzounopoulos & Kraus, 2009; White-Schwoch & Kraus, 2013), and larger-scale temporal features of both music and language impact syntactic comprehension (Schmidt-Kassow & Kotz, 2008; Schmuckler & Boltz, 1994). Working memory is positively correlated with music syntax task performance in music (Fiveash & Pammer, 2014), and as is particularly relevant to the current investigation, improved working memory is associated with improved reanalysis of ambiguous syntactic structure in language (Friederici, Steinhauer, Mecklinger, & Meyer, 1998; Vos, Gunter, Schriefers, & Friederici, 2001).

Since individual differences in musical expertise, temporal perception, and working memory affect meter perception and syntactic comprehension in such multifaceted ways, a neurocognitive account for meter impacting syntactic integration, in both music and language domains, should optimally acknowledge the role of individual differences and monitor them in empirical investigation.

## 1.2 Current approach

This dissertation was motivated by these four points in the literature, which together suggested meter to impact syntactic integration in both music and language domains, subject to individual differences. The current investigation therefore explored a possible unifying explanation for these observations (Figure1.3): Firstly, the brain perceives and entrains to meter in music and language, with progressively greater entrainment as metrical regularity increases. Secondly, entrainment to meter facilitates syntactic integration in both domains.

Finally, individual differences in meter perception and syntactic comprehension may influence this process.

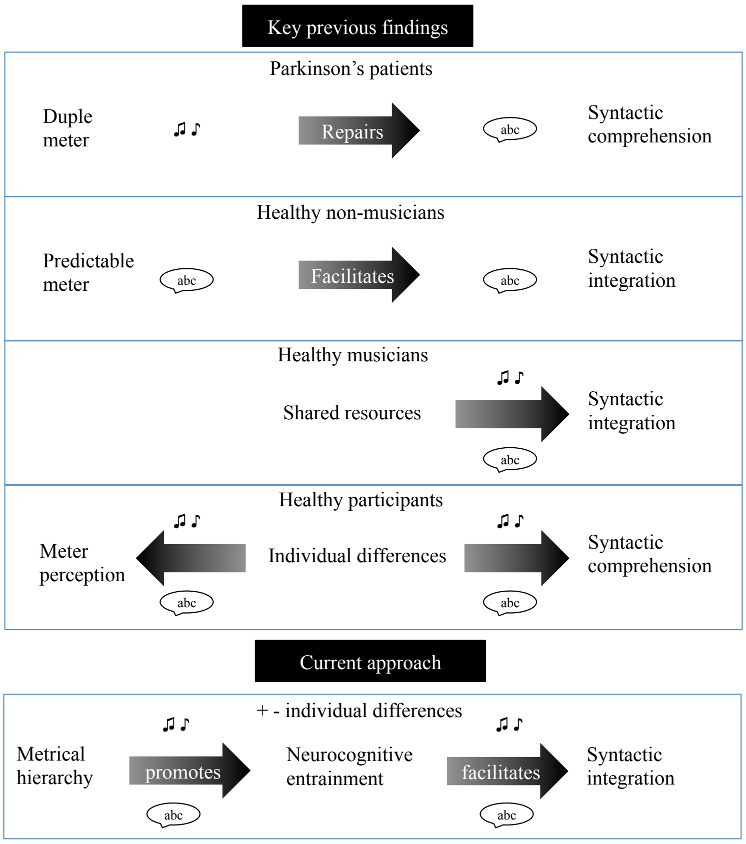



*Figure 1.3: An explanation for key phenomena observed in the literature: Neurocognitive entrainment is encouraged by the temporal predictability in music and speech that metrical regularity can provide. This in turn facilitates the integration of syntactic units. Both processes are subject to individual differences.*

## 1.3 Overview of chapters

The theoretical framework supporting the current approach will be introduced in the coming chapters. **Chapter 2** explains how multiple cognitive- and neural-based accounts of entrainment can comprise a neurocognitive entrainment framework, which should have similar implications in music and language domains. Chapter 2 also outlines basic principles of meter in music and language, and aims to present a representative sample of relevant EEG and MEG studies that show entrainment to meter in both domains. **Chapter 3** outlines basic principles of syntax in music and language domains, and introduces background syntax-ERP literature relevant to the current investigation. Then, the neurocognitive entrainment framework is extended, and mechanisms that could underpin entrainment's impact on syntactic integration are suggested. Next, the importance of the individual-differences approach is highlighted in **Chapter 4**, which presents empirical evidence of how individual differences in musical expertise, temporal perception, and working memory could impact both meter perception and syntactic comprehension, in both music and language domains.

The methods used for the current approach are introduced in **Chapter 5**. Studies 1, 2 & 3 then proceed to systematically address cognitive entrainment, neural entrainment, and syntactic integration while taking into account individual differences. In Study 1 (**Chapter 6**), Experiment 1a statistically evaluates the psychological processes that are cognitively linked to diagnostic scores collected from all participants, and provides each participant with two 'cognitive factor' scores to be correlated with further experimental measures. Experiment 1b validates the stimuli, to assure that the syntactic structures are discriminable (via the task), while doubly serving to assess behaviorally whether groups of musicians and non-musicians cognitively entrain to the regular-meter melodies and sentences in a 2 x 2 design with factors Metricality (regular, irregular) by Domain (music, language). Study 2 (**Chapter 7**) evaluates the frequency spectrum of both stimuli (audio files) and EEG for evidence of frequencies representing the group or individual level of the metrical hierarchy. Experiment 2a assesses which peaks (if any) represent levels in the metrical hierarchy of stimuli, and Experiment 2b assesses whether peaks in EEG reflect neural entrainment to hierarchical levels in stimuli. The amplitude of found peaks is compared in a 2 x 2 x 2 design of factors Metricality (regular, irregular) by Hierarchy (group, individual) by Domain (Music, Language). Study 3 (**Chapter 8**) investigates the comparability of syntax processing in both domains in an ERP analysis, and whether entrainment impacts syntactic integration comparably across domains. Experiments 3a and 3c compare amplitudes of found ERP effects respective in music and language domains, both in a 2 x 2 design with factors Metricality (regular, irregular) by Syntax (preferred, non-preferred). Experiment 3b pilots German relative clause attachment preference for the purposes of coding the preferred syntax conditions in Experiment 3c. The 'individual differences' approach to this dissertation on the one hand manifests in stimuli parameters which highlight individual differences in meter perception and syntactic comprehension, and on the other hand the correlation of participants' cognitive factor scores (from Experiment 1a) with experimental results across all three studies.

The results from the three studies are discussed with a broad perspective in **Chapter 9**, and **Chapter 10** suggests some future investigations that could continue this research.

# Chapter 2

# Entrainment to meter

Entrainment is simply the phase coupling between two oscillators, such as two pendulums gradually shifting their trajectories until they are swinging in sync. This chapter focuses on human perceptual entrainment to meter, offering mutually compatible accounts of cognitive and neural entrainment. Metrical theory, as well as empirical findings supporting meter perception and entrainment to meter, are outlined for both music and language domains. Finally the meter-related aspects of the current EEG paradigm are introduced. Though perceptual principles introduced here may apply to multiple modalities, the current scope is limited to the auditory modality.

## 2.1 Entrained attention

This section introduces two hypothesized accounts of attention that monitors auditory signals. The dynamic attending theory (DAT; e.g. Jones, 1976) proposes attending that qualitatively adapts to the degree of temporal predictability in the signal. The attentional bounce hypothesis (Pitt & Samuel, 1990) proposes that attention to speech is modulated by the location of stressed syllables. These two accounts are compatible, in that the attending mechanism proposed in the DAT could also subserve attention 'bouncing' among stressed syllables.

### 2.1.1 The dynamic attending theory

The dynamic attending theory (DAT) accounts for online (dynamic) attention, or attending, to the environment. Originally conceived by Jones (1976) and fine-tuned in decades since (Jones, 2008; Jones & Boltz, 1989; Large & Jones, 1999), the theory proposes two alternate forms of attending: 1) future-oriented attending, which relies on temporal predictability inherent in external stimuli to guide attention to future points in time, and 2)

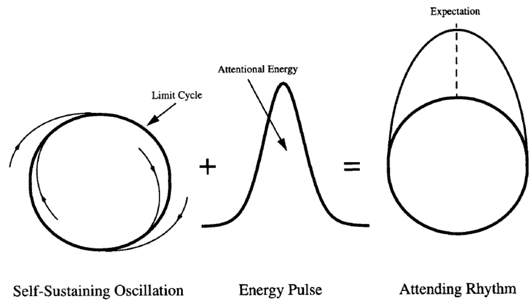



*Figure 2.1: The attentional energy pulse together with the self-sustaining attentional oscillation or limit cycle represents the attending rhythm, which is entrained to a stimulus when its temporal expectation is reinforced. Figure from Large and Jones (1999).*

analytic attending, a more local-level processing which monitors signals that are not temporally reliable (e.g. missing a temporal hierarchy, or multiple nested levels of predictable temporal intervals). It has recently been suggested that these two attending mechanisms exist on a continuum as opposed to being discrete processes, the most appropriate attending driven by stimulus properties (Henry & Hermann, 2014).

Future-oriented attending may be described by a canonical oscillator model (Large & Jones, 1999). According to this model, cognitive attention is represented by an attentional energy pulse. When the attentional energy peaks, the system is most perceptive of the stimulus being attended. The timing and sharpness of the attentional pulse relies on a second feature of the attending system—an oscillation underlying the attentional pulse. This underlying oscillation is self-sustaining (a stable limit cycle), in the absence of stimuli idling at its natural cycling rate (with an inherent period), but when exposed to an external stimulus with temporal regularity it synchronizes its phase with the available phase information from the signal. Once the internal attentional oscillation is synchronized to the external signal, the attentional pulse, which occurs at some regular point 'on top of' the attentional oscillation (Figure 2.1), is thereby aligned with the external signal and attention is entrained. Phase and period correcting components to this model allow future-oriented entrainment despite temporal perturbations in the signal.

Analytic attending does not capitalize on such entrained internal oscillations per se. When no temporal hierarchy exists in the incoming signal, the analytic attending mode rather relies on the next best predictable, non-temporal task-related feature in the incoming signal by which to structure the information (Jones & Boltz, 1989). For example, if the task is related to temporal judgments in the absence of temporal hierarchy, a strategy employed in the analytic attending mode might be counting beats. Further research would be required to clarify whether analytic attending to non-temporal, highly predictable features may also be described by a canonical oscillation model.

If the environmental stimulus has a hierarchy of periodic events such as a series of

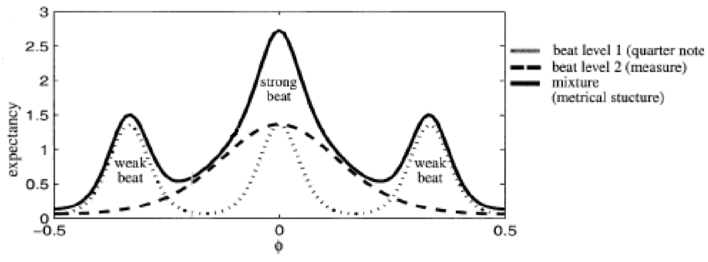



*Figure 2.2: The behavior of the attentional oscillation incorporates nested metrical levels, increasing expectancy at metrically strong beats where the coupled oscillations overlap. Figure adapted from Large and Palmer (2002).*

beats with a metrical organization, future-oriented attending processes are engaged and the attentional oscillations become nested (underscored in Jones, 2008). That is, discrete attentional oscillations will follow the different metrical levels of the stimuli, joining where the metrical levels overlap, e.g. at strong beats in the meter. The attentional pulse will be greater at strong beats where multiple oscillations overlap. Large and Palmer (2002) detailed the behavior of attentional energy during nested attentional oscillations (as a function of expectancy—the system provides attentional pulses when salient events are expected), shown in Figure 2.2 in the case of entraining to music with a ternary meter.

Suggestions as to the natural cycling rate of the DAT's attending oscillator may be inferred from a comprehensive tapping study (McAuley, Jones, Holub, Johnston, & Miller, 2006). Based on a composite measure of preferred listening tempo and motor tapping tempo, the preferred oscillation period among healthy adults was approximately 1.7 Hz. This preference was shown to systematically slow down over the human lifetime, ranging from approximately 3.3 Hz in small children to 1.4 Hz in older adults. Considering that this preferred period corresponded with the greatest attending accuracy (lowest tapping variability in the greatest range of tempi), the preferred period determined by this study is an eligible candidate to represent the natural cycling rate of the attending oscillator proposed in the DAT.

### 2.1.2 Attentional bounce hypothesis

Another theory of attention has emerged specifically in the context of attending to speech, the attentional bounce hypothesis (Pitt & Samuel, 1990). This theory claims that in speech, stressed or strong syllables encourage more attentional salience than unstressed or weak syllables, and attention to stressed syllables is boosted by a context of predictable stressed syllables. Particularly relevant to the current approach, this theory implies that predictably located beats in metrical context can guide attention when participants listen to a speech stream.

The attentional bounce hypothesis, established in a behavioral study with three experiments (Pitt & Samuel, 1990), explains findings that participants were better able to identify target phonemes that were located in stressed as opposed to unstressed syllables, and that furthermore, target phonemes were significantly better perceived when occurring in strong syllables that appeared in an extended context of consistent syllable stress patterns (word lists) compared to metrically unstructured sentence contexts. Participants identified target phonemes (e.g. /p/ or /m/) in two-syllable target words (e.g. 'permit'). Target words changed lexical meaning depending on whether the first or second syllable contained stress (e.g. 'permit' is a noun when the first syllable is strong, PERmit, and a verb when the second syllable is strong, perMIT), and contexts for the target words were established either in full sentences or with word lists. Importantly, target words were themselves digitally processed to have equal stress on both syllables—the surrounding lexical context was designed to generate expectations as to which syllable was stressed. Thus sentences generated expectancies for target word stress based on word category (in the first experiment, e.g. a preceding 'a' generates a noun expectancy), and word lists generated expectancies for target word stress based on the stress patterns of preceding words (in the second experiment, e.g. 'Olive, VILlage, TISsue, KAYak' should encourage PERmit). A third experiment controlled for extraneous word-position effects. Behavioral results (error rates and reaction time) in the sentence experiment showed a general trend that participants performed better when target phonemes were located in (perceived) stressed syllables. In the word list experiment, effects were comparatively boosted (significantly improved perception of target phoneme in stressed syllables), demonstrating that a metrically structured context generated stronger word-stress expectancies. The control experiment demonstrated that word position had little effect, validating the word-list results.

The notion that attention 'bounces' to predictable strong-syllable locations (Pitt & Samuel, 1990) may be rephrased in terms of the oscillator model proposed to describe future-oriented dynamic attending (e.g., Large & Jones, 1999): the attentional oscillator entrains to the (relatively) temporally consistent strong syllable, allowing attention to peak at the strong syllables that in turn improves detection of task-relevant phonemes. Moreover, the fact that contexts with less predictably placed strong syllables demonstrate less attention to strong syllables (Pitt & Samuel, 1990) is consistent with analytic attending, the DAT's alternative to future-oriented attending. Thus language meter can potentially also set up an 'entrainable' metrical context consistent with the above musical meter example in Section 2.1.1 (Figure2.2).[1]

Two mutually inclusive theories, DAT and attentional bounce hypothesis, consistently

[1] At least in stress-timed Germanic languages (Abercrombie, 2006); syllable-timed languages such as French have stricter stress rules, such that the final syllable of an utterance receives stress, rendering such 'musically' metrical structures more difficult.

describe attention entrained to temporally predictable auditory cues. The next section focuses on the neural aspect of entrainment.

## 2.2 Entrained neurons

This section describes several aspects of neural entrainment: Neurons entrained to a periodic signal can impose increased order on internal communication of neural assemblies, thereby improving the general efficiency of cognition (Singer, 2013). Neural entrainment to external signals is suggested to be propagated by oscillations in the delta-theta range (1–8 Hz) and those internal oscillations in turn phase-couple with gamma oscillations as a means to direct attention to stimuli (e.g., Schroeder & Lakatos, 2009). Finally, once neurons entrain to a (rhythmic) external source, their oscillation frequencies represent an embodied perception of metrical levels (e.g., Large, 2008). These concepts are expounded below.

### 2.2.1 Synchronous neural populations in cognition

Neurons entrain to each other as well as to external rhythmic sensory sources (Thut, Schyns, & Gross, 2011), such as direct current, a visual stimulus, or as is relevant here, an auditory stimulus. Communication among neural assemblies occurs via synchronized firing rates in local and distant populations, and compounding phase-coupled or entrained neural assemblies have been proposed to be the basis of cognition (Varela, Lachaux, Rodriguez, & Martinerie, 2001). Although much empirical work needs to be done to trace exact behavior of neuronal populations from sensory-level entrainment to altered higher-order functioning (such as entrained attention), a general explanation presents itself from a recent review by Wolf Singer (2013). Singer described the way that cortical networks in high-dimensional state space oscillate across a broad frequency spectrum, and hypothesized that when networks increase their coherence via entrainment to an input, this imposed order (narrowing multi-frequency-band oscillations to a more uniform meaningful frequency band among multiple networks via synchronization) becomes salient and impacts downstream processing. Thus the functional roll-off effects of neural entrainment to an external source may spread from sensory-level coupling to complex cognitive processes via increased synchronization among dynamic neuronal populations. This view is supported, for example, by observations of increased neural synchrony in learning-related networks when participants listen to sung word lists (with rhythmic and temporal coherence) as opposed to temporally unregulated spoken word lists in memory recall tasks (Thaut, Peterson, & McIntosh, 2005).

### 2.2.2 Delta, theta and gamma coupling

There have been specific mechanisms proposed to describe the coupling of neurons in the auditory pathway to acoustic input such as musical rhythms or speech. Schroeder and Lakatos (2009) proposed that neural oscillatory activity in delta and theta ranges couple to rhythmic acoustic cues such as a beat, and that in a resulting 'rhythmic listening mode' the phase of the three frequency bands delta (1–4 Hz), theta (4–8 Hz) and gamma ( 30–80 Hz) couple in hierarchical fashion, such that delta phase (coupled to the external stimulus) modulates theta activity, which in turn modulates the inhibition/excitation of gamma activity, which is associated with attention. An alternative 'vigilance listening mode' occurs in the absence of rhythmic stimuli, where lower frequency activity is suppressed and gamma activity is in more constant excitation. Thus the sensory coupling in the lower frequency range serves as an intermediary to assure attentional resources peaking at useful temporal intervals. Henry and Hermann (2014), noting the striking compatibility of this account with DAT, proposed that the neural oscillatory hypothesis may be the neural mechanism that underlies the internal attentional oscillator (the delta/theta activity) and attentional peak (gamma activity) described in association with dynamic attending (e.g., Large & Jones, 1999), the rhythmic mode corresponding to future-oriented attention and the vigilance mode corresponding to analytic attending. Similarly, a model of speech perception (Giraud & Poeppel, 2012) focused on the entrainment of theta-range neural oscillations to temporal cues in the spoken syllable, which were hypothesized to couple to gamma-range neurons that in turn encoded the phonetic information for further cognitive processing. Thus the mechanism of attentional modulation by temporally predictable cues is compatible with accounts of neural entrainment in both musical (Schroeder & Lakatos, 2009) and spoken (Giraud & Poeppel, 2012) contexts.

### 2.2.3 Neural resonance

Neural entrainment has also been claimed to embody the perception of musical meter.Large and Kolen (1994) demonstrated that oscillatory units (with bounded frequency ranges) can entrain to driving rhythms; several units in a system were able to entrain simultaneously to different metrical levels, and the authors proposed that such a system of oscillators could describe neurons coupling to various metrical levels. Specifically the nonlinear nature of canonical oscillation models has been used to describe the embodiment of meter perception in neurons, called neural resonance (Large, 2008; Large & Snyder, 2009). Essentially, neural resonance claims that metrical levels will be perceived at the frequencies at which neurons oscillate, whether or not that frequency was actually present in a stimulus (applicable, for example, in musical passages where syncopation can induce the sense of a pulse). In

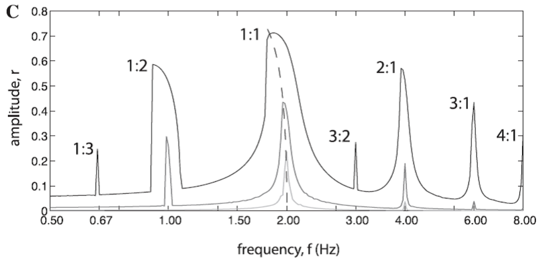



*Figure 2.3: The higher-order nonlinear resonance of a bank of oscillators, stimulated at 2 Hz frequency (at three different amplitudes). The canonical oscillators can model neural oscillation. Note the high peaks at the 2:1 harmonic of the stimulus frequency. Figure adapted from Large (2008).*

models describing neural resonance, the resonance at non-stimulated frequencies can easily occur at simple ratios to the stimulated frequency (such as 2:1 or 1:2; Figure 2.3).

Incidentally, the canonical oscillatory model from Large and Kolen (1994) is akin to the formal model describing the attentional oscillator of the DAT (e.g., Large & Jones, 1999; Large & Palmer, 2002). Thus several aspects of entrainment proposed here overlap: A similar oscillatory model describes both neural resonance that embodies meter perception (e.g., Large, 2008) and attention harnessed by metrical structure (e.g., Large & Palmer, 2002). Moreover, since the low-frequency delta-theta-gamma coupling that describes attention to rhythm or speech (Giraud & Poeppel, 2012; Schroeder & Lakatos, 2009) is applicable to the DAT's attentional oscillator (Henry & Hermann, 2014), this mechanism may plausibly describe meter perception. The current approach assumes combined neural and cognitive aspects of entrainment to meter, presented in the next section.

## 2.3 Neurocognitive entrainment in the current approach

As has been shown in the above sections, entrainment in the brain can be both cognitive and neural. Similar models can be used to describe the entrainment of attention to an external source (Large & Jones, 1999) and embodied perception of meter (Large, 2008; Large & Kolen, 1994), and moreover coupled neural oscillatory activity can describe entrained attention (observed by Henry & Hermann, 2014), entrainment to rhythm (Schroeder & Lakatos, 2009) and speech perception (Giraud & Poeppel, 2012). Neural assemblies entrained to one another in cognitive processing can be influenced by coupling with external rhythmic sources, imposing order in the communication and resulting in enhanced cognition (Singer, 2013; Thaut et al., 2005). The possibility for these various proposed mechanisms to overlap speaks for a broad definition of 'neurocognitive entrainment' that on the one hand mutually includes entrained attention, meter perception, and a corresponding neural oscillatory mechanism of coupled activity in delta, theta and gamma frequency, and on the other hand implicates increased neural synchrony that aids cognition via efficient neural processing. Biologically speaking there is a blurred distinction between what is neural and what is cognitive, because cognition relies on neural activity in the first place. In some places throughout this dissertation, cognitive vs. neural entrainment may be emphasized more strongly (e.g., Study 2 measures neural resonance, and Studies 1 and 3 are more concerned with cognitive consequences of entrainment).

In the current paradigm, neurocognitive entrainment is expected to result when participants are presented with predictable, temporal regularity in either music or language. Metrical structure can provide that temporal organization; hierarchically organized meter based on melodic contour or natural prosody (stressed vs. unstressed syllables) can entrain auditory neurons and attention, resulting in perception of meter, attentional peaks at metrically salient points, and more synchronous communication in the neural network, all of which should increase with increasing temporal predictability in the stimuli.

Now that neurocognitive entrainment has been defined, the metrical structures to which entrainment should occur shall be presented. Organizational principles of meter in music and language, as well as electrophysiological indices of meter perception, are presented in the next sections.

## 2.4 Entrainment to meter

A composite account of neurocognitive entrainment to meter has just been described. Here, theoretical principles that guide composition of meter are introduced for music and language domains, along with supporting EEG evidence (and some magnetoencephalography, MEG) of both general perception and neurocognitive entrainment to meter. The principles presented here aim to show common theoretical ground where hierarchical meter can be comparable across domains, as well as offer empirical indices that suggest comparable cross-domain entrainment to meter.

### 2.4.1 Music

This section presents theoretical and empirical evidence related to meter perception in music. Western Tonal meter is conventionally hierarchical and regular in composition, and the Generative Theory of Tonal Music (GTTM, Lerdahl & Jackendoff, 1983) presents metrical composition in relation to the multiple levels that a listener perceives. Early negative (Fitzroy & Sanders, 2015) and late positive (e.g., Brochard, Abecasis, Potter, Ragot, & Drake, 2003) ERPs as well as time-frequency MEG data (Snyder & Large, 2005) show that meter impacts the focus of attention to stimuli, specifically that enhanced event detection occurs in locations perceived to be metrically strong compared to weak. Frequency EEG data show that the brain represents frequencies of both acoustically present and imagined metrical levels (e.g., Nozaradan, Peretz, Missal, & Mouraux, 2011).

### Theory

In music, rhythm is the local pattern or group of relative duration of beats, whereas meter is the abstract organization of those beats as the music unfolds in time (London, 2012). In other words, meter is responsible for the 'feel' of a beat; for example, a march has a clear ONE-two-ONE-two pulse, and a waltz ONE-two-three-ONE-two-three. Meter in music emerges with a combination of melodic and temporal elements such as contour change or dynamic accents (Hannon, Snyder, Eerola, & Krumhansl, 2004). The organization of meter is hierarchical in nature, typically one of the lower levels being the tactus or salient constant pulse of a piece. The perception of tactus often coincides with the basic beat assigned by the time signature, although not always, as a fast tempo can shift perception of the tactus further up the metrical hierarchy, or a slow tempo shift perception down the hierarchy to a subdivision of the basic beat ("Meter", 1986).

**Meter classification** In convention dating back to the Baroque period (Wright & Simms, 2006), meter can be commonly classified as duple, triple, or quadruple, beats occurring respectively in groups of two, three or four. (In the above example, the march is duple and the waltz triple.) Meter is considered simple if the beat groups occur once per bar or measure, however if measures are subdivided and the beat groups occur more than once per measure, the meter is considered compound. If the meter changes frequently it is considered mixed or irregular; irregular meters were scarce until 20th century compositions. The time signature, a fraction located at the onset of a piece (or meter change within a piece), indicates the meter. The denominator tells which note receives the basic beat unit (a 4 assigning the quarter note the beat, an 8 assigning the eighth note the beat, and so on), and the numerator indicates the number of beats that occur in a repetitive pattern per measure. Thus, for example 6/8 meter is compound duple (two groups of three beats per measure), and 3/4 meter is simple triple ("Meter", 1986).

**The Generative Theory of Tonal Music (GTTM)** In their 1983 GTTM, Lerdahl and Jackendoff include meter as one of several elements that accounts for a listener's perception of musical structure (other elements are introduced in Chapter 3 Sections 3.1.1 and 3.2.1). According to the metrical portion of the GTTM, a beat present higher in the metrical hierarchy must also be present lower in the hierarchy. The smallest subset of beats (for example, played by eighth-notes) has a naturally emerging stronger beat that occurs with equal number of weak beats between, which can in turn be grouped at a next level (quarter-notes),

*Figure 2.4: Metrical structure: the dots represent the beat at which respective notes are played, stronger beats occur also at the next higher (shown lower) level in the hierarchy. Figure from Lerdahl and Jackendoff (1983).*

and so on. The hierarchical organization of beats is depicted in Figure 2.4. From the smallest beat perceived to the largest coherent pattern of the piece is about five or six metrical 'levels'.

### EEG

Physiological evidence supports human perception of the abstract metrical principles just outlined. Early negative ERPs demonstrate perception of converging metrical levels (Ladinig, Honing, Háden, & Winkler, 2009) in general meter perception. Entrainment to meter is evident in multiple paradigms, demonstrated by early negative (Fitzroy & Sanders, 2015) and late positive (e.g., Brochard et al., 2003) ERPs, as well as induced and evoked gamma activity (shown with MEG, Snyder & Large, 2005) and EEG steady-state evoked potentionals (SSE-EPs, e.g., Nozaradan et al., 2011).

**General Meter Perception** Early negative ERP effects demonstrate the perception of meter. One paradigm (Ladinig et al., 2009) presented (non-musician) participants with various computer-generated drum patterns (6.7 Hz presentation rate). Standard patterns (arranged in a meter) were presented for 90% of trials and deviant drum patterns (metrically crucial beats omitted from the standard, creating a strong or weak syncopation of the meter) comprised the remaining trials. Participants either watched a silent film (passive condition) or detected randomly inserted intensity changes that were unrelated to meter (unattended condition). The EEG component elicited by the omitted tone in deviants was a mismatch negativity (MMN). (The MMN, found either with EEG or MEG, typically appears around 100–200 ms after critical stimulus onset, and indicates pre-attentive detection of a deviance from a standard pattern; see Näätänen & Alho, 2007). Results showed that in both listening conditions, the MMN in response to strongly syncopated meter was larger than in response to weakly syncopated meter, indicating that participants had formed stronger representations of metrically salient elements of the drum patterns, thus attesting to the general perceptual salience of converging levels in metrical hierarchy.

**Neurocognitive entrainment to musical meter** The following studies show physiological evidence of neurocognitive entrainment to music meter, in time and frequency domains. ERP studies show evidence of temporally modulated attention, indexed by N1 (Fitzroy & Sanders, 2015) and P3 (e.g., Brochard et al., 2003) effects. Time-frequency gamma activity shows temporal expectation of metrical events (Snyder & Large, 2005), and SS-EPs show frequency peaks consistent with neural resonance at perceived metrical frequencies (e.g., Nozaradan et al., 2011).

**ERPs** A recent study (Fitzroy & Sanders, 2015) showed that musical meter promotes "hierarchical allocation of attention" as indexed by the **N1**, an attention-modulated anterior negative component peaking near 100 ms after auditory onsets (see Näätänen & Picton, 1987, for a review of the N1). Musicians and non-musicians were presented with triple- and quadruple-meter MIDI-generated melodies (notes presented at 2.2 Hz rate), with the task of judging whether an unrelated burst of white noise after each trial matched the meter. For both musicians and non-musicians, the N1 was larger in response to notes at metrically strong beats compared to those at metrically weak beats across meter types, indicating that attention was greater at points of convergence between metrical levels, and in turn that early auditory perceptual processing was influenced by the metrically modulated attention. Negligible difference between musician and non-musicians showed the impact of meter-modulated attention on early auditory perception to be unrelated to musical expertise (but see Chapter 4 of this dissertation). Moreover, the same melodies with a slower (1.6 Hz) presentation rate failed to elicit the same metric strength effect in the N1, indicating a lower limit of note presentation rate below which early auditory perceptual processing is not affected by (entrainment to) meter. The fact that the modulated N1 showed more attention to beats that were salient to multiple, nested metrical levels is directly in line with the DAT (e.g., Jones, 1976), indeed conclusively showing evidence of future-oriented attending.

Two studies have reported temporally modulated attention in response to hierarchical meter perception to be reflected by a P300 effect. Examining a documented behavioral effect labeled the 'tick-tock' phenomenon, wherein participants perceive isochronous identical sounds to have alternating accents (Bolton, 1894), Brochard et al. (2003) presented isochronous computer-generated pure tones to musician and non-musician participants in an ERP oddball paradigm. Standard pure tones (1.7 Hz presentation rate) were occasionally replaced by volume-deviant (quiet) tones, either on positions perceived as strong metrical positions (odd tones) or on perceived weak metrical positions (even tones). A posteriorly distributed late P300-like positive component (400 to 600ms post-stimulus onset) elicited by deviant tones was significantly larger when the deviant occurred in perceived strong compared to weak positions. The **P300**, a positive component typically peaking near 300 ms post-stimulus onset, is associated with novelty detection and its increased amplitude has been associated with attention to events and ease of event detection (Polich, 2007; Sussman, Ritter, & Vaughan, 1998). Thus in the tick-tock paradigm, the larger P300 amplitude suggests that perceived-strong-beat deviances were easier to detect, and in turn that cognitive attention was increased where the perceived metrical downbeat occurred. Moreover, the late positivity was earlier in musicians, suggesting that musicians potentially had "more efficient processing or stronger temporal expectancies." The impact of musical expertise on meter perception is revisited later in a later chapter (Section 4.1.1). The interpretation that attention modulated the P300 amplitude is direct support for future-oriented dynamic attending (but this interpretation was criticized for potential other influences on P300 modulation besides attention, see Fitzroy & Sanders, 2015).

Another study (Jongsma, Desain, & Honing, 2004) presented participants (musicians and non-musicians) with a series of tones organized in either a duple (2 Hz presentation rate) or triple (3 Hz presentation rate) metrical pattern, with the task to judge the temporal appropriateness of a probe tone that appeared after a silent interval (i.e., the appropriateness of the silent interval). Hypotheses were that if participants based their expectation of the probe on the silent interval before, a probabilistic expectation would be greatest at the repeated interval of the silent interval, but if participants based their expectation on hierarchical prediction, expectation would be based on probable locations of various metrical beats. A correlation coefficient, comprised of combined appropriateness judgments and amplitude of an evoked **P300**, showed that rhythmically trained participants indeed based their expectation of the probe beat on the present intervals. Thus the participants adjusted their expectations for metrical subdivisions of the isochronous sequences, consistent with an internal representation of meter that modulated temporal expectation, i.e. entrainment to the meter. Musical expertise influencing meter perception is revisited in Chapter 4.

**Frequency domain** Doelling and Poeppel (2015) recently showed that participants entrain to individual note onsets, or the lower metrical level, in human-performed musical excerpts. Stimuli were nine piano clips, whose filtered envelopes (the broadband sum of several narrow-band filters) were categorized by their note rate per second. Musician and non-musician participants listened to the stimuli, and judged whether trials contained a pitch distortion. The hypothesis was that individual-trial **phase synchronization** between MEG and piano clips could classify which brain response belonged to which stimuli, and that this would be evidence of entrainment to note rate by delta-theta neuronal oscillations. Indeed, it was found that the phase pattern per stimulus item was recognizable in the phase pattern of averaged brain responses to the respective stimulus. Classification strength increased with musicians, and with tempi above 1 note per second. Considering the task performance (pitch distortion detection) also correlated with the phase synchronization, findings were interpreted as showing a functional connection between "temporal predictions and content-based processing." This interpretation is consistent with the current approach: entrainment to information in the delta-theta range links to higher forms of cognitive processing (e.g., Giraud & Poeppel, 2012; Schroeder & Lakatos, 2009).

Snyder and Large (2005) showed evidence of entrainment to meter using auditory **evoked and induced gamma band activity** (20 to 60 Hz range). Following previous evidence that gamma activity was associated with auditory onset processing at various presentation rates (e.g., Snyder & Large, 2004) "linked to integrative sensory, motor, and cognitive functions" (e.g., Pantev, 1995), they hypothesized that induced gamma activity could represent metrical expectancy. While MEG was recorded, participants (with a wide range of musical expertise) were presented with duple-meter pure tone sequences, presented at approximately 154 beats per minute (bpm) or 2.56 Hz rate, with tones omitted on either metrically strong or metrically weak beats. Strong beats were acoustically louder and had a stronger attack velocity than weak beats. Using time-frequency analysis (frequency-power fluctuations over time), a control condition (no omitted tones) revealed that induced gamma activity was prominent exactly at or slightly before tone onsets, and evoked gamma surfaced shortly (ca. 50 ms) after tone onset and was further proportional in power to metrical strength. Experimental conditions showed that induced gamma activity was largely present regardless of tone omissions while evoked gamma activity was largely reduced after tone omissions, thus induced gamma may have represented the priming of auditory perceptual resources at times of temporal expectancy, and evoked gamma the 'primed' perception of auditory onsets. This allocation of auditory neural resources to locations of temporal expectancy provides direct evidence in support of neurocognitive entrainment, in line with the above-presented accounts (in Section 2.2.2) that attribute gamma activity, coupled to delta-theta activity, to external event decoding (Giraud & Poeppel, 2012; Schroeder & Lakatos, 2009).

Evidence from the frequency domain supports neural resonance theory (Large, 2008, see Section 2.2.3), that oscillatory behavior of neurons reflects meter perception. In one study (Nozaradan et al., 2011), participants (with wide a range of musical experience) heard a continuous isochronous 2.4 Hz beat for 33-second trials while their EEG was recorded, and were asked to imagine either 2/4 or 3/4 meter to organize the isochronous beats. An analysis using **steady-state evoked potentials (SS-EPs)**, which are frequency-amplitude peaks in the data that are significantly different from zero once the averaged amplitude of neighboring frequencies is subtracted, showed neural synchronization to both the real isochronous beat frequency and the imagined higher metrical levels, and moreover at subharmonic frequencies of the imagined meter. Thus the imagined metrical structure extracted from the isochronous beat and its subharmonic were represented among oscillating neurons with the same perceptual reality as the acoustically real beat, indicating that the perceived meter was embodied by the neural resonance.

Another study found evidence of entrainment in the frequency domain. Nozaradan, Peretz, and Mouraux (2012) presented participants (with a wide range of musical experience) with 33 seconds of metrical patterns (various groups of a 1.25 Hz beat) while recording EEG. Stimuli contained both metrically salient and metrically irrelevant frequencies, of varying acoustic prominence, and participants were tasked with detecting an unrelated temporary tempo change in the pattern. Examining the frequency domain using **SS-EPs**, it was found that the frequencies representing different levels in the stimuli's meter were present in the EEG response, with several remarkable characteristics. First, the frequencies most prominent in the EEG were consistent with metrically salient frequencies, independent of acoustic prominence. Second, the brain up-regulated metrically salient frequencies and down-regulated frequencies present but unrelated to the meter. Thus the EEG seemed to reflect a top-down cognitive process, consistent with dynamic attending, which perceived frequencies based on internal representation of beat and not simply based on the sound envelope. The metrically salient frequency peaks additionally replicated the previous results (Nozaradan et al., 2011) that neurons resonated at the frequencies perceived as metrical levels, indicating embodied meter perception.

Recently, Tierney and Kraus (2015) preserved natural timing in their stimuli, and showed that entrainment to musical meter was increased at strong-beat locations. A professional drummer tapped to the downbeat of popular musical excerpts, and this drumbeat track was presented together with the musical excerpts while listeners' EEG was recorded. Crucially, in one condition the drumbeat occurred at the strong beat locations, while in an alternate condition the track was displaced, occurring consistently at weak beat locations throughout the excerpt. **Spectral analysis** of the EEG revealed that participants entrained to the drumbeat in both conditions, indexed by a frequency peak at the rate of the drumbeat. A second frequency was found at the 2:1 harmonic of the drumbeat, only in the strong beat condition. Results were interpreted to mean that participants were entraining to both strong and weak beat locations, but that entrainment was greater at the strong beat, indicated by the additional presence of the 2:1 harmonic.

The EEG and MEG studies presented support the general perception of and entrainment to musical meter as it is theoretically described (Section 2.4.1), and also support the current theoretical stand that neurocognitive entrainment occurs to musical metrical structures. The latter three studies in particular supported neural resonance theory (Large, 2008, see Section 2.2.3), in that perceived metrical levels (whether acoustically present or not) manifested as prominent frequencies at which neurons were oscillating, both at stimulated frequencies and harmonics of those frequencies (Nozaradan et al., 2011, 2012; Tierney & Kraus, 2015).

This demonstrates the appropriateness of EEG frequency spectra to index musical meter perception; a frequency peak analysis is used to measure meter perception in Study 2 (Experiment 2b). Now that music meter has been theoretically and empirically surveyed, the theoretical background turns to language meter.

### 2.4.2 Language

In language, poetry, the metrical grid (e.g., Prince, 1983), and prosody (e.g., Nespor & Vogel, 1986) offer compositional principles that are hierarchical in nature. Early negative (e.g., Rothermich, Schmidt-Kassow, Schwartze, & Kotz, 2010) and late positive (e.g., Schmidt-Kassow & Kotz, 2009b) effects show that participants respectively detect and reanalyze disruptions to metrical structure established in sentences, indicating implicit and explicit perception of language meter whose hierarchical structure is comprised of stressed and unstressed syllables. While physiological evidence of entrainment in the language domain has not focused on meter, ERP and MEG frequency domain studies suggest that the brain synchronizes with or 'tracks' speech acoustics likely tied to the syllable (e.g., Astheimer & Sanders, 2009; Doelling, Arnal, Ghitza, & Poeppel, 2014). Taken together, the presented theory and physiological data compliment an approach to language meter that treats the syllable as the organizing unit to establish hierarchical relationships, and suggests that participants will perceive and entrain to the meter.

#### Theory

There are several veins of the study of meter in language. Poetry concerns itself with meter in an aesthetic sense, and within the study of phonology one finds sub-veins of metrical theory and prosodic phonology that formalize aspects of spontaneous productions of native speakers in a linguistic sense. The accounts of meter share the hierarchical organization of smaller phonological units into larger constituents, but differ in the nature of hierarchies; poetry is concerned with the form of metrical constituents defined on a per-line basis (Fabb, 2001), metrical theory is concerned with the ultimate strong-weak relationships among metrical units in speech (Prince, 1983), and prosody is concerned with meaningful perceptual units that are defined by phonological and phonemic aspects (Nespor & Vogel, 1986). The three approaches to meter are sketched below.

**Poetry** The study of meter in poetry reveals an arbitrariness of the smallest units, being often language-specific, which is governed by rules that for example say how smaller constituents are to be juxtaposed or where in a line they may or may not occur. Different concepts of a smallest unit include that of a mora (such as in Japanese Haiku), which weight

*Table 2.1: Meter in poetry. Example of dactylic tetrameter in German, with (dactylic tetrameter) gloss. Excerpt from Ermunterung by Johan Gaudenz von Salis Seewis, lines 1 – 4.*

| |
|---|
| Seht! wie die Tage sich sonnig verklären! |
| Blau ist der Himmel und grünend das Land. |
| Klag' ist ein Misston im Chore der Sphären! |
| Trägt denn die Schöpfung ein Trauergewand? |
| *Look how the days do transfigure with sun!* |
| *Blue are the heavens and verdant the lands.* |
| *Grumbling strikes discord in songs among angels!* |
| *Wears then creation a veil, does it mourn?* |

syllable vowels and consonants, or a discrete syllable (such as in French Alexandrine), or a foot, which is a group of smaller constituents (comprised for example of syllables in Germanic-language poetry or morae in Greek epics; Fabb, 2001). Examples of foot categories among others include trochee (a unit of two constituents, stressed then unstressed as in "sweater") and dactyl (a unit of three constituents, stressed followed by two unstressed as in "overalls"). Meter is conventionally labeled according to how many small units occur in a given line; dactylic tetrameter for example consists of four dactyl feet per line, shown in Table 2.1, one of many metric profiles.

**The metrical grid** Metrical theory typically accounts for meter in speech, with two main components extending from the surface structure of an utterance: first, a metrical tree is formed, which assigns relative prominence or strong-weak relations among constituents in an utterance, and second, a metrical grid exists into which the metrical trees may be mapped (Liberman, 1975; Liberman & Prince, 1977; Prince, 1983), shown in Figure 2.5. The metrical grid provides a temporal-rhythmic structure for the tree; this particular aspect of metrical grid, as well as others, borrows from musical concepts of meter (Figure 2.6). The theory as it is described here specifically pertains to the stress-timed rhythm class of languages that includes English, German, and Dutch (Abercrombie, 2006).[2]

Figure2.5 also demonstrates a key rule of the metrical grid that reassigns stress when two consecutive syllables (or lowest units) receive the strong stress: "Tennessee," when spoken alone, receives stress on the last syllable. However when it is placed next to a strong beat (as in "air" in this phrase), stress shifts to avoid a strong-strong stress clash, thus a relative strong-weak pattern is maintained, as is a within-constituent hierarchy with "air" at the highest level. By maintaining the recursive strong-weak syllable pattern, reassigning stress when necessary, the metrical grid forms a hierarchy in spoken speech with multiple nested levels. According to the current approach, the hierarchical relationship among stressed

[2]Metrical theory that describes multiple rhythm classes may be found in Hayes (1995).

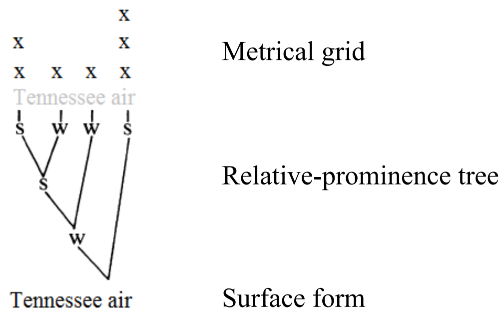



*Figure 2.5: From the surface utterance-constituent, metrical theory assigns relative strong (s) - weak (w) stress prominence with a tree, and subsequently represents each beat by a grid which represents a hierarchy of beats. Figure adapted from Liberman and Prince (1977).*

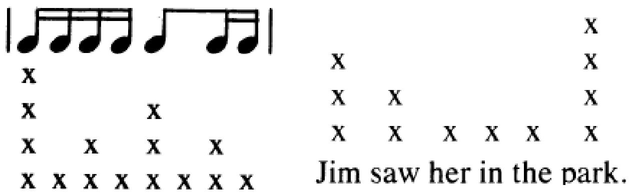



*Figure 2.6: The stress patterns in speech as described by the metrical grid are likened to meter in music, in that within a phrase there is a hierarchy of beats, such that the lowest is the smallest unit and weakest per se, and the higher up the hierarchy, the stronger the metrical value, with one highest-stressed beat per phrase, left or right aligned to the edge of that phrase. Figure from Prince (1983).*

syllables may serve as an attentional guide to the listener (Jones, 1976; Pitt & Samuel, 1990).

Poetic and grid elements of language meter may converge in describing formal speech, such as the oratory genre. Loosely defined, oratory is the public use of rhetoric for evocative or persuasive purposes (e.g. sermons, presidential addresses). Poetically defined meter and the metrical grid may all be invoked to describe, for example, excerpts from Dr. Martin Luther King's "I have a dream" 1963 speech. An examination of a recording of the speech and its transcription (*Martin Luther King, Jr.: I have a dream*, n.d.) reveals 'poetic' lines that transition from iambic to trochaic meter throughout the sentence, and use primarily stressed syllables to create a metrical hierarchy (combining or ignoring weak syllables, see Figure 2.7).

**Prosody** Prosodic constituents serve as the domain of application of specific sounds and phonological rules; prosodic phonology concerns itself with the principles that govern those sounds and rules. "Constituent" refers to a segment of the mental representation of speech, which is denoted by such speaker cues as pitch intonation, lengthening, and loudness (Ne-

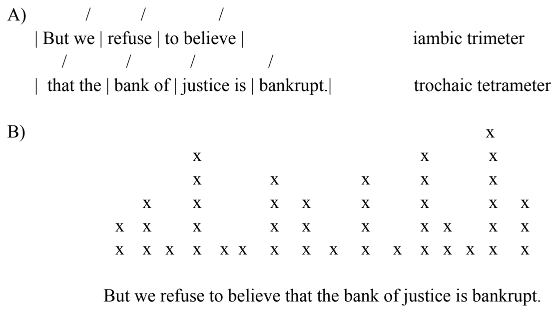



*Figure 2.7: An example metrical analysis from oratory. A) Using poetic terminology, the first half of the sentence could be described as an iambic trimester line, the second half a trochaic tetrameter line. Stressed syllables are accented. B) A possible interpretation of the metrical grid shows relative stress prominence.*

spor & Vogel, 1986). Theories of prosody generally concur in the existence of a prosodic hierarchy as in Figure 2.8, and a strict layer hypothesis asserts that each level consists of and dominates only the level directly below it (Nespor & Vogel, 1986; Selkirk, 1984). In other words, levels cannot recursively dominate constituents of their equal level as is possible in syntax, nor can domination skip levels. The current investigation is limited to the lowest two units in the prosodic hierarchy, the syllable and foot, which overlap with the poetic and metrical grid accounts of language meter. Prosodic constituents are dissociated from syntactic constituents, indeed comprised of conflicting internal rules in the governing of constituent hierarchies. In Section 3.2.2, a theory by Selkirk (2011) will be outlined, which asserts an interface of prosody — and by extension meter —- and syntax, despite these differences.

Now that theoretical aspects of meter in language have been outlined, EEG indices of meter perception, as well as MEG studies showing entrainment to syllable characteristics, will be presented.

### EEG

In stress-timed languages such as German, the current language of investigation, ERP paradigms investigating language meter perception can create a metrical structure using alternating strong and weak syllables in a series of words. Event-related potential responses to critical words that either sustain or break the metrical pattern can be interpreted as evidence of meter perception (e.g., Rothermich et al., 2010; Schmidt-Kassow & Kotz, 2009b). Entrainment to language meter per se is less well documented; studies are included here (including two MEG studies) that demonstrate entrainment to acoustic properties of the speech signal (Astheimer & Sanders, 2009; Doelling et al., 2014; Luo & Poeppel, 2007) that strongly point toward collective neurocognitive entrainment to syllable onsets. Though

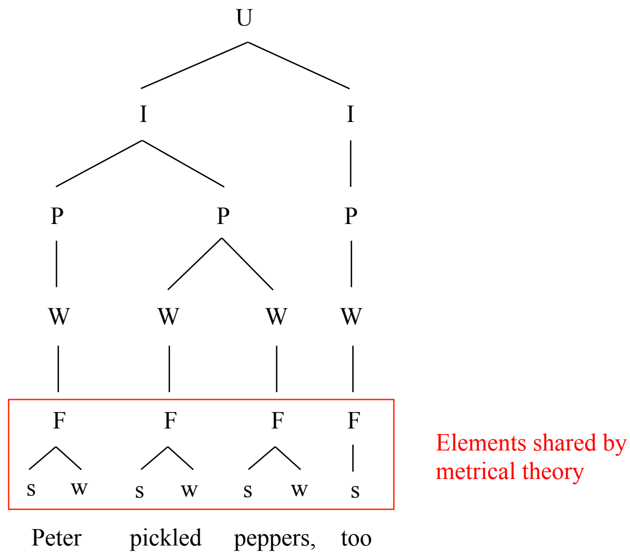



*Figure 2.8: Prosodic hierarchy. U = Utterance, I = Intonational Phrase, P = Phonological Phrase, W = Prosodic Word, F = Foot, s = Strong syllable, w = weak syllable. Category types commonly used in the prosodic hierarchy (Selkirk, 2011). Concepts Foot and Syllable overlap with metrical theory (and meter in poetry).*

falling short of showing entrainment to language meter, evidence of entrainment to the syllable supports the current approach in using the organization of stressed and unstressed syllables to establish metrical structure in language.

**General meter perception** The following studies show ERP responses to disruptions in language meter rules. Where stimuli created hierarchical relationships among stressed syllables (in the stress-timed languages, Dutch, English or German), early negativities (e.g., Rothermich et al., 2010) and late positivities (Schmidt-Kassow & Kotz, 2009b) reflect respectively detection of violated metrical rules and reanalysis thereof, thus supporting the prospect of hierarchical language meter perception among healthy participants.

**ERPs** Early negativities have been reported to represent the detection of violated linguistic meter. Rothermich, Schmidt-Kassow, Schwartze, and Kotz (2010) presented native German speakers with spoken Jabberwocky (syntactically preserved, semantically nonsense) sentences, which were composed in a repeating strong-weak metrical pattern and spoken by a trained German native speaker (e.g. all words two-syllable, stressed syllables indicated with single quotation: 'Tro-geumpf 'hät-te 'Knäpf-eint 'pei-le 'dö-gent 'na-pen 'sol-len / *Trogeumpf should have napped Knäpfeint peile dögent*). When the sentences contained a disruption in the strong-weak metrical pattern (e.g., 'Trog-eumpf 'hät-te 'Knäpf-eint 'pei-le 'dö-gent na'PEN 'sollen), an **early negativity**, broadly-distributed and peaking near 250 ms, was elicited in response to the metrically deviant pseudo words both when participants completed a metric and syntactic task. The effect was interpreted as representing domain-general detection of rule violation, the ‘rule’ in this case being the metrical structure extracted from the sentences.

Another ERP study in English shows a similar early negativity effect (Magne, Gordon, & Midha, 2010), more strongly attributable to hierarchical relationships in language meter, thus supporting the authors’ interpretation of the early negativity in Rothermich et al. (2010). Magne et al. (2010) visually presented healthy participants with word lists consisting of a context of several two-syllable words with either trochaic (first syllable stress) or iambic (second syllable stress) meter, and a final target word with either trochaic or iambic structure (2 x 2 design), with a semantic task unrelated to metrical structure. Event-related potentials time-locked to the target word onset showed an early fronto-central negativity effect (250–400 ms window) elicited when the target word mismatched with its metrical context, regardless of the particular metrical identity (trochaic or iambic) of the context or target words. Since this effect resembled the **N400**, a negativity conventionally associated with semantic processing (but also discourse processing, math, and other phenomena, reviewed by Kutas & Federmeier, 2011), authors interpreted the early negative effect to reflect “perturbed” lexico-semantic processing brought on by a mismatching metrical context and target.

Considering the early negativity found in Rothermich et al. (2010) and their interpretation that the negativity reflects detection of a violated metrical rule, the N400-like negativity found in Magne et al. (2010) may well also reflect the disruption of a perceived metrical rule, i.e. a syllable stress pattern. The relevant function of the early negativity in these two studies, in light of the current investigation, is that its task independence indicates implicit perception of hierarchical relationships in language meter.

A study in Dutch found both negative and positive effects in response to disruptions in metrical patterns established by two-syllable word lists (Böcker, Bastiaansen, Vroomen, Brunia, & Gelder, 1999). Context and target words were either trochaic or iambic, and target meter was either congruent or incongruent with the context meter. Participants either listened passively or actively judged congruity of the target word. A **negativity** named an N325, peaking near 325 ms with fronto-central distribution, was elicited in both tasks, and a late positivity, peaking near 816 ms with left posterior distribution, was elicited in the explicit meter task only. This study highlighted the role of the early negativity as indexing metrical awareness due to both its overall sensitivity to metrical stress (surfacing when participants heard iambic words, an infrequent lexical stress pattern in Dutch, regardless of their context) and enhancement in incongruent as opposed to congruent conditions as well as enhancement in a metrical task compared to passive listening. The **late positivity** was attributed to the task demand (recognition memory to assess congruity) and not to metrical processing, however; the presence of a late positivity in response to deviated metrical patterns persists in the literature (e.g., see below).

In two French studies (Magne et al., 2007; Marie et al., 2011), biphasic ERP effects were elicited in healthy participants by unusually lengthened penultimate syllables in sentences (a deviation, since in French, usually only the last syllable receives stress and lengthening; see Magne et al., 2007). Participants listened to sentences and performed metric and semantic tasks relating to the critical final (multisyllabic) word. In the first study with non-musician participants, an early negative effect, right-lateralized in the 250–450 ms time window, was interpreted as an **N400** that reflected difficult semantic processing of the metrically manipulated words. A **late positivity**, broadly distributed in the 500–800 ms time window, was interpreted as reflecting "integration difficulties resulting from the violation of the typical metric pattern" at the end of sentences. Moreover, the N400 was present in metric and semantic tasks, whereas the late positivity occurred in the meter task only. A second study conducted with musician-only participants replicated the task-independent early negative and task-dependent late positive findings (Marie et al., 2011; musician and non-musician effects are compared in Section 4.1.1). Thus the early negative effects likely further reflected automatic or implicit processing of the meter, whereas the late positivity was dependent on controlled or explicit processing of the meter.

A study conducted in German also found that violation of language meter elicited both an early negativity and a late positivity (Schmidt-Kassow & Kotz, 2009b). German (non-musician) native speakers listened to sentences composed with a repeating strong-weak syllable pattern, spoken by a trained German native speaker (e.g. 'Vera 'hätte 'Christoph 'gestern 'morgen 'duzen 'können. / *Vera could have addressed Christoph informally yesterday morning*), with either a metrical or syntactic task. A critical syllable misplaced the syllable weight in half of the sentences, creating a multi-faceted metrical structure violation, both in the sentence-long pattern, the local strong-weak relationships, and in the critical word itself (e.g. 'Vera 'hätte 'Christoph 'gestern 'morgen du'ZEN 'können.). After the onset of the critically violated syllable, an early negativity and a late positivity were elicited in both tasks. The **early negativity**, bilaterally distributed with latency approximately 250-450 ms in metric and syntactic tasks, was interpreted as reflecting a disrupted word segmentation that was hindered by unusual lexical stress assignment. The positivity, bilaterally distributed with a latency of 550-1150 in the metric task and 750 – 1050 ms latency in the syntactic task, was interpreted as being a **P600**. The P600 is broadly reported in the literature, in syntactic (e.g., Osterhout & Holcomb, 1992), semantic (reviewed in Kuperberg, 2007), and metrical contexts (e.g., Schmidt-Kassow & Kotz, 2009a), and can be attributed to a "general integration mechanism" that performs reanalysis when an element does not match its context (see Brouwer, Fitz, & Hoeks, 2012). The presence of the P600 was interpreted as representing reanalysis of violated language-meter rules, attention dependent since the meter task P600 had a significantly earlier latency. Although it is not clear in this case whether the P600 was in response to hierarchical or local word-level metrical violations, its elicitation is indication that participants extracted a metrical structure from the sentences and reanalyzed deviant metrical constituents.

A follow-up German study proved the latter point more clearly, showing that native speakers indeed form a mental representation of meter that informs expectations of future spoken events that occurs in the absence of word-level meter violations. Schmidt-Kassow and Kotz (2009a) presented German (non-musician) native speakers with the same sentences as in the above study, but replaced the metrically violated critical word with a correctly spoken critical word of a different syllable count. The critical word disrupted the global strong-weak pattern established by the sentence but kept metrical integrity of the critical word (e.g. 'Vera 'hätte 'Christoph 'gestern 'morgen ver'HÖHnen 'können. / *Vera could have mocked Christoph yesterday morning*). When participants were tasked with judging the regularity of sentence meter, a posteriorly distributed **P600** was elicited by the critical word. The similarity of the elicited ERP to the violated meter in the previous study (Schmidt-Kassow & Kotz, 2009b) supported the interpretation that the effect represented reanalysis after a disruption of hierarchical meter established on the scale of the whole sentence, and not just reanalysis of a lexical stress violation.

These combined ERP studies show that the brain perceives hierarchical metrical relationships among stressed syllables, and reacts to deviations in the pattern whether or not attention is directed to the meter. Though none of these paradigms established physiological evidence of entrainment to the language meter, their results are compatible with the current account of neurocognitive entrainment to hierarchical metrical structures regardless of domain: meter perception is synonymous with entrainment to various levels according to neural resonance theory (Large, 2008, see Section 2.2.3), and the above paradigms arguably demonstrated at least implicit meter perception among participants. The next section describes more directly entrainment-attributable physiological responses to speech, albeit in the absence of language stimuli that establishes hierarchical metrical relationships.

**Entrainment to speech** Considering the role of the syllable in the current approach to language meter (e.g. its role in the prosodic foot, Nespor & Vogel, 1986; see Section 2.4.2), several studies focusing on entrainment to the syllable, investigating N1 modulation in the time domain (Astheimer & Sanders, 2009) and phase and power information in the frequency domain (e.g., Buiatti, Peña, & Dehaene-Lambertz, 2009; Luo & Poeppel, 2007), show the potential for language meter to influence entrainment. Though these paradigms do not attempt to establish hierarchical relationships in language meter, their demonstration of cognitive or neural entrainment to syllable information validates the current treatment of syllables as basic metrical units to investigate entrainment to language meter.

**ERPs** An ERP study in English demonstrated that attention is increased near the onset of words (Astheimer & Sanders, 2009), consistent with cognitive entrainment (e.g., Jones, 1976) occurring at these locations. Astheimer and Sanders (2009) presented participants (no report of musical expertise) with excerpts of a long speech, with linguistic probes (the syllable 'ba') inserted on or near word onsets (experimental condition) and at other random spoken positions within the excerpts (control condition). An auditory evoked **N1** in response to the experimental probes was greater compared to control probes, indicating that auditory attention was greater near word onsets compared to other locations in speech. The modulated N1 at word onsets in Astheimer and Sanders (2009) is consistent with cognitive entrainment to word onsets, since the N1 is associated with attention (e.g., Näätänen & Picton, 1987). As argued above in Fitzroy and Sanders (2015), increased N1 at metrically strong notes in a melody indicated entrainment to meter. An audio recording of the speech revealed that the speaker was very rhythmic in his spoken delivery of the material (Edward Abbey's reading of Freedom and Wilderness, 1987; Fundi, 2008); thus here the increased N1 at word onsets could be further evidence for entrainment to rhythmic speech. Relevant to the current approach, the study espouses word onsets as an appropriate marker for metrical boundaries in language, as is the case in current sentence stimuli (see Section 5.3.2 for a full stimuli description). Moreover, although target locations in stressed vs. unstressed syllables was not reported in Astheimer and Sanders (2009)[3], a distinct possibility was that attention was directed to strong syllable onsets, which is consistent with the attentional bounce hypothesis (Pitt & Samuel, 1990).

**Frequency domain** One MEG study suggested internally driven oscillations in the theta range to entrain to syllable information in intelligible speech (Luo & Poeppel, 2007). Healthy participants listened to nine sentence conditions (three sentences with three intelligibility levels) with the task to judge whether two successive sentences were the same or not (the first sentence always intelligible, the second sentence always unintelligible). The hypothesis was that individual-trial **phase synchronization** between MEG and speech stimuli could classify which brain response belonged to which stimuli in proportion to intelligibility. Indeed, it was found that the phase pattern per stimulus item was recognizable in the theta-range phase pattern of averaged brain responses to the respective stimulus, clas-

[3]In an example utterance, three target locations were in partially or fully stressed syllables (e.g., not the lowest level in the metrical grid): in bold, "The landscape of the Colorado **pla**teau lies still beyond the reach of **rea**sonable words or unreasonable **re**presentation."

sification strength increasing with intelligibility of stimuli (while power in the theta range played no role in classifying trials). Considering the average syllable rate across multiple languages to occur in the theta range (e.g., Greenberg, Carvey, Hitchcock, & Chang, 2003), results were interpreted to reflect entrainment to acoustic syllable information. Interestingly, while theta tracking became less robust with low intelligible sentences, a follow-up experiment indicated that theta phase tracking remained intact when the stimuli were compressed (0.5 compression ratio) but still intelligible. This indicated an intrinsic role of theta oscillations (localized in this study to the auditory cortex) in tracking intelligible speech, “resetting according to speech dynamics” when the task required discrimination of sentence-level utterances, regardless of the particular rate of speech.

Another MEG study further suggested that entrainment to syllable information is based on the sharpness of temporal acoustic cues in the amplitude envelope (Doelling et al., 2014). Doelling and colleagues (2014) presented healthy participants with spoken word lists, whose amplitude envelopes were processed to create varying degrees of intelligibility and peak sharpness in the delta-theta range (corresponding to syllable rate information). MEG was recorded while participants listened to word lists, and then the brain data was compared to the amplitude envelope of the stimuli using **envelope-tracking**, which measured coherence of the respective stimuli and brain phase information based on a Phase-Locking Value. The envelope tracking analysis showed that regardless of stimulus intelligibility, stimuli with sharp envelope peaks in the delta/theta range had most coherence with brain data. (When the low-intelligible, sharp temporal peaks condition was removed, phase coherence was proportional to intelligibility, as was originally hypothesized). Thus the brain entrained most to the signal when the sharp envelope peaks provided temporal acoustic cues about the syllable; the brain was even ‘tricked’ into entraining to this information when it was meaningless, in a low-intelligibility condition. The acoustic cues to which the brain entrained were interpreted as cues for segmentation of speech; data is also in line with the assumption in the current approach that the brain will entrain to regular, temporally predictable cues provided by metrical structure.

A very recent MEG study showed that linguistic boundaries can be hierarchically entrained (Ding, Melloni, Zhang, Tian, & Poeppel, 2016). English and Chinese participant groups listened to 4 Hz isochronous artificial syllable streams in either Chinese (Chinese and English groups) or English (English group only). Entrainment to the speech was evaluated using the **frequency tagging** approach, which stimulates specific frequencies and then examines a power spectrum for peaks at these ‘tagged’ frequencies. Importantly, there was no coarticulation among the syllables, thus any linguistic boundaries above the syllable had to be based on internally generated structures. As hypothesized, both participant groups entrained to the individual syllables (monosyllable words), indexed by a 4 Hz peak in the

MEG power spectrum. When participants listened to their own language, additional peaks representing syntactic boundaries (phrases and sentences) also surfaced, consistent with entrainment of these internally generated syntactic structures. Crucially, when English participants heard Chinese materials, only the 4 Hz syllable peak was present. Thus entrainment to hierarchically arranged linguistic boundaries was tied to comprehension-related processes. What Ding et al. (2016) showed may apply to meter perception as well as syntax perception; in the current paradigm, hierarchical metrical, as opposed to syntactic, boundaries are investigated. If neurons resonate at entrained linguistic boundaries that are syntactic, they may also do so with metrical boundaries in accordance with neural resonance theory (e.g., Large, 2008).

Finally, entrainment to syllables was investigated in an EEG study using the **frequency tagging approach** (Buiatti et al., 2009). Participants listened to artificial syllables presented isochronously at ca. 4.2 Hz, with the task to discover imaginary words in an artificial language. Pseudowords were determined by transitional probability in an artificial grammar. Syllable streams were presented in four conditions (2 x 2 design), as randomly concatenated (with or without 25 ms pauses inserted every three syllables) or as pseudowords (with or without 25 ms pauses inserted every three syllables). Participants entrained to the syllable rate of random syllables with and without pauses inserted, indexed by a peak in the EEG spectrum near 4.2 Hz. Participants did not, however, entrain to the syllable rate when exposed to pseudoword streams—the 4.2 Hz peak diminished in these conditions. While no peaks were significant in the no-pause pseudoword condition, a peak near 1.4 Hz surfaced in the inserted-pauses pseudoword condition. In other words, when pauses existed between learned pseudowords, participants entrained to the learned word boundary (every three syllables) and no longer to the individual syllable. Although this paradigm was not treating syllables as metrical units, the hierarchical arrangement of syllables into groups of three has certain parallels with hierarchically structured meter. Thus from the current perspective, Buiatti et al. (2009) may have been showing a qualitative difference in the way that participants entrain lower as to upper metrical levels in speech. This possibility will be re-explored in Study 2.

The above studies show the potential of the syllable to serve as the metrical constituent that is temporally predictable enough to entrain to; while the above studies did not consider stressed and unstressed syllables in their paradigms, the current paradigm aims to exploit hierarchical relationships among strong syllables to promote entrainment commensurate with metrical regularity.

### 2.4.3 Cross-domain comparison

Domain-specific aspects of metrical theory and corresponding physiological responses to metrically salient features have just been outlined. This section emphasizes particular theoretical and acoustic parameters of music and language stimuli—hierarchy and presentation modality—that need to be controlled in order to optimally compare neurocognitive entrainment to meter in both domains.

#### Theory

In metrical theory, music meter is classified in Western Tonal tradition according to the pattern of strong and weak beats within units called measures, and across a piece those units combine to create multiple nested hierarchical levels perceived by the listener. Language meter may be described in terms of poetry and phonology; poetry offers a classification of meter based on strong-weak weighting of smallest-units (syllables, in German) across a line, the metrical grid offers nested hierarchical organization of strong-weak constituents when words are spoken, and prosody assigns stress to syllables in metrical 'feet.' Taken together, these theoretical principles offer a converging compositional structure within which music and language meter may be comparable: melodies and sentences of equal length, which have notes and syllables with respectively identical strong-weak patterns. Each domain can thus be constructed to create the same perceptual hierarchy of nested strong-weak constituents, which in turn has potential to similarly encourage entrainment, provided the temporal predictability in the signals remains comparable.

#### Stimulus timing parameters

Considering the role of temporal predictability in entrainment, an important parameter to consider in a cross-domain comparison, once meter is controlled, is that timing is comparable in music and language. One striking difference between the auditory music and language meter paradigms presented in this chapter is the difference in temporal precision offered by stimuli in the two domains. Musical note onset in experimental settings is typically timed with computer-precision, whereas human-spoken sentences inherently offer natural temporal variation in syllable onset rate (though more regularity is perhaps perceived than exists, see Darwin & Donovan, 1980). Two recent notable exceptions are Tierney and Kraus (2015) and Doelling and Poeppel (2015), who investigated entrainment to human-timed music. These recent exceptions aside, differences in meter perception across the domains cannot be entirely dissociated from differences in temporal perception; for example, even within domain, two identical metrical structures with different tempo could have different outcomes regarding participants' entrainment to the signal (Fitzroy & Sanders, 2015). A way to avoid different execution of timing is to, first of all, use the same medium to perform both the music and speech, e.g. either both be human-performed or both be artificially generated with computer software, and moreover to ensure comparable presentation rate of notes and syllables. Human-spoken speech compared to computer-generated music within the same paradigm would confound the temporal comparability of the meter in respective domains.

Furthermore, to create optimal circumstances for neurocognitive entrainment to meter, temporal parameters in cross-domain stimuli should establish salient metrical levels in the delta and theta frequency ranges. Such a construction would take advantage of the delta-theta range oscillations that seem to drive perception to speech (Doelling et al., 2014; Luo & Poeppel, 2007) and rhythm (Schroeder & Lakatos, 2009) and that have been proposed to underlie the attentional oscillator described by the DAT (Henry & Hermann, 2014).

## 2.5 Open issues

The working hypothesis proposed here is that a neurocognitive mechanism behind entrainment to metrical structures could be responsible for entrainment in both music and language domains, providing that the metrical structure in respective domains is comparable. Two particular studies suggest, however, evidence to the contrary: fMRI and EEG findings suggest separate temporal processing (Abrams et al., 2011) or metrical grouping (Schmidt-Kassow, Rothermich, Schwartze, & Kotz, 2011) in music and language domains. These findings would indicate, despite similarity of the acoustic signals, that entrainment to the metrical structures in respective domains would not be comparable. The studies and their implications for the current investigation are addressed in the next paragraphs.

In Abrams et al. (2011), similar resources were found to underlie music and speech perception, but the particular pattern in which the resources responded to similar ‘scrambling’ of musical and spoken excerpts were different across the domains. Non-musician participants listened to excerpts of symphonies and speeches while images of their brains were scanned. Half of the excerpts were temporally scrambled (while controlling for the amplitude envelope between the two conditions), such that ca. 350 ms chunks were randomly reordered (which preserved small note groups and words but disrupted the syntax of the excerpts); a separate rating experiment assured emotional, arousal, and familiarity aspects among all stimuli. Focusing on hypothesized regions of interest in frontal and temporal lobes, univariate analysis showed similar activations for the music and speech excerpts (in combined original and scrambled conditions), but multivariate analysis (original-minus-scrambled conditions) showed functional differences in the pattern of activation among voxels in these regions. These findings were interpreted such that ’dynamic temporal processing' in music and language use the same functional networks, but that the implementation of the network has different nuances in the two domains.

Nuanced implementation of neural resources that monitor temporal structure is precisely what this paradigm intends to measure with EEG. Thus despite the exclusion (in the current investigative scope) of anatomical regions associated with meter and syntax, what Abrams et al. (2011) suggest has large implications for the current study: that the nature of neurocognitive entrainment cannot be not generalized across domains, and by extension, that EEG responses to melodies and sentences (with comparable temporal qualities) would be qualitatively different. However, Abrams et al. (2011) had the disadvantage that stimuli did not account in any way for syntactic differences across domains; "temporal processing" was synonymous with scrambled syntax, and no comparison of perceptual differences in syntax between and speech and symphonic excerpts was included in the study description. Thus the reported differences in temporal processing may be attributable to asymmetrical detection or reanalysis mechanisms to violated syntax across domains. The current investigation may reveal whether qualitatively different temporal processing occurs when other parameters including syntactic parameters are comparable across music and language domains.

A second study proves challenging for the current cross-domain, neurocognitive entrainment approach: Schmidt-Kassow et al. (2011) suggested that discrete neural mechanisms are responsible for the perceptual grouping of beats in linguistic and non-linguistic domains. In an ERP study, the same sentences as Schmidt-Kassow and Kotz (2009b, see above Section 2.4.2) were used, which created a metrical structure with alternating strong-weak syllables and then violated the structure. Where German native speakers elicited a biphasic response to the metrical violation, French native speakers who were late proficient learners of German did not elicit any ERP responses to the metrical violation. (A syntactic effect indicated that the sentences were understood, that syntactic processing was intact.) Thus the French speakers did not register their second-language (L2) metrical violation as a violation.

A second experiment in Schmidt-Kassow et al. (2011) was conducted in order to see if the different metrical processing between native French and German speakers was due to a general grouping preference, or to a language-specific meter perception. Replicating an experiment from Brochard et al. (2003, see Section 2.4.2), a series of isochronous tones was presented with occasional deviant quiet tones. Both French and German speakers perceived strong-weak grouping (as opposed to weak-strong or no grouping) of the isochronous beats, reflected by an increased P300 effect elicited by deviant tones on odd number locations than by even number locations. Thus the absence of ERPs responding to the metrical violation in their L2 was not a result of domain-general differences in grouping preference (such as always assigning stress to the last in a group of beats), and this was interpreted such that different neural mechanisms underlie linguistic and nonlinguistic meter processing.

The latter study (Schmidt-Kassow et al., 2011) offers convincing evidence that the perception of meter is qualitatively different in music and language. Still, one could argue that the two experiments were not testing cross-domain resources comparably enough to claim the same perceptual phenomenon across domains. The linguistic experiment effectively measured reanalysis of a violated metrical rule (in German speakers) or not (in French speakers), compared with a music paradigm that showed entrainment to tones (no reanalysis of a violated rule). As far as the current investigation is concerned, entrainment is a distinct phenomenon from rule-violation reanalysis; that the French-participant ERPs did not indicate the observance of an L2 metrical violation does not necessarily imply that there was no entrainment to the their L2 language meter. Entrainment to the language meter could have in fact occurred in the French participants, despite a local breakdown in the metrical pattern; in line with this possibility, Snyder and Large (2005) showed that evoked gamma activity, present for strong metrical beats in a tone sequence, was present in some participants despite the absence of an actual acoustic tone.

The two studies presented here suggest that cross-domain entrainment to meter may not be comparable, since temporal processing on the one hand and metrical grouping on the other were reported to be different in cross-domain, within-participant paradigms. Counter-arguments were offered here to negate the interpretation of these findings, namely, that temporal processing may have been confounded with dissimilar syntax conditions, and that a key distinction between metrical rule reanalysis and entrainment distances the impact of that study from the current investigation. The next section presents some specific parameters of the EEG paradigm to be used in Study 2, where music- and language-meter perception will be measured with EEG frequency analysis.

## 2.6 Current approach

In this chapter, entrainment to meter has been successfully shown in music, and entrainment to the syllable has been successfully shown in language; what the current investigation offers is a cross-domain comparison that structures music and language meter analogously, to see if entrainment to the meter is subsequently analogous. This paradigm uses melodies and sentences that treat eighth notes and syllables in a one-to-one relationship, comprising metrical organization of beats that offers a comparable strong-weak beat pattern across domains. Considering the importance of timing in the signal, i.e., that temporal predictability should come from the meter and not from dissimilar sound generation in music and speech, current stimuli are both human-performed (by a trained pianist and a trained speaker). Moreover, a controlled tempo for eighth note and syllable presentation rate, approximately 5 Hz, capitalizes on perceptual benefits reported for acoustic stimuli in the theta range (e.g., McAuley et al., 2006; Schroeder & Lakatos, 2009). A full description of the paradigm may be found in Section 5.3.

A primary research question in this dissertation is whether participants entrain similarly to meter across domains, when the meter and temporal fluctuation of presentation is consistent. The empirical physiological evidence included in this chapter points to a likely fruitful approach to cross-domain meter perception that uses an EEG frequency-domain method. Based on neural resonance theory (e.g., Large, 2008), the frequencies at which neurons resonate (in response to metrical information) embody the perception of metrical levels, and resonance at those frequencies is achieved through neural entrainment to the meter. Thus identification of significant frequency peaks in EEG frequency-power spectra, if they are peaks that can be traced to beat periods in metrical levels (or harmonics thereof), may simultaneously represent both neural entrainment and perception of meter (see Large, 2008). An EEG frequency analysis of cross-domain meter perception is performed in Study 2 (Experiment 2b).

A secondary, accompanying research question is whether the auditory signals of human-performed melodies and sentences are objectively comparable in their (frequency) representation of metrical levels, when meter and timing have similar parameters. Drastically different acoustic signals might call into question the theoretical cross-domain comparability of meter and complicate the interpretation of brain data from respective domains. Therefore, before EEG is analyzed, Study 2 (Experiment 2a) performs a frequency analysis of stimuli materials .

## 2.7 Summary

This chapter has introduced a working definition of neurocognitive entrainment that compiles several mutually compatible cognitive and neural theories. Principles of metrical organization in music and language meter have been outlined, and EEG and MEG findings have built a case (including some counter-evidence) that neurocognitive entrainment to meter can occur comparably in both domains. Finally, the metrical aspect of the current paradigm was shown to provide appropriately similar stimulus conditions in the two domains in order to test the hypothesis that neurocognitive entrainment to meter is similar across domains.

The overarching aim of this dissertation is to show whether entrainment to meter can impact syntax processing comparably across domains, and what role individual differences play in the process. Now that the metrical portion of the paradigm has been presented, the next chapter turns to the syntax portion of the investigation.

# Chapter 3

# How entrainment to meter can facilitate syntactic comprehension

The previous chapter was concerned with how neurocognitive entrainment to meter could be comparable in music and language domains. This chapter brings the comparison one step further, exploring how neurocognitive entrainment could impact syntactic comprehension similarly across the two domains. First, the syntactic organizing principles of music and language, along with relevant ERP indices, will be outlined. Then, a theoretical and neurocognitive account of how entrainment to meter can influence syntactic comprehension will be presented.

## 3.1 Syntax

Before addressing how entrainment can impact syntactic comprehension, the theoretical and empirical aspects of syntax must be understood. In the next sections, generative theories from music (Lerdahl & Jackendoff, 1983) and language (Chomsky, 1957) will be outlined, which were respective influential milestones in their fields[1], and which offer common principles across the two domains that may be used for empirical comparison. Event-related potential studies that report indices of relevant syntax perception will also be presented. Particularly syntactic manipulations that create correct but more complex or less expected structures are of interest; error-detection mechanisms add a component of executive control that is irrelevant and potentially confounding.

[1]For historical perspectives and comparison of various musical and linguistic theories, consult Christensen (2006) for music (see also the recent Rohrmeier, 2011) and Graffi (2001) for language.

### 3.1.1 Music

Western tonal music encompasses most of the genres heard in popular media and classical concert arena in largely European or European-descended societies. Here, some fundamental elements of music syntax are introduced, highlighting relative key ambiguity, as well as harmonic aspects of the GTTM (metrical aspects shown in the previous chapter, Section 2.4.1; Lerdahl & Jackendoff, 1983). Relevant ERP paradigms are highlighted, which demonstrate key effects known in the literature to represent violation detection (e.g., the ERAN; Koelsch, Gunter, Friederici, & Schroeger, 2000) and harmonic integration (e.g., the N500 and P600; N500, Koelsch et al., 2000; Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013; P600, Featherstone, Morrison, Waterman, & MacGregor, 2013).

#### Theory

**Basic principles** The Western Tonal system of tonality organizes individual notes on the scale in a hierarchy that can lead to rule-governed large works. From a bottom-up perspective, the music notes themselves are arranged in a scale, within a unit called an 'octave.' An octave is a 2:1 frequency ratio of a given note; the frequency is halved or doubled but still retains the perceived identity of that given note, only lower or higher. A chromatic scale is a conventional twelve-note division of the octave separated by pitches one-half interval apart, or a 'second' interval.

The diatonic scale (Figure 3.1), used in stimuli for the current paradigm, employs eight of the twelve notes, organized around the base pitch, calling it a 'tonic.' Each note may also be termed 'degree' and referred to as its degree away from the tonic. Certain nomenclature corresponds to the scale degrees: the tonic (first degree of the scale), supertonic (second), mediant (third), subdominant (fourth), dominant (fifth), submediant (sixth), and leading tone or subtonic (seventh). The eighth degree doubles the tonic in frequency, and the following octave repeats the scale. This set of pitches organized in a scale around the tonic is called a 'key.' Tonal keys may be major or minor. A natural minor scale (of the same tonic) is created by lowering the third, sixth, and seventh scale degrees by one semi-tone of the major scale. A harmonic minor scale lowers only the third scale degree.

**Relative keys** Two relative keys share the same key signature and notes, but are organized around different tonic identities. A relative minor is built upon the submediant or sixth scale degree of its corresponding major key (Harrison, 1995). The common key signature and notes shared by the relative keys C-major and A-minor are shown in Figure 3.1.

Musical works may treat relative keys ambiguously or modulate between them, often using harmonic context to establish or change the listener's sense of tonality. The usage of


*Figure 3.1: The diatonic scale in Western Tonal harmony. (a) C (natural) minor. The third, sixth and seventh scale degrees are lowered from C major (b) C major. The pitches proceed as whole-, whole-, one-half-, whole-, whole-, whole-, one-half-step. (c) A (natural) minor. All of the notes are the same as in C major, organized around A as the tonic.*

secondary dominants – the fifth scale degree of a note other than the tonic, which 'resolves' to said scale degree – is one common element used by composers in the common practice period (Piston, 1941). The relationship among keys, their relative keys, and dominants are depicted in Figure 3.2.

**The GTTM** The **GTTM**, whose metrical aspects were described in Section 3.1.1, also describes the harmonic structure and final stage of music perception to which an experienced listener arrives (Lerdahl & Jackendoff, 1983): rules of well-formedness establish the hierarchy of musical components; transformational rules permit modifications of actual music examples to fit the model; and preference rules do the major analysis as to which structures in the music most influence the listener. The concept of 'reduction' is an important part of this theory: the listener attempts to organize all the pitch-events of a piece into a single coherent structure, such that they are heard in a strict hierarchy of relative importance, the less-important events having a specified relationship to their surrounding more-important events.

The GTTM addresses harmony in terms of hierarchies of tension and relaxation. Prolongational reduction describes the intuited tension and relaxation in a musical piece. The most stable point of relaxation is the end of a group, the next most stable is the beginning, and events in between are considered points of tension, all operating from the top-most level of the hierarchy (the whole opus) to its smallest group. This is intuitive when considering the general sense of closure accompanying the end of a piece of music, the climax of a piece at its point of greatest tension in compared with its beginning, and individual phrases perceived as having the same holographic form. The relationship between levels is recursive, as any given node (after the top) can be considered an elaboration of the tension relation-

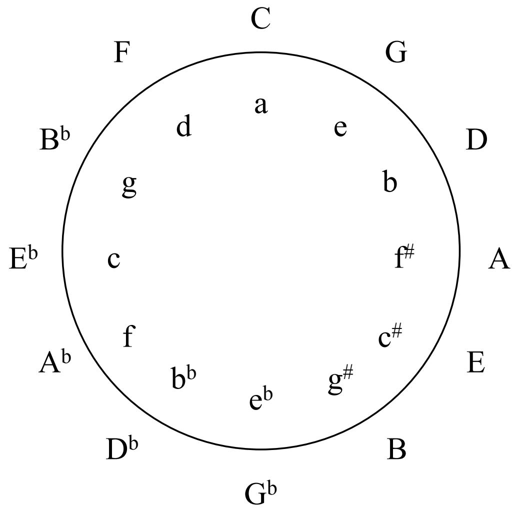



*Figure 3.2: The circle of fifths, first published by Johann David Heinichen in 'Musical Circle' in 1711 (Wright & Simms, 2006). Major keys are in the outer ring in capital letters, their relative minor keys in the inner ring in lowercase letters. The proximity of keys in the circle reflects their actual harmonic proximity, e.g., C major is harmonically close to G major and A minor, and harmonically distant from G flat major.*

ship of its superordinate node. Figure 3.3 shows the prolongational reduction of a simple musical phrase.

Another element of musical syntax is grouping structure. A group is a set of musical elements within a piece that naturally occur together in units, perceived as themes, phrases, periods, theme-groups, and sections. Groups are perceived hierarchically, such that two smaller groups can comprise a larger group and so on, so that themes are perceived as sets of sub-themes and themselves part of larger sections (below Figure 3.5 depicts this grouping structure hierarchy, in Section 3.2.1). The theory has a strict classification of groups, such that there is never an overlap of group boundaries within the hierarchy; aside from the topmost and bottommost groups, each intermediate group can be dominant over its own smaller component sub-groups or subordinate to a larger group in the other direction; this results in a recursive structure wherein each subgroup could be considered an elaboration of its dominating group (next highest node of the hierarchy).

The syntactic nature of the model lies in the hierarchical organization of its subcomponents. Relaxation and tension patterns are organized into a prolongational reduction, and grouping structure, meter, and time-span reduction (introduced below in Section 3.2.1.) combine to represent the musical surface as it is finally perceived by the experienced human listener. All of these structures are derived by the well-formedness rules, and then start to undergo a more personal interpretation for the listener via application of transformational rules and consideration of underlying properties.

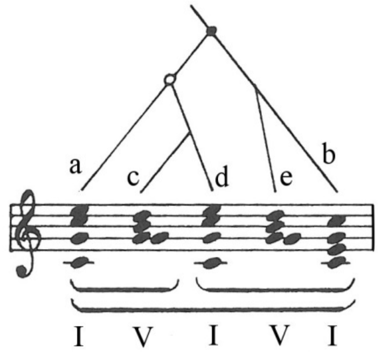



*Figure 3.3: This prolongational reduction shows group a – b as starting and ending on the tonic (in C major), with points of tension and relaxation within the group (at subgroup levels). The tonic d is considered here a repetition, and therefore embellishment, of tonic a. Dominant c is a point of tension relaxing to tonic d (a cadence). The dominant e is the point of tension in the final cadence of the phrase, relaxing into b, which is the greatest point of relaxation in this phrase (Figure from Lerdahl & Jackendoff, 1983).*

### ERPs

The human brain is sensitive to disruptions in the harmonic relationships outlined above, e.g. violations of the harmonic tension-resolution relationships (Lerdahl & Jackendoff, 1983). This section provides a rough chronological overview of music-syntax ERP paradigms.[2] Considering the salience of musical expertise in the perception of harmony, musical expertise of the participants is mentioned here; musical expertise as it impacts syntactic comprehension is addressed primarily in Chapter 4. Most relevant to the current paradigm, this section describes a study that shows ERP evidence for preference of certain syntactic structures over other alternatives, even when there is no explicit harmonic violation (e.g., Koelsch et al., 2013).

**Responses to harmonic error** Research in music syntax perception began using EEG methods in the late eighties. The early paradigms often used familiar melodies, and focused on violations that were overt and multifaceted, such as tonality, contour, and timbre being violated by the same event (e.g., Besson & Macar, 1987). The P600 was among the first ERP effects to be reported for music syntax perception: One early ERP study addressed the perception of harmonic incongruence as a function of tonal distances from harmonic context, familiarity, and musical expertise (Besson & Faïta, 1995), and found the various manipulations to elicit both late negative (N500) and late positive effects (P600), the latter reflecting musical expectancy. In the first ERP experiment, musicians and non-musicians were presented with familiar and unfamiliar melodies, each of which ended in either in-key

[2]Music-syntax ERPs from cross-domain paradigms are addressed in Section 3.1.3

or out-of-key deviants. The task was to categorize the computer-generated piano melodies (of "various" tempos) according to varying degrees of harmonic congruence. A negativity (200 ms to 600 ms latency, anterior distribution) was elicited in response to deviant endings primarily in unfamiliar melodies, and the effect was driven by musicians. A late positivity (300 ms to ca. 1200 ms latency, right anterior distribution) was elicited by deviant endings as well, which was larger in amplitude and earlier for: musicians compared to non-musicians, familiar as opposed to unfamiliar melodies, and out-of-key as opposed with-in key deviants. A second experiment with a less explicit task (general questions over blocks of stimuli) revealed the late positivity to have a smaller amplitude compared to the explicit task, and the musician-non-musician difference to disappear. The late negativity was not interpreted beyond being distinct from an MMN or N2. The late positivity was identified as a late positive component (LPC), which is in light of later studies the same as a **P600**. The P600 was interpreted as reflecting musical expectancy, being larger when (violated) expectations were stronger and when explicit attention was given.

Subsequent work focused on unfamiliar chord sequences (as opposed to melodies) with the aim to isolate components related to music syntactic knowledge (e.g., Koelsch et al., 2000). The ensuing paradigms typically found two negative effects, the ERAN and the N500 (but rarely found a P600). In a pioneering paradigm, Koelsch et al. (2000) reported two novel negativities to reflect respectively violated expectancy of sound and integration of sounds into their previous context: an early right anterior negativity (named **ERAN**) peaking near 150 ms, and a late, bilateral frontal negativity peaking around 500 ms (named **N500**). The effects were elicited in a study whose four experiments gaged the ERPs' sensitivity to sequence position of a violation, degree of harmonic dissonance, task-relatedness and detection in 1.7 Hz MIDI chord progressions. The amplitude of both effects, in response to harmonic violation (Neapolitan chords that contained out-of-key notes), increased when expectancy for harmonic chords was increased (e.g., at the end of sequences, when a return to a tonic was harmonically expected, and when the probability of the Neapolitan chord was 25% as opposed to 50%). The two negative effects likely reflected discrete cognitive processes of detection and integration of harmonic violations; the ERAN and N500 amplitudes were reduced in independently varying degrees according to task-relevance and probability. The N500 was also present in response to in-key chord onsets, with amplitude reduced as in-key chords formed a sequence (e.g., it was larger after the first in-chord than after the fourth in-key chord). This amplitude reduction as sequences progressed was interpreted to mean that the N500 reflected harmonic integration, in other words the process that integrates chords into their preceding context; the better the fit, the lower the N500 amplitude. This was similar to the behavior of the N400 in language studies, where word onsets in a sentence demonstrated progressively reduced amplitude as more words fit their preceding semantic context (e.g., Van Petten & Kutas, 1990). Participants were all non-musicians, thus these effects reflected enculturated awareness of harmonic relationships as opposed to education in musical syntax theory. The elicitation of the ERAN and N500 has been replicated in many subsequent studies, noted by Featherstone et al. (2013) to commonly have used non-musicians and chord progressions that ended with low-probability deviant chords (e.g., Koelsch et al., 2000; Steinbeis & Koelsch, 2008).

Recently, an ERP paradigm successfully undertook the goal of reconciling the seemingly mutually exclusive P600 and N500 effects found in earlier music syntax paradigms (Featherstone et al., 2013). Featherstone et al. (2013) presented musicians and non-musicians with musical passages, MIDI tones presented at several rates balanced across stimuli (1.2 Hz, 1.7 Hz, or 2.2 Hz), with the task to rate each item according to various properties including arousal, aesthetic pleasantness, and oddness. Stimuli consisted of three conditions: congruous, incongruous-resolved, and incongruous-unresolved. Incongruous-resolved started one key, made a tonal detour to a new key (either closely or distantly related, balanced across stimuli), and returned to its original key. Incongruous-unresolved made a harmonic detour to a new key and remained in that key. Control congruous stimuli consisted of the same key as the harmonic detours of the incongruous conditions. Triggers were placed at the point of harmonic detour (identical across the three conditions) and the point of resolution (either return to the original key or not), thus testing neurophysiological responses to both congruence and harmonic resolution. Event-related potentials time-locked to the first chord of the key modulation (compared to the congruent condition) showed a late posterior positivity that was significant between 500 and 700 ms in both groups but driven by the musicians, and interpreted as a **P600**. Event-related potentials time-locked to the final chord of the sequence (which either resolved to the original key or not) showed a similar P600 in musicians, but revealed a near-significant right-hemisphere late negativity (570 ms peak latency) in non-musicians. Despite distributional differences (attributed to differences in rhythmic makeup of the melodies) compared with earlier reports, this negativity was interpreted as an **N500**. This study suggested the N500 to reflect a lack of harmonic resolution (as opposed to harmonic integration), and that the P600 was rather a marker of harmonic integration. The differences in ERP effects elicited between musicians and non-musicians was attributed to a local, analytical integration of an incongruous element into its context by musicians, as opposed to a holistic listening strategy by non-musicians that is sensitive to global tension or resolution. The study shifted the terminology from syntax vs. semantics to congruity and resolution, which can incorporate language semantic and syntactic paradigm manipulations and seems to resolve the differences in reported ERP effects in music studies.

Although the studies included above are not an exhaustive sample, the most consistent emerging pattern seems to be that an ERAN is elicited regardless of task, while the N500 and P600 are present in multiple, differing circumstances of task-relatedness, musicianship, and position of a violation in a sequence. Thus the latter two effects may be consistently associated with cognitive function that performs an ongoing structural integration of syntactic elements (or the degree to which such element is expected, given the context), although the nuances of the circumstances are not transparent across the literature.

**Responses to hierarchical error** One recent ERP study (Koelsch et al., 2013) shifted the paradigm from local syntax violations to include hierarchical relationships (i.e., deepening the comparability to language syntactic comprehension). Thus Koelsch et al. (2013) extended the repertoire of the **ERAN** and **N500** to include hierarchical processing of music harmony over long distances, demonstrating listeners' preference for harmonic closure within a musical piece, a preference moreover common across musicians and non-musicians. Participants (musicians and non-musicians) listened to musical passages with the unrelated task to detect chords played by deviant instruments. Stimuli were excerpts (or altered excerpts) from Bach chorales, presented with MIDI audio files at 1.7 Hz rate. The chorales consisted of either an original ABA or altered CBA musical structure (thus the first section established the context, the last section either returned to original tonality or introduced a new tonality). Event-related potentials time-locked to the final chords of the passages showed effects highly reminiscent of the ERAN and N500, with no statistical difference in musicians compared to non-musicians. Thus even in the absence of a local harmonic violation, the expectancy of harmonic events included an implicit awareness of hierarchical tension-relaxation relationships, regardless of musicianship; a return to the beginning key of a piece provided the greatest sense of harmonic resolution.

Importantly for the current paradigm, which aims to test the impact of meter on the processing of correct, non-preferred structures, the latter study showed that preferences for non-violated Western syntax are strong enough to elicit event-related responses in the absence of local violations; the hierarchical irregularity in Koelsch et al. (2013) is a violation of preference rules such as those described by the GTTM, since the tension created in the middle of the passage should optimally be 'relaxed' to the original key.

**Sensory vs. cognitive music perception** One major issue in music syntax research is whether ERP effects originate from sensory or cognitive processes, since experimental conditions that manipulate harmonic relationships can easily also create sensory dissonance. For example, the ERAN has on the one hand been suggested to be elicited by sensory processes (Poulin-Charronnat, Bigand, & Koelsch, 2006), and on the other hand cognitive or purely syntactic processes (Koelsch, Jentschke, Sammler, & Mietchen, 2007). The P600, the focus-ERP in the current investigation, is typically designated as cognitive or syntactic in nature (e.g., Featherstone et al., 2013; Patel, Gibson, Ratner, Besson, & Holcomb, 1998, see Section 3.1.3), but has also been claimed to be sensory-based (Regnault et al., 2001). Indeed, previous paradigms reporting the music syntax P600 often included out-of-key violations, which created a tonal distance between critical events and their contexts that could trigger sensory responses. The current paradigm shall attempt to avoid sensory dissonance in the harmonic conditions, since all conditions will use within-key critical notes (relying on interval relationships as opposed to notes to determine syntactic belongingness to a key), and critical notes will always lie within the range of melodic contour previously established in the melody (i.e., the target note is never the highest or lowest point in the melody).

The sensory vs. cognitive argument is a large and multifaceted discussion, and is not fully addressed here beyond the care taken to avoid violation-type dissonance with present music stimuli. Two recent studies show that sensory-based processes could alternatively explain many claims for cognitive/syntactic processes reported by previous behavioral and EEG studies (Bigand, Delbé, Poulin-Charronnat, Leman, & Tillmann, 2014; Collins, Tillmann, Barrett, Delbé, & Janata, 2014). Bigand et al. (2014) reviewed behavioral and EEG datasets that claimed cognitive-based music syntactic comprehension, and found a model based on the role of psychoacoustics and auditory working memory able to describe almost all previously reported results. Such a strict perspective on the role of sensory processing in music essentially undermines empirical investigations based on traditional accounts of music syntax (like the principles described in Section 3.1.1), since the physical characteristics of the sound can always predict neurocognitive data independently of whatever syntactic relationships exist among those sounds. Collins et al. (2014) applied a similar model that considered psychoacoustics, auditory working memory and musical 'closure,' and found their model to account for a large set of reaction time data previously attributed to 'cognitive' music syntactic comprehension. Collins et al. (2014) suggested music cognition to therefore encompass a continuum of multiple sensory and cognitive processing stages as opposed to discrete (and opposing) sensory vs. cognitive accounts. Fully addressing this debate may require innovations in empirical design as well as combined behavioral, physiological and neuroimaging datasets to determine the presence of purely cognitive processes, and is outside the scope of this dissertation.

These above sections have shown theoretical and ERP evidence for syntactic comprehension in music, as part of the cross-domain comparison of syntactic comprehension—and whether it is impacted by entrainment to meter—in this dissertation. Before music is compared to language (in Section 3.1.3), the next sections necessarily introduce theoretical and empirical aspects of language syntactic comprehension.

### 3.1.2 Language

Syntax in language is a vast discipline, therefore only principles relative to this dissertation will be mentioned here: Chomskian theoretical concepts of syntax (Chomsky, 1957) are briefly outlined and the particular case of ambiguous relative clause attachment in German, central to Study 3, is introduced. The empirical portion focuses on ERP effects elicited by syntactic violations, as well as highlighting effects elicited by syntactic manipulations that are not errors but rather local ambiguities.

#### Theory

**Basic principles** Syntax is the study of the principles and processes by which sentences are constructed in particular languages, whose goal it is to account for the sentences by a grammar (Chomsky, 1957). A Chomskian account of grammar includes that language is common to all humans, and the mechanics of all languages operate with a common set of core principles (universal grammar). Core linguistic principles include phonemics, morphology, and phrase structure (example phrase structure is in Figure 3.4). The inarguably varied manifestations of different languages come not from fundamental differences among the languages themselves per se, but rather from differences in external parameters, which make slight alterations from a language's core and cause the surface differences apparent to the casual listener. Grammars have a core sequence of rules, which govern phrase structure and the way that morphemes reach spoken word form, as well as transformational rules, which apply structural changes to phrases (such as rearranging word surface word order depending on whether a sentence is a statement or question in English). Sentences may be recursive, or have nested phrases within phrases, which permits an infinite possible combination of sentences with a finite set of words.

Several decades after Chomsky proposed his concept of grammar, other linguists proposed a potential way that language-particular parameters may interact in order for surface form to be reached. This was called Optimality Theory (Prince & Smolensky, 2004). Optimality Theory creates a dichotomy between language-universal core components—much as those outlined by Chomsky—and language-particular parameters or constraints. The constraints operate in a language-specific hierarchy, and 'choose' the correct surface form from the possible set of outputs created in the universal core grammar. Thus principles of universal grammar are theoretically applicable across languages, and of particular relevance here, extend from English, the language in which they were conceived, to German, the language used in the current set of experiments.

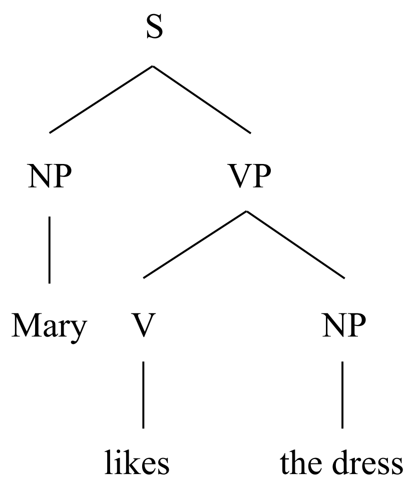



*Figure 3.4: Phrase structure in an English sentence. S = Sentence, N = noun, V = verb, P = phrase. Figure adapted from Baker (2001).*

**Relative clause attachment** A relative clause is a sentential clause that serves to modify some antecedent (Example 3.1.1). The relative clause may contain a relative pronoun ("who" referring to the antecedent), which may occur in any case (nominative, accusative, and so on).

**Example 3.1.1.** The boss who went on strike crossed the street.
"who" is clearly referring to "the boss."

**Example 3.1.2.** The boss of the bus drivers who went on strike crossed the street.
"who" could refer either to "the boss" or "the bus drivers."

In the second example (3.1.2), the relative clause is ambiguous, attached either to the boss or the bus drivers. Usually context would play a role for a listener or reader; the intonation of a speaker or implicit prosody of a reader would somehow cue a listener as to who crossed the street (Augurzky, 2006; Fodor, 2002). One particular feature of German syntax is that the functional word *die*, ("who") when appearing as a pronoun, has the same surface form whether it refers to a feminine, singular noun or any-gender plural noun (in nominative and accusative cases). An example with *die* is in the below example (3.1.3).

**Example 3.1.3.** Die Chefin der Busfahrer, die streikte/streikten, ueberquerte die Strasse.
*The boss (fem.) of the bus drivers who went (fem. sing. or - plural) on strike crossed the street.*
The relative pronoun *die* is ambiguous in its appearance at the beginning of the relative clause, and is subsequently disambiguated by the conjugation of the verb *streiken*.

This particular usage of *die* (Example 3.1.3) allows a sentence to only be locally ambiguous; that is, a reader or listener may assign the relative pronoun based on the conjugation of the verb "to strike." Such a setup is ideal in psycholinguistic research, as, for example, the ERP-response at the verb onset may indicate which verb conjugation the participant was expecting, thus allowing an online measure of which attachment was preferred.

In the above construction (Example 3.1.3), if the relative clause refers to the first noun or subject in the sentence ("boss"), the attachment is considered 'high,' and if the relative clause refers to the second noun of genitive object/possessor ("bus drivers"), the attachment is considered 'low.' Previous research in auditory German sentence processing showed no consistent preference for ambiguous high- vs. low relative clause attachment (Augurzky, 2006), although the preferred reading of such sentences have been reported to always be 'high' (Hemforth, Konieczny, & Scheepers, 2000). It is important to note here that whether native German speakers have high- or low relative clause attachment preference is outside the scope of the current study; the interest here is to rather investigate the neural process that occurs when participants encounter their non-preferred attachment (and whatever impact entrainment may have on that process). Attachment preference for current German sentence stimuli is therefore determined in a separate experiment (Study 3, Experiment 3b), and assumed to be limited to the current stimuli, without attempting a systematic explanation for individuals' generalized attachment preference. Further reading on German relative clause attachment preference may be found in Augurzky (2006), Hemforth et al. (2000), and current state of the art for language ambiguity processing may be found in Sanz, Laka, and Tanenhaus (2013).

### ERPs

The above section has just described some basic principles of language syntax, highlighting relative clause attachment, which is the syntactic manipulation used by the current paradigm. This section accordingly focuses on ERP responses to non-error, structural syntactic manipulations (but briefly surveys ERP responses to violated syntactic structure). Important here is that when sentences have more than one possible syntactic structure, speakers commonly demonstrate a preference for one resolution over another. The most relevant ERP response is the P600, a late positive effect elicited when native speakers encounter a complex structure that conflicts with their expectations (e.g., Osterhout & Holcomb, 1992).

The three most common ERP effects related to syntax (violation) perception were demonstrated by one early study (Neville, Nicol, Barss, Forster, & Garrett, 1991). Neville et al. (1991) tested participants' response to syntactic violations (such as "...Max's of proof the theorum" and "Ted's about films America") where word order disrupted the phrase structure. Participants were instructed to judge grammaticality while sentences were visually presented, word-by-word at a rate of 2 Hz. Time-locked to the words that first marked the deviant structure, two negative and one positive effect were elicited: an early left anterior negativity (**ELAN**) peaking near 125 ms; a later, left anterior negativity (**LAN**) with a latency between 300 and 500 ms (to be distinguished from the N400, which is posteriorly distributed, e.g. Kutas and Hillyard 1980); a late positivity with centro-parietal distribution with and latency between 500 ms and 700 ms (**P600**). These three effects were not interpreted by this study beyond that they represented processing of ungrammatical, violated syntactic (as opposed to semantic) structure (a possible interpretation of the late positivity was that it "patched up" the sentence after a violation); however, these three effects proved to consist of the primary repertoire of syntax ERPs that represent syntactic violations. Incorporating ERP and anatomical imaging findings, neurocognitive models of auditory language comprehension by Friederici (1995, 2002) suggested the ELAN to represent the detection of syntactic violations (such as word category), the LAN to represent the integration of morphosyntactic information, and the P600 to represent reanalysis and repair (but see Steinhauer & Drury, 2012, for a criticism of both this model and the validity of reported early negative ERP effects).

Event-related potential studies in the decades since have investigated the similarities and differences of ERP responses to violated structure as opposed to grammatically correct, non-preferred experimental manipulations of syntactic structure. The next few paragraphs are a representative sample of these studies, which almost exclusively report similarities and differences in the P600.

In an influential study, Osterhout and Holcomb (1992) presented English participants with sentences with syntactic violations or ambiguities over two experiments, with the task to judge whether sentences were acceptable according to semantic content and grammaticality. Participants read sentences in a word-by-word presentation (ca. 1.5 Hz presentation rate). In the first experiment, sentences contained a temporary syntactic ambiguity known to cause readers to stumble in their initial comprehension and reanalyze to achieve understanding, called a *gardenpath* effect. For example, "The broker persuaded to sell the stock" is a reduced relative clause describing the broker (in expanded form, "the broker who was persuaded to sell ..."), but an initial reading by native English speakers would assume an act of the broker to persuade someone until seeing the word "to". Event-related potentials time-locked to words that induced syntactic reanalysis, e.g. "to," elicited an anterior late positivity peaking near 600 ms, named a **P600**. A second experiment replicated this effect while showing that structural anomaly elicited by the word "was" in "The broker hoped to sell the stock was sent to jail," (treating the first part of the sentence as a reduced relative clause, when it was not) elicited a similar P600 with slightly more posterior distribution. Thus this late positivity consistently reflected structural reanalysis in contexts where the participant had a different preferred first reading of a local ambiguity or gardenpath (with centro-anterior distribution), compared to repairing violated phrase structure (with centro-posterior distribution).

The different distributions of the P600 were later explicitly investigated, in an effort to see whether syntactic preference is subserved by a similar cognitive function to repair violated rules. Friederici, Hahne, and Saddy (2002), replicated Osterhout and Holcomb (1992) findings that the more anterior distribution follows syntactic complexity whereas the more posterior distribution follows syntactic violation. Friederici and colleagues (2002) presented German participants with correct and incorrect sentences with major and minor complexity (2 x 2 design; word-by-word visual presentation at 2 Hz rate) with the task to judge sentence correctness. Event-related potentials were time-locked to critical words establishing complexity or correctness. Compared to simple sentences (one NP), complex sentences (two NP) elicited a centro-anterior positivity with 800 ms to 1100 ms latency (**P600**). In response to incorrect sentences, a biphasic response (reminiscent of the early Neville et al., 1991 report) was elicited, a negativity with central distribution and 350 ms to 450 ms latency, followed by a positivity with 500 ms to 1100 ms latency with centro-posterior distribution (P600). Where time windows overlapped, the two P600 distributions were significantly different. Thus it was argued that the complex (but correct) structure was activating a distinct neural network from the violated structure, reflected in the different distribution. Important to the current investigation was the confirmation that syntactic complexity that is grammatically correct elicited a centro-anterior P600, which reflected syntactic reanalysis (as opposed to repair after structure is violated).

However, other studies have found no difference in P600 distribution when comparing syntactic ambiguity resolution to violated syntax, indicating that the cognitive processes underlying the two effects are rather similar. Kaan and Swaab (2003) manipulated complexity (one versus two noun phrases in a sentence), structural preference (e.g. “I cut the cake beside the pizzas that were brought by Jill” is preferred compared to “I cut the cakes beside the pizza that were brought by Jill”), and grammaticality (e.g., noun-verb number agreement) in native English speakers. A **P600** effect was elicited by critical words that determined the condition type (complex, non-preferred, and ungrammatical): the non-preferred and ungrammatical were posteriorly distributed, while the complex condition was anteriorly distributed. Thus both reanalysis of (grammatically correct) non-preferred structure and repair of ungrammatical structure seemed to be represented by similar underlying mechanisms, reflected in the posterior P600.

A more recent study sought to determine whether the **P600** elicited by various syntactic manipulations relies on different neural generators (i.e. whether scalp distribution was different; Gouvea, Phillips, Kazanina, & Poeppel, 2010). Participants answered unrelated comprehension questions after sentences were visually presented word-by-word (2 Hz presentation rate). Event-related potentials were time-locked to critical words determining condition type. Gardenpath sentences elicited a posterior P600 (300 to 1300 ms latency), as did ungrammatical sentences (posterior P600 with 500 to 1300 ms latency). Despite similar scalp distributions, the earlier latency in the gardenpath condition was interpreted such that reanalysis information was available to the participants sooner in the gardenpath condition (since preceding words already contained the alternate correct interpretation). The ungrammatical sentence had no cues to an alternative interpretation, thus the 'repair' started after the possibility for alternative interpretations was exhausted. This study adopted an interpretation of syntactic reanalysis and repair such that a common core of syntactic processes attend different structural manipulations, and are expressed differently (e.g. latency) based on the complexity and information available in the encountered structure.

One recent hypothesis for the language comprehension underlying ERP effects assumes syntactic integration to be maintained by a general integration mechanism that is responsible for any number of 'integration' functions, including semantic integration[3] (Brouwer et al., 2012). According to this account, which frames language comprehension as "mental representation of what is being communicated" (MRC hypothesis), a 'retrieval' mechanism (represented by an N400) retrieves lexical features from long-term memory and a subsequent 'integration' mechanism (represented by a P600) integrates lexical information into an ongoing mental representation (the paper left incorporation of other syntax effects, such as the LAN, for future work). Thus in light of the studies mentioned here, this account would favor the perspective of Gouvea et al. (2010) that the P600 reflects a set of resources and not a highly specific syntactic function. This account also offers a palatable explanation for domain-general function of the P600; integration of new information into a current mental representation could also describe syntactic integration in music. Accordingly, the syntactic process underlying the P600 in the current investigation will be referred to as 'integration' as opposed to 'reanalysis' or 'repair.'

To sum up this section, event-related potential paradigms investigating grammatically correct structural complexity, such as ambiguous relative clause attachments or gardenpath effects, have elicited a P600, with either anterior-or posterior-central distribution (e.g., Kaan & Swaab, 2003; Osterhout & Holcomb, 1992). This empirical evidence on the one hand suggests highly nuanced neural processes to attend specific aspects of structural integration (e.g., Friederici et al., 2002), and on the other hand there are counterarguments that claim a more monolithic integration mechanism to subserve syntactic reanalysis (e.g., Brouwer et al., 2012; Gouvea et al., 2010); the current investigation remains neutral to the two views of

[3]This model offers a common ground to controversy surrounding the syntactic specificity of the P600 (see Kuperberg, 2007, for a review).

language syntax integration, and cannot dissociate between them since only non-preferred and not violated language syntax will be employed. Nonetheless, a P600 is hypothesized for the current language syntax manipulation (Study 3, Experiment 3c) and its distribution may support one account over the other. The next section compares syntactic comprehension across music and language domains.

### 3.1.3 Cross-domain comparison

As presented in this chapter so far, music and language both have syntactic organization principles, or a 'grammar' governing its structural unfolding in time (Chomsky, 1957; Lerdahl & Jackendoff, 1983), and ERP effects representing syntactic comprehension seem on the surface to have similarities, such as early negative detection-related effects and late negative and positive integration-related effects (Featherstone et al., 2013; Friederici, 1995; Koelsch et al., 2000). This section introduces a specific hypothesis for shared syntax resources in the two domains (Patel, 2003), and directly compares syntax-related ERP effects across the domains.

#### Theory

In an influential paper, Patel (2003) noted the comparability of distance-relationships among syntactic constituents in music and language syntactic theory (e.g., Gibson, 2000; Lerdahl, 2001), and—valuable to empirical comparison of cross-domain syntactic comprehension—that the more distant a syntactic element is from its preceding context, the more difficult to integrate. This theoretical observation was juxtaposed with converging anatomical and functional evidence that similar neural activation (e.g. Broca's Area; Friederici, Wang, Herrmann, Maess, & Oertel, 2000; Maess, Koelsch, Gunter, & Friederici, 2001) and similar ERP responses (e.g. Patel et al., 1998, see below section) seemed to underlie syntactic comprehension in both domains, as well as the acknowledgment of patient reports of impaired music (e.g., amusia) or language (e.g., aphasia) function while function in the other system was preserved.

**Shared Syntactic Integration Resource Hypothesis (SSIRH)** Thus the **Shared Syntactic Integration Resource Hypothesis (SSIRH)** proposed that overlapping neural resources access separately stored musical and linguistic syntactic representations, and that these resources are at play in the online integration of syntactic constituents in both domains. Since the SSIRH was conceived, much behavioral, electrophysiological, and anatomical evidence has surfaced to support the presence of overlapping resources underlying syntactic comprehension in music and language. (A challenge to the SSIRH, i.e. an alternative account that attributes evidence of shared syntactic resources to cognitive control, is presented in below Section 3.3; Slevc & Okada, 2015).

### ERPs

The previous sections dedicated to ERP paradigms showed separate music and language experiments to elicit electrical potentials in response to both explicit syntactic error (Koelsch et al., 2000; Neville et al., 1991) and to more subtle manipulations such as unresolved long-distance harmonic tension (Koelsch et al., 2013) or local ambiguities or gardenpath effects (e.g. Osterhout & Holcomb, 1992). The current section presents the few ERP paradigms that addressed music and language syntactic comprehension within-participants, which focused largely on syntactic errors.

A pioneering music-language EEG interaction paradigm showed that the early (detection) processing of musical syntax violations interfered with effects attributed to language morphosyntactic integration when the two streams were simultaneous (Koelsch, Gunter, Wittfoth, & Sammler, 2005). Participants were presented with five-element auditory chord progressions and visual sentences (simultaneous word-and-chord 1.7 Hz presentation rate, chords were computer-generated MIDI files), with the final chord or word either correct or consisting of a syntactic violation (music: a Neapolitan chord, similar to Koelsch et al. 2000, described above; language: German noun gender-mismatched to its preceding adjective). The task was to judge the grammatical correctness of sentences while ignoring music. Event-related potentials were time-locked to the critical last chord or word. Main effects in music (e.g., incorrect chords and correct sentences) were an **ERAN** (slightly right lateralized, peaking near 190 ms) and a **P600** (centro-parietal, peaking near 500 ms). Main language effects were a **LAN** (frontally distributed, peaking near 390 ms) and a **P600** (parietally distributed, latency 450 ms to 700 ms). Syntactic comprehension processes interacted across the two domains: the LAN elicited by gender disagreement was reduced when chords were Neapolitan, e.g., when it was preceded by an ERAN. (No interactions were reported for the P600). In a second experiment, the same sentences were presented along with tone sequences, the last tone either deviant or not. An MMN in response to the deviant tone did not interact with the LAN, indicating that the interaction in the first experiment was due to early syntactic processing and not to general auditory deviance processing.

In Koelsch, Gunter, et al. (2005), an interaction in the P600, an effect attributed particularly to syntactic integration in both music (Featherstone et al., 2013) and language (Brouwer et al., 2012), would have fully supported the original claim of the SSIRH. There was however no report of a P600 cross-domain interaction in either direction. It was speculated that different task demands (e.g. the unattended music versus explicit language syntax task) could have a different impact on the ERP effects, which could also account for no interaction in integration-related processes. Nevertheless, this study (without showing shared integration resources) interpreted findings such that the SSIRH could extend its shared resources from integration-only to those processes that subserve integration, such as structure-building resources. The fact that the violation detection process in music interacted with the language syntax integration suggests that a more optimal paradigm to test the SSIRH would be to test a musical condition that could potentially bypass violation detection and elicit rather later integration-related comprehension processes.

Steinbeis and Koelsch (2008) showed in a similar interaction paradigm that language processing could impact music processing, i.e., that the interference of cross-domain syntactic comprehension was bi-directional. Identical five-chord sequences (Neapolitan vs. tonic endings) and similar five-word sentences (final word correct, syntactically incorrect or semantically incorrect) were used as in Koelsch, Gunter, et al. (2005); the primary difference was that an additional semantics language condition was added, which will not be reported here. The Neapolitan chords replicated the **ERAN**, but elicited an **N500** instead of a P600 (attributed to a different music task requirement, detecting timbre deviants). The syntactically incorrect words replicated the **LAN** and **P600** findings. Relevant to the current study, the ERAN was modulated by the incorrect language syntax, and the LAN in turn was modulated by the Neapolitan chords. This bi-directional, cross-domain syntax interaction was interpreted as support for the SSIRH, although similarly as in Koelsch, Gunter, et al. (2005), syntactic *integration* resources were not shown in music.

Recently, a similar paradigm replicated the above LAN interaction under slightly different musical conditions, pairing English language syntax violations with melodic expectancy violations as opposed to harmonic violations (Carrus, Pearce, & Bhattacharya, 2013). Participants were tasked with judging the acceptability of sentences, and heard five-element melodies (presumably computer generated, since specific ms duration was listed) paired with visually presented five-word sentences (1.7 Hz presentation rate). The final note was either a high- or low-probability note (based on a computational model taken from a corpus of English hymns, Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010; since all notes were derived from a set of preexisting notes in melodies, both low- and high-probability conditions likely consisted of within-key notes), and the last word was either correct or syntactically incorrect (e.g., noun-verb number disagreement). Compared to high-probability notes, low-probability notes elicited a larger N1 component (fronto-centrally distributed, peaking near 100 ms) when words were correct. Compared to syntactically correct words, incorrect words elicited a LAN (fronto-central distribution, 250 ms to 550 ms latency) and P600 (broadly distributed, 500 ms to 800 ms latency) when notes were high-probability. The LAN was reduced for low-probability notes, while there was no interaction in the P600. This study concluded that neural resources for language syntax competed with those used for melodic expectation, and that further the low-probability note condition likely only influenced early-stage syntactic processing as opposed to later-stage integration.

These three studies (Carrus et al., 2013; Koelsch, Gunter, et al., 2005; Steinbeis & Koelsch, 2008) clearly demonstrate overlapping neural resources across domains, but not overlapping syntactic integration resources per se; while the LAN is attributed to morphosyntactic integration (e.g., Friederici, 2002), the ERAN is attributed to early violation detection mechanisms (e.g., Koelsch, Gunter, et al., 2005) and the N1 is attributed to processing unexpected melodic elements (Carrus et al., 2013). Thus these paradigms indicated that the neural resources detecting a harmonic violation interfered with language syntactic integration, which does not necessarily negate the SSIRH, but does not show proof of overlapping integration resources either (but see the next paragraph).

A landmark study focusing on within-subject comparison of music and language syntactic integration found a P600, in both domains, to be proportionally modulated by the difficulty of integrating an element into its preceding context (Patel et al., 1998). In the music experiment, the task was to judge the harmonic appropriateness of musical passages (a 2 Hz-tempo MIDI chord progression) with tonally nearby or distant deviant chords (both of which contained notes outside of the original key). In the language experiment, participants judged grammatical correctness of sentences that contained either a gardenpath effect (e.g."Some of the senators endorsed promoted an old idea of justice") or a syntactic phrase structure violation (e.g., "Some of the senators endorsed the promoted an old idea of justice."). ERPs were time-locked to a critical word after the deviation took place (in given examples, "*an* old..."). In response to the deviant chords, a negativity with right anterotemporal negativity, peaking near 350 ms was elicited, named **RATN**, and a late positivity peaking near 600 ms with posterior distribution was elicited, identified as a **P600**. The RATN was cautiously interpreted as reflecting the processing of music-syntactic violations, but left for future investigation. (Koelsch et al. 2000 noted the potential similarity of the RATN to the ERAN.) In response to the language conditions, a late positive component with posterior distribution, peaking near 900 ms, identified as a **P600** (the later latency was attributed to the auditory as opposed to visual modality). Importantly in this paradigm, the P600 was similarly modulated in amplitude for the lesser and greater 'distances' that the critical elements were from their contexts (smaller amplitude for nearby-key deviants and gardenpaths, larger amplitude for distant-key deviants and phrase structure violations). Thus the P600 was interpreted as a cross-domain marker of syntactic integration, whose amplitude reflected the difficulty with which elements could be integrated into their preceding syntactic context.

These ERP studies, which compared music and language syntactic comprehension within participants, all point toward some degree of neural overlap for the resources underlying syntactic processing in each domain. The dual-stream interaction paradigms (Carrus et al., 2013; Koelsch, Gunter, et al., 2005; Steinbeis & Koelsch, 2008) suggest the early syntactic processing of musical violations (but not unattended integration of those violations, see Koelsch, Gunter, et al., 2005) to interfere with language syntax integration, thus the greatest support for the SSIRH—which claims that *integration* mechanisms are shared across domains—comes from the study (Patel et al., 1998) that showed similar P600 modulation for respective more and less difficult integration of syntactic elements in separate music and language experiments (indeed this study helped to found the SSIRH; Patel, 2003).

It should be pointed out, however, that Patel et al. (1998) paired syntactically correct, non-preferred sentences and syntactically violated sentences only with deviant chords that could be classified as dissonant, sensory violations, since out-of-key notes were used in both deviant types. Of the three studies, only Carrus et al. (2013) used within-key 'deviants,' and at that, the musical stimuli was designed from the perspective of expectancy and not musical syntax, i.e. the tonal relationships were not accounted for beyond their likelihood of occurrence (and syntactically violated language stimuli were used). Thus when comparing syntactic integration resources in music and language, the literature calls for a within-participant paradigm that focuses on correct, complex or non-preferred structures that require syntactic integration processes in both domains without the accompanying registration of a syntactic violation. The avoidance of a violation in both domains, while introducing a non-preferred structure, ensures that resources attributed to syntactic integration may be independently triggered from violation-related mechanisms. Such a manipulation is attempted by the current paradigm, in order to ensure that elicited ERP effects are comparably attributable to syntactic integration in music and language.

So far this chapter has addressed the state of the art regarding music and language syntactic theories and relevant ERP paradigms, focusing on relevant aspects for a cross-domain empirical comparison of syntactic integration mechanisms. The additional goals of the dissertation include addressing how syntactic comprehension is influenced by entrainment to the metrical composition of music or language (and what role individual differences in musical expertise, temporal perception and working memory might play in the process—addressed in the next chapter). Entrainment and syntax across domains shall be addressed in the next sections.

## 3.2 Entrainment drives meter-syntax interactions

Theoretical and empirical aspects of entrainment to meter were addressed in Chapter 2, and the previous sections in this chapter addressed theoretical and empirical aspects of syntactic comprehension. The forthcoming sections provide a two-fold account of how entrainment to meter may impact syntactic comprehension across domains: Firstly, the theoretical boundaries of metrical and syntactic constituents often overlap in both music (Lerdahl & Jackendoff, 1983) and language (Selkirk, 2011), thus attention directed to metrically salient boundaries via entrained attention to meter coincides with attention to syntactically salient boundaries, which in turn can facilitate the integration of syntactic constituents. Secondly, entrained neurons generally contribute to a better-synchronized dynamic system, which may lead to improved cognitive efficiency (Singer, 2013; Thaut et al., 2005).

### 3.2.1 Music

This section describes the remaining aspects of the GTTM that support the current account that entrainment can impact syntactic comprehension. Subsequent EEG evidence is presented that also supports the notion that entrainment can impact musical syntactic comprehension (Nittono, Bito, Hayashi, Sakata, & Hori, 2000), however; the study only peripherally addressed the current topic, demonstrating the potential of the current investigation to add to this literature.

#### Theory

Meter is encompassed in the **GTTM** syntactic account of music, particularly in how the perceptual boundaries of metrical and harmonic groups overlap with temporal units. Section 2.4.1 described metrical structure, and Section 3.1.1 described grouping structure and preference rules of harmonic tension and relaxation. This section describes a final key element of the GTTM: time-span, which is the temporal manifestation of the other musical elements.

A time-span is defined simply as the span of time, or temporal space, occurring between two adjacent beats. As the metrical structure is a hierarchical organization of regular beats, the time span reflects this organization; time spans are larger at higher levels in the metrical hierarchy because the occurrence of the corresponding strong beat at that level is lapsing over the weaker beats of its sub-levels (see Figure 3.5). Any group is considered a time-span; the most structurally or musically important event in a time-span is considered its 'head' (usually corresponding to the most important event in a group).

It is clear from Figure 3.5 how the boundaries of multiple metrical and harmonic components overlap with the time spans in this account of music syntax. Thus, neurocognitive entrainment to the temporally predictable, time span units implies attention to the other simultaneous events, which are salient to other co-occurring dimensions of syntactic structure in the musical piece. It is in this way that syntactic comprehension can be facilitated by neurocognitive entrainment at the theoretical level: enhanced perception of salient struc-

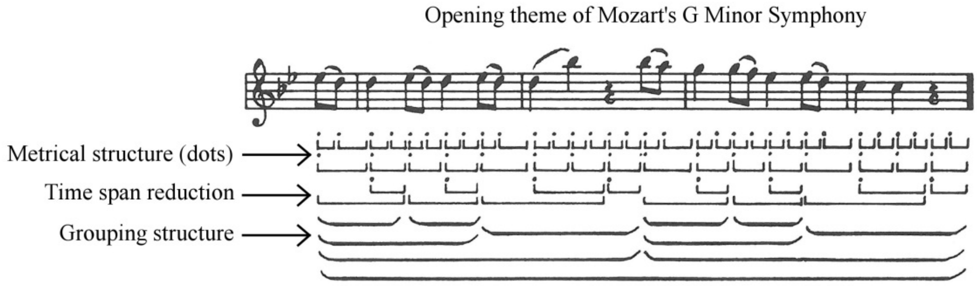



*Figure 3.5: Meter is combined with other elements of musical syntax to account for a listener's perception. Note that the time-spans and groups have the same boundaries (Figure from Lerdahl & Jackendoff, 1983).*

tural events contributes to fluid integration of new syntactic elements into their preceding musical context.

This structural facilitation combines with other converging consequences of entrainment to facilitate syntactic comprehension from multiple angles. Increased neural synchrony as a result of entrainment also leads to overall more-efficient cognition (e.g., Singer, 2013; Thaut et al., 2005, see Section 2.2.1), which contributes to syntactic comprehension from the point of view of biologically efficient processing.

### ERPs

Event-related potential studies investigating the impact of meter or timing on syntactic comprehension in music are sparse (but see Section 4.2.2 for a behavioral account of temporal coordination in music affecting syntactic judgments; Schmuckler & Boltz, 1994). The study presented below is not a paradigm geared toward entrainment, but entrainment likely occurred to the stimuli since contexts of isochronous beats were provided, and processes attributable to syntactic comprehension were consequently impacted by a temporal jitter (that likely disrupted entrainment).

Nittono et al. (2000) investigated the effect that a temporal deviation might have on the processing of deviant final notes in familiar melodies. Non-musicians were presented with familiar popular melodies (arranged as single-tone piano MIDI files presented at 2 Hz rate) with the task to rate the congruity of the ending. The final note of melodies was either well-timed or delayed (by 750 ms), correct or incorrect (raised or lowered one semi-tone) in a 2 x 2 design. An **N1** (fronto-central, peaking near 100 ms) was elicited for delayed vs. well-timed final notes. Early and late **P300** effects were elicited by both well-timed and delayed deviant final notes (P3a: centro-parietal distribution, peaking near 350 ms; P3b: parietal distribution, peaking near 450 ms). A later positivity ('slow wave'; centro-parietal distribution) was elicited for deviant well-timed notes only, which may have been a **P600**.

The amplitudes of the late positive effects were significantly reduced in the delayed condition. The reduced positive effects were interpreted such that two deviant events (omission and deviant pitch) were too close together to evoke a large P300 amplitude, i.e. in terms of deviance processing and not syntactic comprehension, though deviant notes were tonally distant from their correct counterparts.

In light of the current approach, the results of Nittono et al. (2000) could be explained such that the attentional peak from the DAT was in recovery mode; the delayed note was not even on the same cycle as a beat (a beat and a half delay), thus the attentional peak was unprepared for this important information occurring at a different point in its peak cycle. The P600 disappeared when the final note was delayed, thus syntactic reanalysis or harmonic integration were clearly affected by temporal jitter, which supports the theory adopted in the current approach.

What is not clear from this paradigm, nor from any of the ERP paradigms previously listed (Sections 3.1.1 and 3.1.3), is whether syntactic integration processes are affected by natural temporal jitter in human-played performances, particularly temporal jitter propagated by irregular metrical structure in melodies. The current paradigm shall attempt to elicit ERPs representing syntactic integration of non-preferred melodic structure, moreover in more- as opposed to less- temporally predictable musical conditions (based on the metrical composition of the melodies). Under such circumstances, it should be clear whether entrainment to the meter facilitates syntactic integration. The current approach is elaborated below in Section 3.4 and fully described in Section 5.3.

### 3.2.2 Language

Language meter has been described in Chapter 2 as it appears in poetry and phonology (metrical theory and prosody), and core concepts from language syntax have been introduced in above Section 3.1.2. Here, a theory of how phonology interfaces with syntax is introduced (Selkirk, 2011), as well as EEG evidence suggesting entrainment to impact syntactic integration (Roncaglia-Denissen et al., 2013; Schmidt-Kassow & Kotz, 2008).

#### Theory

A dissociation of governing rules for syntax and prosody is generally acknowledged (Nespor & Vogel, 1986). For example, the strict layer hypothesis posed that within prosodic hierarchy, a constituent at a particular level may never dominate a constituent of its equal level. This is fundamentally at odds with the principle of recursion found in syntax (but see Wagner 2010 for an account of prosodic recursion), however; a recent theory suggests syntactic and metrical boundaries in language to overlap more often than not (Selkirk, 2011).

In the *Syntax-Phonology Interface*, Selkirk (2011) draws from Optimality Theory (Prince & Smolensky, 2004, see Section 3.1.2) in claiming that actually the respective constituents of word, phrase, and clause in syntax and prosody should and often do match each other, but that hierarchically arranged constraint-parameters select surface permutations that sometimes cause prosodic constituent boundaries to mismatch with syntactic constituent boundaries. The assertion that constituents should be faithful to one another is termed **Match** theory (faithfulness meaning that their respective edge boundaries should align), and particularly states that syntactic clauses should correspond to prosodic intonational phrases, syntactic phrases should correspond to prosodic phonological phrases, and syntactic words should correspond to prosodic words.

Although 'faithfulness' is default, the faithful alignment of syntactic and metrical boundaries can sometimes be prioritized lower than other surface rules in a language. For example, "prosodic markedness constraints"—e.g., minimally or maximally-sized prosodic words, assigning the left-most constituent automatic strength (called "strong start"), stress prominence assignment, or tone assignment to prosodic stress—may be ranked higher than Match faithfulness constraints when a surface structure is formed. Thus it is still possible for languages to produce utterances that contain conflicting prosodic and syntactic boundaries.

In light of the current approach, Selkirk (2011) provides a theoretical explanation for how syntactic comprehension may be influenced by entrained attention (e.g., the DAT, Jones, 1976; attentional bounce, Pitt & Samuel, 1990). Selkirk (2011) claims that phonological and syntactic boundaries align as a rule, at multiple constituent levels. Therefore, if the phonology is organized in such as way that the meter is regular, attention entrained to meter simultaneously entrains to salient syntactic boundaries. As mentioned in the previous chapter (Section 2.1.1), events occurring at or near the peak of the attentional pulse will be best perceived (e.g., Large & Jones, 1999); when an event occurs while the attentional pulse is far from the peak, perception of that event will be comparatively poorer. Thus in the current example where entrained attention to (regular) meter co-occurs with attention to syntactic boundaries, the attentional peak at the syntactically salient locations can enhance syntactic processing compared to (irregular) metrical context that does not encourage entrainment. This in combination with increased neural synchrony resulting from the entrainment, which can enhance cognitive efficiency (e.g., Singer, 2013; Thaut et al., 2005, see Section 2.2.1), speaks for facilitated syntactic comprehension in contexts where entrainment occurs. The next section presents EEG examples of meter-syntax interactions.

### ERPs

Chapter 1 presented a key ERP study showing syntactic reanalysis facilitated by a regular compared to irregular metrical context in language (Roncaglia-Denissen et al., 2013, see

Section 1.1.2). To recapitulate, Roncaglia-Denissen et al. (2013) presented non-musicians with spoken sentences that contained locally ambiguous subject- and object-relative clauses, and the sentences were composed with either regular or irregular meter (stressed syllables with apostrophe: regular example, ‘Roland triff die ‘Diener, die An’tonio mal ge’stört haben, im Park./*Roland meets the helpers, who once bothered Antonio, in the park.*; irregular example, Bernhard trifft die Ge’hilfen, die Ni’cole mal gestört haben, im Park./ *Bernhard meets the helpers, who once bothered Nicole, in the park.*) When participants listened to the sentences (performing a comprehension task after each trial), a P600 was elicited by the critical word that disambiguated object-relative clauses (non-preferred in German). Crucially, the P600 was reduced when the meter had created a regular as opposed to irregular context.

While effects were interpreted such that regular metrical units were easier for the brain to parse, the hypothesis adopted by the current approach also serves this data: enhanced attention to overlapping metrical and syntactic boundaries leads to an overall better integration of syntactic constituents. The theoretical overlap is underscored by better neural synchronization and, in turn, better cognition. To be clear, the argument is not that neurocognitive entrainment facilitates the syntactic comprehension of individual items occurring at specific beat locations; the improved comprehension likens rather properly greased teeth of gears that fit together smoothly.

Two other ERP studies strengthen the argument that language syntax integration is affected by entrainment to the temporal predictability of strong syllables in words (see Pitt & Samuel, 1990). In Schmidt-Kassow and Kotz (2008), the latency of a P600 effect was time-locked to a temporally predictable syllable as opposed to occurring much sooner, when a syntactic violation became obvious by an omitted word, suggesting language syntactic comprehension to be influenced by temporal predictability of salient lexical items. Then in Schmidt-Kassow and Kotz (2009b), the metrical identity of words was shown to be a necessary perceptual landmark for syntactic reanalysis when a syntactic P600 was canceled-out in syntactically violated words that incorrectly contained reversed strong and weak syllable weights. Thus language syntax resources seem sensitive to metrical features, specifically later reanalysis-related processes show sensitivity to temporal and metrical integrity of words.

Considering these language syntax ERP studies together, it seems a consistent phenomenon that metrical regularity impacts language syntactic comprehension, one which the current neurocognitive-entrainment approach can account for. Particularly in instances where stimuli triggers syntactic integration in correct but non-preferred structure (e.g., Roncaglia-Denissen et al., 2013), it would seem that entrainment to metrically regular contexts can facilitate cognition, marked by reduction of the P600.

### 3.2.3 Cross-domain comparison

The above sections have shown that there is a theoretical overlap of metrical and syntactic boundaries, in both music (Lerdahl & Jackendoff, 1983) and language (Selkirk, 2011). Therefore, enhanced attention to predictable metrical hierarchy (described by, e.g. Large & Jones, 1999; Pitt and Samuel (1990), see Section 2.3) implies enhanced attention to the coinciding syntactic events. This entrained attention is only part of the 'neurocognitive entrainment' assumed in the current approach that can lead to facilitated syntactic comprehension: As presented in the previous chapter (Section 2.2), increased synchrony among neuronal populations (propagated by delta- and theta range neural oscillations that underlie the perception of rhythmic beats and syllables, e.g. Schroeder & Lakatos, 2009; Giraud & Poeppel, 2012) is perpetuated throughout associated dynamic neural communication, leading to generally increased synchrony among neural populations which perform the cognitive operations (Singer, 2013).

When empirically testing the impact of entrainment on syntactic comprehension across domains, an ideal process that surfaces in the literature is syntactic integration. The SSIRH already posits that the online integration of syntactic constituents is subserved by a resource shared by music and language domains (Patel, 2003), indexed by similar ERP responses (the P600; e.g., Patel et al., 1998), thus this integration resource should show transparent cross-domain modulation by entrainment. Considering the literature surveyed (in sections 3.1.1, 3.1.2, and 3.1.3), the syntactic construction most likely to engage purely integration mechanisms (as opposed to violation-detection mechanisms) is one that is technically correct but containing a structural complexity or non-preferred structure.

As argued here, the possibility exists in music and language domains for entrainment to affect syntactic comprehension. However, the only EEG paradigm to give concrete evidence for this was in the language domain: Roncaglia-Denissen et al. (2013) showed syntactic integration of correct, non-preferred structure to be facilitated when participants had the opportunity to entrain to a regular metrical context (though results were interpreted within a different, but not incompatible, framework). In the music domain, neural entrainment has been demonstrated to coincide with meter perception (Nozaradan et al., 2011, see Section 2.4.1) and temporal deviance has disrupted syntactic comprehension (Nittono et al., 2000), but (entrainment to) meter and syntax has thus far not been unified within a single investigation. The current approach, described in the Section 3.4, offers a unified approach to cross-domain entrainment to meter and its impact on syntactic comprehension.

## 3.3 Open issues

In the decade since the SSIRH was proposed, counter-theories have emerged to account for similarities of cross-domain syntactic comprehension. Notably, Slevc and Okada (2015) recently suggested that the SSIRH 'syntactic resource' is alternatively better attributed to cognitive control. The paper has suggested that most experiments supporting the SSIRH are not, in fact, showing overlapping syntactic integration resources but rather showing overlap in (supra-domain) cognitive control. Slevc and Okada (2015) argue that the "conflict detection" and "resolution" processes employed in syntax paradigms, in either domain, are not specific to syntactic processing but rather general to cognitive control mechanisms. A prime argument for their viewpoint was a result from Perruchet and Poulin-Charronnat (2013), that showed a semantic gardenpath sentence to interact with music syntactic comprehension. The "shared syntactic integration resource" was therefore more general than syntax. Thus interactions initially thought to be driven by shared neural syntactic integration resources (that could not repair simultaneous music and language errors, e.g. a behavioral study by Slevc, Rosenberg, and Patel 2009) were claimed by Slevc and Okada 2015 to have been confounded with cognitive control allocation. This paper did not say to which degree or how syntactic integration is different in music and language, and acknowledged that much empirical work is necessary to clarify the nuances of purely syntactic versus purely cognitive integration mechanisms.

The current investigation cannot dissociate between shared syntactic resources and cognitive control. Experimental manipulations requiring conflict detection and resolution of complex structure are argued by Slevc and Okada (2015) to trigger cognitive control; these are precisely the types of manipulations suggested by the literature to highlight syntactic integration mechanisms and to bypass violation detection, and what is used by the current approach comparably across domains. Thus cross-domain similarities in ERP responses to non-preferrd syntactic structure may be a result of cognitive control employed similarly across the two domains. This bleeds into a larger debate, acknowledged by Slevc and Okada (2015), that argues cognitive control mechanisms to confound results in psycholinguistics, and is outside of the current scope.

## 3.4 Current approach

The arguments in this chapter have outlined ideal circumstances to test cross-domain comparability of syntactic integration, as well as provided an account for how entrainment to a regular metrical context could facilitate syntactic integration. Thus in the current approach, the metrical context of respective melodies and sentences will have alternative regular and irregular versions, the regular intended to induce entrainment. A primary research question is whether cognitive resources entrained to the regular meter facilitate syntactic comprehension. (The metrical experimental conditions were introduced previously in Section 2.6 and a full description may be found in Section 5.3.1.)

Syntactic integration is an ideal process to test the effects of entrainment. However, a cross-domain comparison of integration-only mechanisms is still wanting in the literature; previous within-participant, music-language syntax paradigms have either used violations in both domains Koelsch, Gunter, et al. (2005) or asymmetrically used violations and non-preferred syntax across domains (Carrus et al., 2013; Patel et al., 1998, see Section 3.1.3). Therefore a secondary research question is whether non-preferred relative key resolution in melodies is comparable to non-preferred relative clause attachment in sentences, two structures that do not require violation detection. The two syntactic situations are not equivalent, but share the 'alternate ending' option that is not incorrect either way, rather simply based on the preference of the listener (see Section 5.3.1).

The current hypothesis is that a central or posterior P600 will be elicited by non-preferred relative key- and relative clause attachment (the relative clause attachment preference of native Germans is addressed in Experiment 3b), and that a reduction of the P600 in regular- as opposed to irregular-meter contexts will signify facilitated syntactic integration.

## 3.5 Summary

This chapter has presented syntactic theory and accompanying ERP evidence for relevant aspects of cross-domain syntactic comprehension, highlighting syntactic integration of correct, non-preferred structure. A theoretical account was provided for how entrainment to regular meter could facilitate syntactic integration in either domain, and relevant ERP studies in support of this view were surveyed.

One final aspect of the current approach remains: the role that individual differences in musical expertise, temporal perception, and working memory may have with respect to entrainment to meter and syntactic comprehension. The next chapter addresses individual differences in the framework of the current paradigm.

# Chapter 4

# Individual Differences

Introduced thus far are basic concepts of meter and syntax along with their relevant empirical correlates, as well as an account of how neurocognitive entrainment to meter can facilitate syntactic comprehension. The current chapter addresses one final angle to the empirical approach of this dissertation: accounting for individual differences in the healthy testing population.

The study of individual differences extends far and wide, including pathology (psychiatric conditions), personality, intelligence, executive control, memory, and more. Theories of individual differences accounting for individual traits (a long-term, stable characteristic of an individual) and any overlap with states (the immediate contextual circumstance of an individual) include theories by Humphreys and Revelle (1984), Revelle (1993), and Corr (2008). This field espouses the value of accounting for individual differences in empirical research, as traits and states interact to influence data collected in experimental settings and may inform important distinctions in neurocognition associated with particular tasks (Corr, Revelle, Wilt, & Rosenthal, 2010; Yarkoni & Braver, 2010). The scope of this dissertation is narrowed to include an account of individual traits specifically relevant to entrainment, meter perception, and syntactic comprehension, leaving for example effects of personality or anxiety on task performance to future veins of research.[1]

Musical expertise, temporal perception, and working memory are germane to investigating whether entraining to meter impacts syntactic comprehension in music and language. Below, the empirical role that these traits play in meter perception and syntactic comprehension are described, with the intention to demonstrate the utility of accounting for these individual differences in the current approach.

[1]Traits or states that severely affected performance in Studies 1, 2, and 3 were hopefully initially reduced by participant inclusion criteria (no psychiatric illnesses) or subsequently pruned away on account of task-performance inclusion criteria.

## 4.1 Musical Expertise

Members of a culture may develop tacit musical expertise, which allows them to recognize common harmonic structures and identify moods associated with musical passages although they are non-musicians (Bigand & Poulin-Charronnat, 2006; Sloboda, 1991). Musicians become (express) experts by way of deliberate practice (Ericsson, Krampe, & Tesch-Römer, 1993; Ericsson, Prietula, & Cokely, 2007), and they may have natural musical talent to begin with (see Hambrick & Meinz, 2011 and Ericsson et al., 2007 for arguments for and against a mediating role of natural aptitude in expertise).

Empirically, musical expertise is often investigated by grouping participants according to whether they are musicians or not (e.g. Besson & Faïta, 1995) or according to degree of expertise (such as non-musical, amateur, and professional musicians; e.g. Oechslin, Ville, Lazeyras, Hauert, & James, 2013). Other methods include using the age at which musical training started, the number of deliberate practice hours in childhood, or the total years of instrument training as covariates (Brod & Opitz, 2012; Wong, Skoe, Russo, Dees, & Kraus, 2007). Specific aspects of musical aptitude may also be indexed, such as separate melody and rhythm skills (e.g. Montreal Battery of Evaluation of Amusia, Peretz, Champod, & Hyde, 2003; Musical Ear Test (MET), Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010). This thesis employs a musician/non-musician grouping, weekly practice hour estimates, and musical skill assessment (the MET, Section 5.2.3) to account for musical expertise among participants.

Musical expertise is documented to have wide transfer effects beyond musical performance, and musical training is espoused as an ideal paradigm to study brain plasticity (Wan & Schlaug, 2010). For example, musical training is linked to improved verbal intelligence, second language acquisition, and executive functioning (Moreno et al., 2011; Shook, Marian, Bartolotti, & Schroeder, 2013; for review see Miendlarzewska & Trost, 2013). Musical training (compared to basic levels of enculturation) is also associated with plastic changes in central auditory processing (e.g. in the auditory cortex), in the motor network (e.g. motor cortices associated with instrument playing), and in multimodal integration (e.g. brain regions attributed to audio-visual, auditory-motor, and auditory-visual-sensorimotor integration; Pantev & Herholz, 2011; Zendel & Alain, 2012; for review see Herholz & Zatorre, 2012). Specifically relevant to this thesis, musical expertise also impacts meter perception and syntactic comprehension across music and language domains; these areas of study, outlined below, highlight the value of accounting for participants' musicianship in the current approach.

### 4.1.1 Meter perception

Physiological and behavioral evidence presented below suggest that musical expertise offers an advantage to meter perception in both music and language. In music, this advantage is related to more efficient processing of small metrical constituents and to better awareness of structural hierarchy in meter (Geiser et al., 2010; Palmer & Krumhansl, 1990). In language, an improved perception of syllables is hypothesized to drive better access to meter (Magne et al., 2007; Marie et al., 2011).

#### Music

Behavioral evidence suggests that musicians 'tune-in' to meter at a finer timescale than non-musicians (Palmer & Krumhansl, 1990). A behavioral experiment asked musicians (with at least five years of musical training) and non-musicians (with less than two years of musical training) to judge the appropriateness of probe tones played in various temporal positions relative to an underlying context beat. The context beat corresponded to the downbeat of 2/4, 3/4, 4/4, and 6/8 meter when played at respectively 69, 75, 69, and 75 bpm (each tempo was a previously published convention for its meter). Musicians rated the probe tones most appropriate when they represented smaller periodic ratios (corresponding to eighth- and quarter-note divisions per measure) and non-musicians rated larger periodicities most appropriate (corresponding to strong-beat positions per measure). Results were interpreted such that the ability to form abstract metrical representations was present in both musicians and non-musicians, but that additional theoretical knowledge of musicians allowed 'richer' mental representations of metrical hierarchies extending to smaller periodicities.

In line with the above findings, ERP evidence indicates that musicians are more sensitive to perceiving a metrical hierarchy than non-musicians (Geiser et al., 2010): In an MMN paradigm comparing musicians (trained in percussion) and non-musicians (lacking musical training), the ERP response of standard (quiet) tones in a metrically strong position were compared to deviant (loud) tones that were presented either in the same strong beat position (metrically congruous context) or shifted one beat earlier, giving the impression of a changed beat count for one measure (metrically incongruous context). In both musicians and non-musicians, an MMN was evoked for both deviant types, with a larger amplitude for the meter-incongruent context. While non-musicians had statistically comparable MMN amplitude in both deviant types, musicians had a smaller MMN in the metrically congruent context and a larger MMN in metrically incongruous context. This larger difference between meter conditions, specifically the lesser response to meter-congruent and greater response to meter-incongruent deviances, was interpreted by the authors to indicate a higher sensitivity to metrical structure in trained musicians.

#### Language

With regard to language meter, ERP evidence indicates that musicians better integrate metrical structure in speech than non-musicians (Magne et al., 2007; Marie et al., 2011). In a pair of ERP studies, combined results from two tasks showed that musicians' perception of meter was improved over non-musicians in both detection and integration of metrical speech features. In the first task, participants judged whether the metrically manipulated (uncharacteristically lengthened) final word in French sentences was well pronounced or not. An early positivity (P200) elicited by the metrical violation was larger in musicians (mean of 17 years of musical training) compared to non-musicians, which was taken to indicate improved perception of the temporal properties of speech, reflecting the "automatic processing of the temporal attributes of...words' syllable structure," further explored below (see Section 4.2.1). In the second task, participants were asked to judge the semantic congruity of sentence-final words. A P600 was elicited for metrically incongruous words in this semantic task in musicians, but not in non-musicians. Considering a previous interpretation that the P600 represented general reanalysis of violated language-meter rules (Schmidt-Kassow & Kotz, 2009b), this presence of the P600 was taken to reflect musicians' sensitivity to metrical incongruity independently of where attention was directed. Thus group differences in the P200 and P600 implied that musicians have enhanced perception of meter in speech, in both initial perception of syllables and later processes that integrate metrical violations.

#### Summary

The advantage of musical expertise in music meter perception includes improved perception of metrical constituents at a small timescale and greater sensitivity to overarching hierarchical relationships among those constituents. This temporally fine-grained perceptual advantage seems to persist in language, improving general access to meter, although language meter is not often organized in such a way as to be able to test perception of metrical hierarchy comparably to music.

### 4.1.2 Syntactic comprehension

Tacit musical expertise based on enculturation affords detection of overt deviances from expected musical norms (Bigand & Poulin-Charronnat, 2006; Koelsch et al., 2000). Nonetheless, behavioral and ERP evidence presented below show a processing advantage of syntax in groups of adults and children who have had music training, which persists in both music and language. Particularly musicians have better implicit early perception of syntactic violation in both domains (Fitzroy & Sanders, 2013; Koelsch et al., 2002), better explicit integration in music when theoretical knowledge informs harmonic nuance or supplies implied harmony (Besson & Faïta, 1995), and better learning of language syntax structures that parallel hierarchical syntactic relationships found in music (Brod & Opitz, 2012).

### Music

Behaviorally, while non-musicians can easily detect overt harmonic incongruity, musicians are better able to distinguish and classify subtle harmonic incongruity than non-musicians. One experiment asked adult participants to categorize melodies that contained complete congruity, in-key melodic contour incongruity or an out-of-key harmonic/melodic incongruity (Besson & Faïta, 1995). Musicians outperformed non-musicians in general, but little difference existed between groups for out-of-key incongruities while gross differences occurred between groups when categorizing within-key incongruities. Complimentary results exist in children. Primary school children with and without extra-curricular musical training (mean two years) were asked to classify musical excerpts according to the appropriateness of their endings, which were either congruous, markedly incongruous or subtly incongruous (James, Dupuis-Lozeron, & Hauert, 2012). All children on average were able to distinguish harmonic incongruity, indicating implicit learning of musical syntax. Children with musical training were better able to distinguish congruent from subtle- and marked incongruities than their non-musician peers. These studies demonstrate the sufficiency of enculturation to make broad music-syntax distinctions but that distinguishing harmonic nuance requires musical training.

Early implicit detection of music syntax violations is evidenced by ERPs to be enhanced in musicians. A study with adult musicians (with at least ten years of instrument training) and non-musicians (with no musical training outside school) presented participants with chord sequences, ending in either expected or deviant cadences (Koelsch et al., 2002). In response to the deviant chords, musicians showed an increased ERAN amplitude compared to non-musicians. Similarly, in a study with 10-to-11 year old children where slightly and severely deviant musical chords were inserted into chord progressions, an ERAN was elicited in groups with (mean five years) and without musical training, but its amplitude was significantly larger in musically trained children (Jentschke & Koelsch, 2009). An N500 reflecting harmonic integration was elicited in both studies, but showed no effect of musical expertise in either adults or children. Thus the non-musician is indeed a ‘tacit musical expert’ who appropriately detects and implicitly integrates musical syntax violations, but musical expertise contributes an enhanced early detection of harmonic deviances.

Event-related potentials reflecting explicit harmonic integration are present in musicians and non-musicians, with little-to-no group differences when stimuli provides full harmony (i.e. is polyphonic). Two studies that used chord sequences as musical stimuli report similar findings: A late positivity was elicited by both subtle and pronounced harmonic deviances, whose amplitude did not significantly differ between groups of musicians and non-musicians (Fitzroy & Sanders, 2013; Regnault, Bigand, & Besson, 2001). (Fitzroy and Sanders 2013 found expertise-modulated distribution differences in the P600 elicited by subtle harmonic deviance, which may indicate different processing sources, but authors did not interpret this finding.) The studies used different criteria for musicianship, Regnault et al. (2001) using more stringent separation between groups (musicians at least ten years of musical training, non-musicians no musical training) and Fitzroy and Sanders (2013) using less strict classification (musicians and non-musicians had 30% overlap in the number of years of instrument training, although non-musicians had less theory training than musicians). The consistency across studies despite musical-expertise criteria differences suggests that the phenomenon is robust, that musicians and non-musicians have similar explicit, later integration of harmonic deviance when heard in the context of polyphonic sequences.

In contrast, when harmonic information is implied as opposed to overt (such as in monophony), expertise modulates the explicit integration of syntactic deviance. In one ERP study which asked participants to categorize simple piano melodies according to varying degrees of harmonic congruity, it was found that the amplitude and latency of a late positivity (P600) in response to out-of-key incongruities was greater in musicians (with at least seven years of musical training) than non-musicians (with no musical training) for both familiar and unfamiliar melodies (Besson & Faïta, 1995). A late negativity elicited by incongruities (both in- and out-of-key) in unfamiliar melodies was also larger in musicians than in non-musicians. Group differences disappeared in all effects when the task was changed from explicit judgment per melody to general questions at the end of large blocks. Thus the role of expertise affects explicit integration of harmonic deviance when theoretical knowledge contributes underlying harmonic structure not present in monophonic stimuli.

### Language

ERP evidence indicates that musical exposure facilitates development of early implicit language syntactic comprehension in children (Jentschke & Koelsch, 2009). Ten- and 11-yr olds listened to sentences with a syntactic phrase structure violation (a missing critical noun from a prepositional phrase), with the instruction to detect a speaker change in catch-trials. Children with musical training (mean five years) showed an early anterior negativity (ELAN) whereas their non-musician counterparts did not. A subsequent sustained negativity (associated with developing syntactic processing) was also elicited in both groups, larger in amplitude for musically trained children. As the ELAN is not usually elicited in children in this age group (and the sustained negativity verified that their syntactic processing was still under development), musical training was therefore concluded to influence early development of adult-like automatic detection processes in language syntactic comprehension.

Musical expertise modulates early processing of language syntax violations in adults as well (Fitzroy & Sanders, 2013). When presented with task-relevant syntactic insertion violations such as “her” in “. . . write letters to those her friends,” an early anterior negativity (EAN) and a late positivity (P600) were elicited in both musicians (ranging three years to over ten years’ experience) and non-musicians. The EAN was left lateralized and more focal in musicians, while the P600 showed no group differences. A previous, paradigmatically similar study investigated language (as opposed to musical) expertise, presenting high- and low-language-proficiency groups (controlled for non-verbal intelligence) with the same language stimulus materials and task (Pakulak & Neville, 2010). Results from the language-only study suggested that compared to low-proficiency, high language proficiency promoted more efficient early, automatic processes (a smaller, more focal EAN), which in turn allowed for a freed-up, more robust later controlled syntactic integration (a larger P600). Thus when considering both studies, it would appear that musical expertise, like language expertise, improves efficiency of early automatic detection of syntactic violations. It is quite possible that the freeing-up of later controlled language syntax resources also occurs in musical expertise groups, but that more strict separation of music-proficiency is necessary to capture the effect (as mentioned above, Fitzroy and Sanders 2013 musician- and non-musician groups had looser separation criteria, whereas Pakulak and Neville 2010 groups had strict separation criteria in language proficiency).

One study asserts that musicianship only affects language syntactic comprehension insofar as the syntactic structure is hierarchical and reminiscent of hierarchical tonal relationships found in music syntax (Brod & Opitz, 2012). In an experiment concerned with the transfer effects of musical expertise, adult musicians with a broad range of expertise (quantified by the number of deliberate practice hours undertaken before the age of 19) learned an artificial language with both local and long-distance/hierarchical syntactic dependencies. Measures of long-distance-dependency acquisition correlated with musical expertise whereas measures of local-dependency acquisition did not, leading authors to surmise that musical training may result in “an improvement in probabilistic learning of long-distance dependencies in general,” but that musical expertise should not affect local-type language syntax violations.

The claim from this last study, if qualified to assert that musical training modulates only long-distance or hierarchical language syntax integration processes, may easily be reconciled with the first two ERP studies in this section (Fitzroy & Sanders, 2013; Jentschke & Koelsch, 2009), both of which showed modulatory effects of expertise on language syntax violations that did not necessarily involve long-distance dependency. Group differences in ERPs in those studies were limited to early detection processes, and did not manifest in later effects associated with syntactic integration (see Hahne and Friederici 1999 for an account of the separate detection and integration language syntax processes and their corresponding ERPs). Thus musicians, due perhaps to their superior perception of syllables and phonemes (Perfors & Ong, 2012; Strait, O'Connell, Parbery-Clark, & Kraus, 2014), are more efficient at recognizing the violation, but then do not differ from non-musicians in processing the non-hierarchical violation once detected. One may speculate that any later integration-related ERP effects elicited by long-distance dependency language manipulations such as described by Brod and Opitz (2012) would be modulated by musical expertise; a wider range of language stimuli in ERP paradigms investigating musical expertise transfer effects on language syntax is needed to broaden this scope of knowledge.

**Summary**

Enculturation is sufficient to build basic musical syntactic expectancies in non-musicians; however, musical expertise seems to offer an advantage to syntactic comprehension in both music and language. Early automatic detection of cross-domain syntax violations shows enhanced ERP response in musicians compared to non-musicians, and specifically in music, musicians are sensitive to more subtle structural manipulations and can abstract harmonic structure from less acoustic information. These advantages appear as early as childhood with only a few years of musical training. In language, musical expertise may only affect language-syntactic integration in long-distance dependency conditions, although broader stimuli sets are needed to explore this arena.

## 4.2 Temporal perception

Individual differences in temporal perception may influence performance in tasks that involve careful monitoring of music and speech signals; it is therefore relevant in the current paradigm to account for how well participants actually track events. Interestingly, the "sweet spot" for ideal human time estimates reported in the literature (from 500 ms to 700 ms; Eisler, Eisler, & Hellström, 2008; Fraisse, 1957; Friberg & Sundberg, 1995) falls into the same band width that is hypothesized to both represent low-frequency neural oscillations underpinning the DAT attending rhythms (4–8Hz; Henry & Hermann, 2014) and the ideal frequency range for speech perception (Luo & Poeppel, 2007). Thus temporal discrimination thresholds within this range, such as those found by anisochrony detection tasks (described below), may influence the extent to which a person successfully attends to valuable information from the environment (e.g., meter and syntax information in music and language). This section will review the DAT as well as introduce paradigms that measure temporal perception, focusing on anisochrony detection.

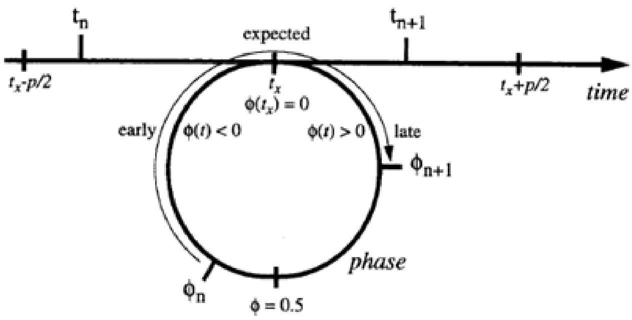



*Figure 4.1: The perception of events in time as described by the DAT. At the expected location in time (tx), the phase of the attentional oscillation is set to zero (Ω(tx) = 0). This is where the attentional pulse will occur (not shown). Events occurring earlier/later in time than expected appear earlier/later in the phase of the attentional oscillation (Ω(tx) < 0/ Ω(tx) > 0), triggering an adjustment to the phase/location of the attentional pulse. Figure from Large and Jones (1999).*

The DAT offers a salient model of perception of events in time, although numerous other models exist. Among others, alternative accounts posit timing perception governed by neural network state dynamics (Buonomano, 2007; Karmarkar & Buonomano, 2007), attentional or memory cognitive mechanisms (Block, 2003), other oscillator mechanisms (Church & Broadbent, 1990; Treisman, Faulkner, Naish, & Brogan, 1990), or an internal pacemaker-counter device (Gibbon, 1977; Treisman, Faulkner, Naish, & Brogan, 1963; Wearden, 2003). Of special relevance here is that the DAT accounts for attention to environmental events that contain a hierarchical timescale while being robust to deviations from perfect periodicity among those events (such as human-performed melodies and naturally spoken speech). A deeper discussion of the DAT compared to other types of timing theories may be found in Large and Jones (1999), and an overview of theories of temporal perception may be found in Grondin (2010).

The DAT describes an internal attentional oscillation that couples to temporally regular external events and the attentional pulse that rides atop this oscillation (explained in more detail in Section 2.1.1). This accounts for temporal perception of events such that events occurring before the peak of attention-—or expected temporal occurrence of the event— are perceived as early, and events occurring after the peak are perceived as late (Figure 4.1; Large & Jones, 1999). When events do not match their expected temporal occurrence, the period and phase of the internal oscillator are adjusted in order to re-attempt synchronization to the environmental stimuli, thereby continuously fine-tuning perceived timing of consecutive events. Many paradigms investigating temporal perception rely on for example interval discrimination (Grondin, 2010): in this case detection of an early event would aid in discriminating a shorter-than-previous interval, and perceived late events would help discriminate a longer-than-previous interval. One way to measure temporal perception is via some discrimination threshold, for example, hearing an interval or series of intervals and judging whether some target interval is the same or different from its previous standard(s). Anisochrony detection is one such test, which–when presented along with an algorithm that finds a discrimination threshold via a psychometric function—finds the minimal time difference that a participant needs in order to perceive the difference between two intervals (see Section 5.2.6; Dalla Bella et al., 2016; Grassi & Soranzo, 2009; Hyde & Peretz, 2004). Other methods include participants reproducing temporal intervals by tapping or button press (see Repp 2005 for a review), and while this estimate is valuable it involves the motor system and sensorimotor synchronization, which is outside the scope of this dissertation. Empirical studies are sketched below, with the intention to illustrate which aspects of cross-domain meter perception and syntactic comprehension may be impacted by individual differences in temporal perception.

### 4.2.1 Meter perception

The nature of meter perception research relates to temporal processing differently in music than in language. Empirical evidence provided below shows that, conversely, meter influences temporal perception in music whereas accurate temporal perception influences meter perception in language. Empirical work within the music-meter framework has shown that the nature of temporal processing changes when events are structured within meter (in line with dissociated future-oriented vs. analytic dynamic attending, see Section 2.1.1) and temporal discrimination threshold is improved by a strong metrical context (Ellis & Jones, 2010; Grube & Griffiths, 2009). While the impact of meter on temporal processing presumably are similar in language (indeed language-meter studies investigating hierarchical perception are rare), language-meter perception and its ties to temporal perception has occupied a much narrower strain of empirical research, one which centers on (stressed) syllable-level encoding: The vowel nucleus or p-center carries the temporal identity of the beat and improved temporal encoding of syllables provides improved access to language meter (Greenberg et al., 2003; Magne et al., 2007; Pitt & Samuel, 1990).

#### Music

A metrical as opposed to non-metrical context influences the nature of temporal perception. One experiment (Ellis & Jones, 2010) presented participants with tone sequences, with the task to judge as fast and accurately as possible whether the last target tone was higher or lower in pitch than preceding tones. The intervals between tones either imitated musical meter (metrical condition), or the same intervals were presented in a different order which did not induce a sense of meter (scrambled condition); intervals directly preceding the target tone, or foreperiods, were meaningless in the scrambled condition but were consistent with coupled periods found in the metrical condition. The nature of participants' reaction-time response was fundamentally different depending on whether the temporal organization of the preceding tone sequence was metrical or scrambled: When presented in a scrambled context, the reaction times occurred in a negative relationship to foreperiods, suggesting probabilistic reactions to the temporal location of targets. When presented in a metrical context, reaction times occurred in a positive as opposed to negative trend, indicating that temporal anticipation was no longer probabilistic. Moreover, reactions were faster in the metrical context, suggesting that participants entrained to the various metrical levels in the tone sequences and accordingly (better) timed their responses. The dissociated temporal perception strategy modulated by temporal predictability/metrical hierarchy of stimuli supports the DAT's future-oriented vs. analytic attending (Jones & Boltz, 1989).

A strong musically metrical context improves temporal discrimination. One study (Grube & Griffiths, 2009) presented participants with standard tone sequences that either strongly or weakly induced a sense of musical meter, and subsequently participants were asked to identify which of two consecutive tone sequences contained deviation from the temporal intervals of the standard sequence. The presentation was adaptive, finding each participant's temporal discrimination threshold of intervals occurring in the target sequences. Participants had better/lower discrimination of temporal intervals in strongly metrical sequences, and moreover found the task easier to perform for metrically strong sequences. These results support the notion of sharper attentional pulses attending events that occur within temporally nested oscillations as described by the DAT (Jones, 2008; Large & Jones, 1999).

As mentioned previously in Section 2.4.3, meter is often synonymous with timing in music experiments. For example, in the above Ellis and Jones (2010) study, the metrical condition is perfectly periodic, thus the findings apply (perhaps strictly) to a hierarchy of perfect temporal periodicities.[2] Considering previous unclear distinction between temporal precision and metrical organization, it is especially salient to monitor the timing abilities of participants when investigating their entrainment to or perception of different metrical structures in the current investigation. Moreover, the emergent picture from these studies is that musical meter influences temporal perception in a top-down fashion; the current approach with natural stimuli opens the possibility for timing to influence metrical perception in the opposite direction (as it is suggested below to be the case in language), for example, if participants have poorer timing they perhaps have less ability to form an abstract metrical representation when performance of music includes human timing errors.

[2]The second-mentioned experiment, Grube and Griffiths (2009), necessarily employed perfectly timed meter because the objective was to determine whether participants could discern small temporal deviations.

**Language**

The fundamental location of ‘beat’ in the complex speech signal, which is necessary for establishing meter, has been the focus of some research. The nucleus of vowels has been suggested to be a perceptual center or **p-center**, and this coincides with where humans perceive the relative stress or de-emphasis of syllables (Greenberg et al., 2003). It has been suggested that a p-center is the ‘event’ whose temporal regulation draws the attentional pulse in the DAT, and further that attention is guided more by strong syllables than weak syllables (Cummins & Port, 1998; Pitt & Samuel, 1990; Port, 2003). Temporal encoding of sounds in the brain is necessary for perceiving this energy peak in a speech signal. This syllable encoding interfaces with meter, as the syllable weight, length, or relative stress comprises metric identity in poetry or meter phonology (Fabb, 2001; Prince, 1983, see Section 2.4.2).

In line with this, two studies suggested that temporal encoding of syllables yields subsequent access to language meter. Magne et al. (2007) (mentioned already in sections 2.4.2 and 4.1.1), stretched the final syllable of French phrases while preserving non-temporal aspects of the signal such as pitch and volume. When participants judged the pronunciation of the sentences, a late positivity (P600) was elicited in response to the temporally violated metrical structure. A similar late positivity was evoked by (attention to) language meter violations that manipulated both temporal and non-temporal acoustic aspects of German target syllables, i.e. typically weak syllables received incorrect strong metrical emphasis (Schmidt-Kassow & Kotz, 2009b). More research is needed to explore timing-only manipulations to meter in various rhythm classes, however; evidence thus far suggests that the temporal integrity of syllables, when alone disrupted, is enough to jar the metrical structure of sentences.

The improved access to language meter exhibited by musicians (Marie et al., 2011) may be due to improved absorption of information at small timescales. Numerous studies have been conducted that show how musical exposure can improve the temporal tracking of syllables in speech (e.g. Strait et al., 2014; Strait, Parbery-Clark, Hittner, & Kraus, 2012; Tierney, Krizman, Skoe, Johnston, & Kraus, 2013); improved temporal encoding (via increased musical exposure) improves the perception of syllables, thus the monitoring of individual differences in timing abilities is valuable when investigation the perception of language meter. Moreover, the current use of natural human timing in melodies and sentences, with comparable hierarchy in the metrical structures (see stimuli description, 5.3.2), may help to reconcile opposite directions of the meter-timing relationship described in the literature across the two domains.

### Summary

Meter perception is entwined with temporal perception in both music and language. Particularly, music studies show that meter alters temporal perception, and language studies show that accurate temporal perception is crucial to accessing meter. The precise (circular) relationship of meter- and temporal perception across domains is an area that globally requires more research, but the few studies mentioned here demonstrate the value of accounting for potential temporal-perception differences among participants when investigating entrainment to meter and the subsequent impact on syntactic comprehension.

## 4.2.2 Syntactic comprehension

The role of temporal perception on syntactic comprehension is at least twofold: On the one hand, temporal encoding of auditory events is necessary to accurately perceive syntactic events such as musical pitches or linguistic phonemes containing morphological information. On the other hand, the temporal arrangement of events can guide attention such as outlined in DAT paradigms. Behavioral and EEG evidence below suggest a mediating role for these aspects of temporal perception on syntactic comprehension in music (Schmuckler & Boltz, 1994; Tzounopoulos & Kraus, 2009) and language (Schmidt-Kassow & Kotz, 2008; White-Schwoch & Kraus, 2013).

### Music

Syntactic comprehension hinges in part on the perception of individual constituents, which in music listening requires accurate temporal encoding of complex sounds. Regarding auditory brainstem representation of complex sound, improved representation of sub-second-timescale timing, pitch and harmonics has been attributed with musicians (Kraus, Skoe, Parbery-Clark, & Ashley, 2009; see Tzounopoulos & Kraus, 2009 for review). As seen above (Section 4.1.2), musicians have globally better perception of music syntax than non-musicians. Thus, more accurate temporal encoding may be a large factor in better syntactic comprehension. Moreover, improved syntactic comprehension may in part be related to attending auditory musical events with more left-lateralized neural resources that operate at the (rapid) timescale of speech, as several neuroanatomical studies ascribe the left hemisphere to rapid temporal processing (Poeppel, 2003; Tallal, Miller, & Fitch, 1993; Zatorre & Belin, 2001) or show more left-hemisphere engagement in music perception in musicians compared to non-musicians (Bever & Chiarello, 1974; Evers, Dannert, Rödding, Rötter, & Ringelstein, 1999; Ohnishi et al., 2001; Tervaniemi, Sannemann, Noyranen, Salonen, & Pihko, 2011).

On a larger timescale, the temporal arrangement of syntactic constituents (pitches) in a phrase influences syntactic comprehension. In one study, chords accompanying a melody were either periodically fixed to a strong beat or occurring at variable beat locations throughout the measures, and the final chord varied in both harmonic expectancy (high, medium, and low) and temporal location (early, on time, and late; Schmuckler & Boltz, 1994). In two separate experiments, participants were asked to rate the appropriateness of the final chord on a 1-7 scale and to answer a forced-choice belongingness question. The results of both experiments confirmed the role of temporal placement of chords in building of harmonic expectation: when preceding chords were temporally periodic (as opposed to temporally variable), the timing of the final chord modulated the perceived appropriateness of highly expected harmony. Another study showed that temporally jittered sequences caused an increase in error rates among musicians (but not non-musicians) in the dissonance judgment of harmonically unexpected chords (Bigand, Madurell, Tillmann, & Pineau, 1999). The fact that performance was poorer in both studies when the target was temporally unexpected confirms that chord processing depends on the temporal organization of the musical sequence.

It is important to note that, as pointed out elsewhere (Section 4.2.1), these latter two music studies (Bigand et al., 1999; Schmuckler & Boltz, 1994), do not distinguish temporal manipulation from rhythm or meter. The stimuli in the experiments were MIDI files with exact timing, such that temporal cues were synonymous with metrical identity of phrases. Distinguishing temporal from metrical information is often not included in music research paradigms, as MIDI files with perfectly timed periodicity render the only temporal manipulations highly salient to metrical structure. This begs the question, then, whether the perception of harmonic structure is influenced by temporal or metrical characteristics. Music stimuli in the forthcoming experiments (Studies 1 – 3) contain human-performed timing errors in two types of meters, isolating the impact of meter on syntactic comprehension while ecologically engaging general temporal perception skills.

#### Language

Improved temporal perception offers more accurate encoding of rapid speech information, which in turn allows for improved general cognitive processing of that information such as syntactic comprehension. For example, neural tracking of the auditory signal at fine temporal scales is integral in perceiving speech-syllables that carry morphosyntactic information vital to syntactic comprehension (Song, Skoe, Banai, & Kraus, 2011; White-Schwoch & Kraus, 2013). In addition to this, links between temporal perception and syntactic comprehension may be made via studies in healthy and patient populations. For example, improved discrimination of rhythm sequences (with varying temporal intervals) was positively correlated with morphosyntactic competence in children (Gordon et al., 2015), and temporal processing deficits in persons with dyslexia or other specific language impairments (Farmer & Klein, 1995; Overy, Nicolson, Fawcett, & Clarke, 2001; Tallal et al., 1993) may underlie their deficits in syntactic comprehension. Patients with Parkinson's disease have also a deficit in temporal processing and a deficit in syntactic comprehension (Friederici et al., 2003; Pastor, Artieda, Jahanshahi, & Obeso, 1992). These studies point to a significant role of timing in speech perception, which may carry over to syntactic comprehension (for a comprehensive discussion of time and speech perception/production, see Kotz & Schwartze, 2010).

One particular ERP study makes a strong case that temporal arrangement of syntactic elements impacts syntactic comprehension (Schmidt-Kassow & Kotz, 2008). In this study, two factors were explored: syntactic correctness and temporal presentation. The syntax manipulation omitted part of a prepositional phrase, creating a condition that required syntactic reanalysis. The temporal presentation was normal speech, isochronously spaced individual lexical items, or isochronously spaced chunked lexical phrases (such as noun phrases appearing together). It was found that a P600 (see Section 3.1.2) was elicited by the omitted noun. Critically, the P600 occurred not when the omission was apparent, but rather at the onset of the next chunked phrase, indicating that entrainment had taken place in the chunked condition and that the syntactic reanalysis of the phrase was also somehow dependent on the entrainment mechanism.

This latter study also provides insight into the nature of entrainment in language, such that temporal regularity is most beneficial when occurring in phrases and not for individual items. The Schmidt-Kassow and Kotz (2008) study also provides an important segue to the current empirical approach. Indeed the chunking condition in this study offered close-but-not-quite temporal isochrony of strong elements in the meter of the sentence, creating a strong case for stimuli in the current studies to induce entrainment via temporally regular metrical units in speech (much the way meter is organized in naturally performed music).

**Summary**

In both domains, syntactic comprehension may be traced back in part to perception of acoustic events at a sub-second timescale and to global temporal placement of events. The former relationship of temporal encoding to larger-scale syntactic comprehension may be inferred rather from a panoramic view of the literature than from single studies, whereas studies investigating the latter perspective show specifically that syntactic expectation in music is fundamentally tied to temporal expectation, and that when temporal periodicity is introduced as an organizing principle of language-syntax phrase constituents, processing of syntactic deviance conforms to those periodic oscillations. Individual differences in temporal perception may play a role both at the sub-second level and with respect to larger-time-scale temporal organization of syntactic events.

## 4.3 Working memory

The term "working memory" refers to the cognitive system attributed to temporary storage and manipulation of relevant information during the performance of a task (Baddeley & Hitch, 1974; Miyake & Shah, 1999). A prominent model of working memory describes a system comprised of a central executive, a phonological loop, a visuo-spatial sketchpad and more recently an episodic buffer (Baddeley, 2000). The components of the working memory system are shown in Figure 4.2. The central executive focuses attention, dividing and switching attention among tasks as necessary. The visuo-spatial sketchpad stores and processes visual or spatial characteristics of an object, and the phonological loop stores and processes verbal information via an articulatory rehearsal system (comparable to sub-vocal speech). The episodic buffer stores shortened higher-order representations of information and provides a link between specialized working memory subsystems and long-term memory. An overview of the separate systems and their function may be found in Baddeley (2010).

Various psychological tests are associated with measuring working memory. For example, spatial rotation tests access the visuo-spatial sketchpad (Shah & Miyake, 1996), and attention-shifting tasks test the central executive (Collette & der Linden, 2002; Stroop, 1935). As this thesis focuses on the auditory domain, tests tapping the phonological loop are particularly relevant; tasks which access the phonological loop via articulatory rehearsal include digit span (Tewes, 1994), non-word repetition (Mottier, 1951), and modified listening span (Daneman & Carpenter, 1980), the latter of which additionally engages the central executive (see Section 5.2 for a full description of these tasks).

In individual-differences literature, measures of working memory are generally associated with fluid intelligence (Unsworth, Fukuda, Awh, & Vogel, 2014) and reasoning (Engle, Tuholski, Laughlin, & Conway, 1999), and especially pertinent to the current approach, aspects of music and language processing. Below is a sketch from the literature as to how working memory relates to meter perception and syntactic comprehension in music and language, illustrating the benefit of accounting for individual differences in (verbal) working memory in the upcoming experiments.

### 4.3.1 Meter perception

When it comes specifically to meter perception, there is a paucity of literature relating working memory processes to either music or language domains. In music, evidence exists for potentially overlapping neural networks involved in working memory and meter

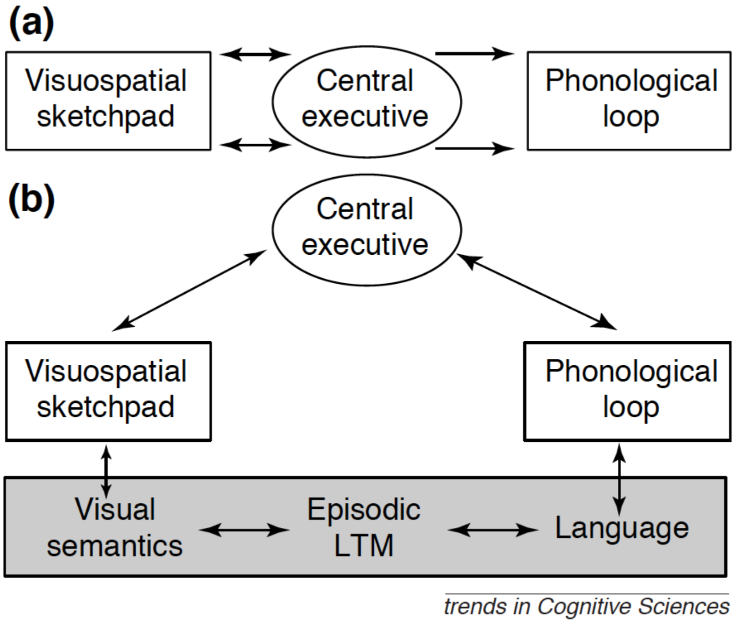



*Figure 4.2: (a) Components of the working memory system according to Baddeley and Hitch (1974). (b) How working memory components interface with cognitive systems according to Baddeley (2000). LTM = long term memory. See text for description. Figure from Baddeley (2000).*

processes (Anvari, Trainor, Woodside, & Levy, 2002; Kuck, Grossbach, Bangert, & Altenmüller, 2003), though these paradigms do not sufficiently isolate musical meter processing from working-memory engagement by their respective tasks. In language, working memory involvement in meter perception may be implicitly inferred from studies showing participants with improved auditory working memory to better encode syllables (Kraus, Strait, & Parbery-Clark, 2012), since syllables contain metrical information (see Section 2.4.2).

### Music

Although the perception of meter has not been directly linked to working memory capacity, evidence from one study suggests that working memory is involved in meter perception when the task requires retention and subsequent comparison of a beat sequence (Kuck et al., 2003). This study recorded EEG slow potentials (time-locked to presentation of entire 12 second trials) while participants (musicians) listened to a standard four-second beat sequence with a metrical structure played on single piano tones, with the task to judge whether the subsequent four-second sequence matched the first or had a deviation in the beat pattern. Deviations were either in the meter (inducing for example a change in time signature) or in rhythm (a substitution of note divisions that did not affect meter). Sustained activity found predominantly over the right temporo-frontal region in both meter- and rhythm-mismatch conditions was interpreted as reflecting “auditory working memory and a pattern recognition module" present in integration of these beat sequences (as opposed to online rapid temporal processing present in other paradigms). Behavioral evidence in children supports this interpretation: rhythm discrimination in five-year-olds correlated positively with working memory capacity (Anvari et al., 2002). The task in this study required the children to compare beat sequences (forced-choice 'same' or 'different'), which is identical to the above process of retention and subsequent comparison of a beat sequence. Though these studies show that meter-related tasks may engage the working memory system, they do not definitively show working memory involvement beyond a same-different task that engages the working memory system.

#### Language

There is little research directly addressing working memory and the perception of specifically language meter. However, language meter and auditory working memory may be indirectly linked by the work of research into auditory perception and musical training: Over many studies, Kraus and colleagues have shown that musicians have improved auditory working memory over non-musicians and that musicians have improved encoding of speech syllables compared to non-musicians (for a comprehensive review see Kraus et al., 2012). Considering that syllable features impact the perception of metrical information (Magne et al., 2007; Schmidt-Kassow & Kotz, 2009b), it follows that musicians, the participants with greater working memory ability, have an advantage to perceive metrical features (indeed, Section 4.1.1 outlined evidence for musical expertise influencing language meter perception). The role of auditory working memory in meter perception mentioned here is indistinguishable from musical expertise; the current series of studies, which involves the perception of language meter, addresses this gap in the literature by collecting working memory measures for all musician and non-musician participants.

#### Summary

There is little research directly addressing working memory and the perception of specifically meter, in either music or language domains. Two studies suggest a link between working memory and musical meter perception, but task demands in those studies involved components of the working memory system, possibly confounding the results. Research as to improved syllable encoding in participants with improved auditory working memory is indirectly related to language meter, since the syllable comprises the lowest level of the metrical grid and carries weight in poetry. The current approach pioneers a direct connection between working memory and meter perception, though the primary motivation for the inclusion of working memory measures in individual assessment is due to its well-established link to syntactic comprehension, addressed in the next section.

### 4.3.2 Syntactic comprehension

The relationship of working memory to syntactic comprehension is an emerging focus in the music domain and thoroughly explored in the language domain. In music, behavioral evidence links syntactic comprehension to working memory (Fiveash & Pammer, 2014), and further indications of this relationship may be found fMRI evidence (Oechslin et al., 2013). In language, ample behavioral, functional and anatomical evidence exists for links between working memory and syntactic comprehension (e.g. Fiebach, Schlesewsky, & Friederici, 2001; Makuuchi & Friederici, 2013; Meyer, 2012). Below, focus is taken to discuss the role of working memory in the specific type of language syntactic comprehension used in this thesis, syntactic ambiguity processing.

#### Music

A recent behavioral study reported evidence that music syntactic comprehension relies on working memory (Fiveash & Pammer, 2014). Participants recalled sentences that had been visually presented during melodies with normal syntax, normal syntax and a deviant-instrument chord, or violated syntax. Sentence recall was detrimentally affected by melodies with syntactic error whereas neither normal melodies nor melodies with a deviant instrument impacted recall. This evidence confirms that syntactic manipulations in music tax the working memory system.

The potential for music syntactic comprehension to interact or be dependent upon working memory is present in other behavioral and fMRI evidence. Behavioral evidence in children has shown positive correlations between measures of working memory capacity (digit span) and same-different discrimination of chords (pitch changes) and melodies (pitch/contour changes; Anvari et al., 2002), which are pre-cursors to detection of musical syntactic elements. Also, a recent fMRI study supported the notion that musical syntax detection and working memory recruit common resources (Oechslin et al., 2013). This study found that the areas of activation in response to harmonic-violation detection coincided with fronto-temporal networks commonly associated with working memory and attention processes.

Musicians have been reported to have, as a group, higher verbal working memory capacity than non-musicians (Chan, Ho, & Cheung, 1998; Ho, Cheung, & Chan, 2003). Though not explored specifically in music syntax studies, the fact that working memory has generally been attributed to being higher in musicians may be a factor in advanced syntactic comprehension in musicians, particularly in later integration processes (see below; working memory correlates with language syntax integration processes). The higher amplitude reported in musicians for the P600 (Besson & Faïta, 1995) is symmetrical to the higher-amplitude P600 in language-syntactic comprehension for groups with high working memory capacity (Bornkessel, Fiebach, & Friederici, 2004).

**Language**

In the working memory model, the phonological loop is inherently connected to language processing; working memory has a long-established link to studies of language comprehension (e.g. Baddeley & Wilson, 1988; Daneman & Carpenter, 1980). Of relevance here is the processing of ambiguous syntactic structures, which have a non-preferred (but not grammatically incorrect) resolution to a local ambiguity. Specifically, the difference in comprehension between participants with high and low working memory capacity has been linked to the ability of those with high working memory capacity to better inhibit information, particularly information that would place non-preferred readings of an ambiguous lexical item or syntactic structure in competition with preferred meanings or structures (Bornkessel et al., 2004; Gunter, Wagner, & Friederici, 2003; Meyer, Obleser, & Friederici, 2013).

In this context, participants with a lower working memory (measured by reading span) have been consistently shown to be at a disadvantage to participants with a greater working memory capacity. Examples may be found from the ERP literature, in studies that compared the ERPs elicited from two types of relative clauses, subject-first vs. object-first. In German, the relative pronoun ‘who’ is expected by the reader or listener to be the subject of its following relative clause, and although the alternative—when ‘who’ is the object and not the subject—is not grammatically incorrect, this is a more complex syntactic structure and requires reanalysis from the reader or listener (Section 3.1.2 gives an example of an ambiguous German relative pronoun). This preference for subject-first word order is attributed to the load in working memory that is greater when the syntactic complexity increases (Schlewesky, Fanselow, Kliegl, & Krems, 2000). In one study, high-span readers demonstrated the ability to integrate an object-relative disambiguation that appeared long after the ambiguous relative pronoun, whereas low-span readers demonstrated this ability only when disambiguating information was immediately available (Friederici et al., 1998). In a similar study which manipulated the distance between the ambiguous relative pronoun and its subject/object disambiguation, low-span participants demonstrated a delayed ERP response (a late positivity) to a syntactic ambiguity when the sentence contained a large amount of information required by working memory processes, whereas the ERP effect found in high-span participants (an early positivity) displayed the same latency regardless of the working memory load (Vos et al., 2001). The amplitude of ERP effects consistently elicited by ambiguous pronoun resolution tends to correlate positively with reading span scores (Bornkessel et al., 2004), also cross-linguistically in languages such as Dutch (Nieuwland & Berkum, 2006). Thus working memory plays an integral role in the comprehension of ambiguous language syntax structures.

**Summary**

In music, behavioral and fMRI evidence suggest functional links between working memory and syntactic comprehension; this dissertation pioneers an approach that compares measures of working memory with electrophysiological response to music syntax manipulation. Measures of working memory have an established relationship to language syntax, including ERP response to structural ambiguity comprehension. Particularly, higher working memory is thought to better inhibit information that would potentially encumber resolution to an ambiguity, and is indexed by higher ERP amplitudes.

## 4.4 Open Issues

As touched-upon in respective sections of this chapter, individual differences in musical expertise, temporal perception and working memory potentially overlap as explanations behind trends of cross-domain meter perception and syntactic comprehension. Musicians are better perceivers of meter and syntax than non-musicians, a fact which has been linked to richer mental representations stemming from a music-specific knowledge base, however, the brain's precise, sub-second temporal encoding of pitches and syllables is critically responsible for comprehension of small syntactic and metrical components. The relative advantages of musical expertise versus basic temporal perception are not clear, particularly when it comes to later-stage processing, hence the current approach that indexes musical skill as well as temporal perception among all participants. Similarly, processing advantages offered by musical expertise and better working memory may in some instances bleed together: Musicians are reported to generally possess higher working memory capacity than non-musicians, and music-syntax-integration-ERP group differences among musicians and non-musicians are analogous to language-syntax-integration-ERP group differences among participants with high- versus low- working memory (namely, that ERP amplitudes are higher for musicians and high-working-memory groups). Thus the current working memory indices alongside indices of musical skill and temporal perception will help to elucidate otherwise indiscernible underlying perceptual correlates to tasks in the current experiments (Studies 1 – 3).

## 4.5 Current approach

This chapter has outlined the relevance of individual differences in musical expertise, temporal perception and working memory as they pertain to meter perception and syntactic comprehension in music and language. To best account for individual variability in these areas, this thesis employs a compilation of seven diagnostic tests, known to measure cognitive ability in music perception (the Musical Ear Test, rhythm and melody subtests; Wallentin et al., 2010, temporal perception (anisochrony detection test; Dalla Bella et al., 2016), and working memory (forward and backward digit span, modified listening span and non-word repetition; respectively Tewes, 1994, Daneman & Carpenter, 1980, and Mottier, 1951). The scores, collected from all participants across three studies, will be evaluated with a factor analysis in Experiment 1a to first find whether the diagnostic scores can be clustered together into factor scores, i.e., to see if diagnostic measures overlap in the cognitive ability they portray. Any resulting factor scores can be correlated with experimental measures in subsequent experiments, thus accounting for individual variability in participants across the whole of the dissertation.

## 4.6 Summary

This chapter has shown how musical expertise, temporal perception and working memory empirically impact studies of meter perception and syntactic comprehension in music and language domains: Musical expertise offers improved pre-attentive auditory processing, which heightens cross-domain detection of metrical and syntactic deviances, and a theoretical knowledge base which allows richer mental representations of music meter and –syntax. Temporal perception is entwined with the perception of meter and syntax across both domains, as sub-second timing impacts actual event perception, metrical structure impacts timing discrimination thresholds, and temporal arrangement of discrete events impacts syntactic comprehension across domains. Working memory is subtly related to music meter processing and this relationship may be inferred in language meter; emerging literature links working memory to music syntax while working memory links to language syntax are well established, most relevant here is that increased working memory capacity aids in resolution of language syntactic ambiguities. The goal in presenting this evidence was to illustrate the advantage of accounting for individual differences among these three traits.

This dissertation aims to investigate whether participants can entrain to a beat in metrically hierarchical melodies and sentences, and if so, index the impact that entrainment would have on the comprehension of syntactic complexity in those melodies and sentences. Monitoring individual differences in temporal processing, musical ability and working memory is crucial to obtaining a complete empirical picture. The next chapter of this dissertation details the methods used for such investigation, including the diagnostic tests employed to account for individual differences and the mechanism by which an EEG recording can capture neural activity relevant to meter perception and syntactic comprehension.

# Chapter 5

# General Methods

The current investigation is concerned with whether the healthy brain will entrain to language or music as a function of its respective metrical structure, and whether that entrainment (if found) can facilitate syntactic processing across domains. Measuring the frequencies present in EEG oscillations—and comparing them to the frequencies present in the stimuli—is a robust approach to observe entrainment; if frequencies coinciding with different levels of the metrical hierarchy are found in the stimuli and brain, one may infer entrainment to the stimuli's meter. Event-related potentials are a well-founded method to observe online comprehension-related processes, and ideal to identify differences in syntactic processing due to a metrical context. In addition to measuring brain waves, an important aspect of this investigation is to address natural cognitive variability in the healthy population and how those differences might contribute to entrainment and syntactic comprehension.

This chapter contains methodological aspects of EEG relevant to frequency and ERP analysis, motivation and description of various diagnostic tests that index individual differences, and finally a description of the experimental paradigm used for the subsequent three studies.

## 5.1 EEG

### 5.1.1 Theoretical overview

#### Biological basis

The signal measured at the scalp by EEG, whose invention is credited to Berger (1929), stems from the summation of postsynaptic potentials generated in assemblies of thousands of similarly oriented cortical pyramidal cells, whose apical dendrites are positioned perpendicular to the surface of the cortex. Per pyramidal cell, the postsynaptic potentials are a result of ions moving into or out of the cell when neurotransmitters bind with receptors on the postsynaptic membrane; these cells need to be similarly oriented so that the electrical potentials sum together as opposed to canceling each other out. The orientation and distance of the cell assembly from respective electrodes on the scalp greatly affect the strength of the signal, because all of the fluids, tissue and bone in between serve as impedance and distort the signal. A detailed description of these processes may be found in Luck (2005, Chapter 1) and Kandel, Schwartz, and Jessell (1991).

#### Oscillations

The communication of neurons is nonlinear and dynamic, and information flow involves synchronized cell activation within and among various functional networks. Human EEG oscillations are categorized into five frequency bands named delta (< 4Hz), theta (4–8Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–80 Hz) (Buszáki, 2006); electrical oscillations in various frequency bands presumably reflect the coupling and uncoupling of communicating cell assemblies (Singer, 2013). High-frequency oscillations may represent communication of neighboring cells within assemblies or nodes and low-frequency oscillations may represent communication across nodes which are separated by some distance within the brain. In terms of data analysis, power, or squared amplitude, is thought to reflect the synchronous activity of proximally located cells, while phase coherence is the stability of the phase (for an introductory description of power and coherence with respect to functional networks, see Bastiaansen, Mazaheri, and Jensen 2012). Specific cognitive functions of networks also may be characterized by increased power in the EEG of certain frequency-bands (for example memory storage and gamma frequency range, or suppression of task-irrelevant information and alpha). Experimental settings can induce changes in oscillatory brain activity, for example Berger (1929) famously noted that amplitude of alpha waves was prominent in relaxed states but then decreased when participants opened their eyes or executed a cognitive task, whereupon oscillations in the beta range were observed. The present studies focus on oscillations at frequencies induced by temporal qualities of stimuli, and frequency-power analysis of stimulus-induced oscillations is described below in Section 5.1.4.

#### Event-related potentials

Examination in the temporal domain of what are called event-related potentials (ERPs) shows the instantaneous voltage changes derived from synaptic activity, which presumably reflect sensory, cognitive, affective, and motor processes elicited within tens or hundreds of milliseconds by an experimental event. Various positive and negative deflections in the volt-

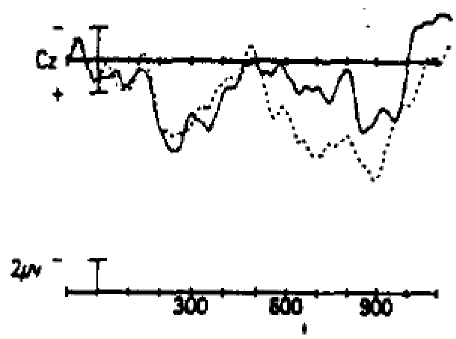



*Figure 5.1: An example of the P600 effect, where the P600 component elicited by a preferred syntactic structure (solid line) has significantly less amplitude than the P600 component elicited by a non-preferred syntactic structure (dotted line). Figure from Osterhout and Holcomb (1992).*

age changes are called components and are associated with specific neural/psychological processes (Kappenman & Luck, 2012). Event-related potential components are commonly labeled with respect to their polarity, latency and/or scalp distribution, for example the name P600 (introduced in sections 3.1.1 and 3.1.2, see Figure 5.1) refers to a positive deflection appearing ca. 600 ms after stimulus presentation. Experimentally relevant ERP components are quite small in comparison to irrelevant time-locked EEG noise, thus many trials per condition are averaged to separate stimulus-response effects from other brain activity (Handy, 2005). When components arising in different experimental conditions are statistically different from each other (typically, in amplitude), this is referred to as an effect (e.g., a P600 effect occurs when the amplitude of the P600 component is significantly greater in one condition). Analysis of ERPs is described below in Section 5.1.5.

#### Poor spatial resolution

EEG reflects instantaneous voltage changes associated with brain activity, and has optimal temporal resolution, however the spatial resolution of the signal is poor; localization of the signal-source is unclear, thus mathematical models are required to predict signal origins. This problem is compounded in ERPs, as components may come from multiple sources simultaneously (see Kandel et al., 1991; Kappenman & Luck, 2012). While EEG is the preferred method to answer the present research questions with respect to entrainment and syntactic integration, further studies incorporating questions of anatomical origin would require methods with more spatial resolution such as magnetoencephalography or functional magnetic resonance imaging.

### 5.1.2 Recording

The voltage potentials that reach the scalp are often recorded with Ag/AgCl electrodes, which typically require preparation of the underlying skin with gel to improve conductance. The recording is of a difference between the potentials in two locations: a bi-polar setup references each active electrode to another active electrode whereas a unipolar setup (used here) references each active electrode to a single remote inactive electrode (e.g., over the right mastoid). A ground electrode (e.g. at the forehead or sternum) helps in the measurement of the difference potentials by subtracting voltages which show at both active and reference points. The recorded signal must be amplified and digitized, and is collected at a sampling rate defined by recording equipment parameters (e.g. 500 Hz). Other facial electrodes may be placed to record artifacts from the eyes or for later re-referencing to the nose. See Teplan (2002) for a detailed description of these EEG measurement steps.

The active scalp electrodes, often sewn into a cloth cap, must be arranged in some order; one common arrangement is the International 10-20 system, a precedent established by the International Federation of Societies for Electroencephalography and Clinical Neurophysiology, which places electrodes at 10% and 20% intervals away from bony landmarks (Reilly, 2005; Sharbrough et al., 1991). In this system electrodes are labeled to indicate left- and right hemisphere orientation (odd numbers: left hemisphere, even numbers: right hemisphere, z: midline) as well as anterior/posterior orientation loosely referring to underlying brain regions (letters, F: frontal, C: central, T: temporal, P: parietal, O: occipital).

### 5.1.3 Preprocessing

**Automated and manual artifact rejection** One cannot help but record spurious signals in an EEG dataset, whether they be the unconscious muscle twitches of a participant or lurches in signal due to starting, stopping, or resetting of recording equipment, etc. Noise causing large amplitude deviation in the recorded potentials can be automatically detected and removed with most EEG software, by defining some noise threshold within a limited search window.

In addition to automatic rejection, manual inspection of the signal allows the experimenter to fine-tune the rejection of data segments, for example capturing clear artifacts missed by the automatic rejection or reclassifying rejected material (for example, perhaps some rejected artifacts are systematic, such as a general muscle tension or blinks) that is best removed in one of the subsequent steps mentioned below.

**Filtering** Filtering is one way to remove artifacts that systematically occur in a limited frequency band. In the filtering process, a time-series signal is transformed to the frequency domain, the amplitude of the signal is attenuated at specified frequencies, and then the signal is transformed back to the time domain with its newly-altered frequency information. (The Fourier transformation process is summarized in 5.1.4 below). High-pass filters attenuate the amplitude at frequencies below a cutoff and let higher frequencies pass; accordingly low-pass filters attenuate the amplitude at frequencies above the cutoff and let lower frequencies pass. Band-pass filters have two cutoff frequencies, letting frequencies pass between them while attenuating frequencies outside the band, and stop-band filters attenuate only a specific set or band of frequencies. The filter attenuates frequency amplitude along a slope which crosses the cutoff frequency, the steepness of which is determined by filter parameters (namely, a higher ‘order’ specification yields a sharper cutoff). The two most common classes of filters are finite impulse response (FIR) and infinite impulse response (IIR), which have respectively shallower and steeper cutoff slopes for the same order. Steep slopes create more distortion artifacts or ringing in the signal than shallow slopes, though some distortion can never be avoided. Other types of artifacts created by filters include a delay of the signal in the time domain as well as a shift in the signal’s phase, both which may be corrected in further processing steps. See Lyons (2004, chapter 5, chapter 6) for a technical description of filtering.

When EEG datasets contain systematic artifacts among many participants in consistent frequency bands, the benefits of filter use can outweigh distortion effects by ‘cleaning’ the signal and improving the signal-to-noise ratio. For example, low-frequency drifts may be cleaned with a high-pass filter or excessive muscle tension (which occurs at higher frequencies) may be cleaned with a low-pass filter—both of these types of artifacts can pervade a participant’s whole dataset when trials are otherwise usable. See Widmann, Schröger, and Maess (2015) for a discussion of filters and their proper usage in electrophysiology.

**Independent component analysis** Some types of artifacts are successfully removed by independent component analysis (ICA), for example blinks, noise from a particular electrode, or the presence of a heartbeat. The ICA method decomposes the EEG scalp data, using an ‘unmixing’ matrix to find maximally temporally independent individual components which comprise the data, including their probable electrode locations. Drifts and extremely large artifacts must be removed before performing ICA in order to assure correct identification of independent components. Once one views the signal and scalp topography of each found component, artifacts (such as eye blinks) may be identified and removed, and then a ‘mixing’ matrix recomposes the EEG scalp data without the removed artifacts. See Makeig and Onton (2012) for EEG-related discussion of ICA and Naik and Kumar (2011) for a description of the computational steps.

**Electrode interpolation** When individual electrodes distort a dataset with noise, their removal (for example during the ICA process) improves the signal-to-noise ratio. In order to keep data input statistically consistent across participants, the removed electrode may be interpolated, meaning that the removed signal is reconstructed by averaging the weighted signal of neighboring electrodes.

### 5.1.4 Spectral analysis

Spectral analysis, a process wherein a time-series signal is decomposed to reveal the magnitude of the signal at frequencies along a spectrum, is essential to the present investigation: entrainment of relevant neuronal populations to a sensory signal may be logically inferred when the amplitude of (hypothesized-for) frequencies peak significantly in both the stimulus and EEG frequency-amplitude spectrum. The following paragraphs explain the basic steps in spectral analysis relevant to this investigation.

**Fourier transform** The discrete Fourier transform is an important concept for the processing of digital signals such as audio files or EEG. The 'transformation' is of a number of samples from a discrete signal from time to frequency domain (or vice versa), made by correlating each sample input point with many sinusoidal waves and summing the term-by-term products. The exact frequencies of the sinusoids depend on both the number of sample input points and the sampling rate ($fs$) of the original signal. The output (from time to frequency domain) is the magnitude (amplitude) and phase angle of the summed sample input points at frequency bins along a spectrum. The bins occur at frequency intervals determined by the fundamental frequency (sampling rate divided by the number of sample input points): other frequencies present in the input signal which are not multiples of the transform's fundamental frequency are not represented. Zero padding, or appending zeros to the end of the time-series data, is a convenient way to increase the amount of frequency output bins (minimize the fundamental frequency). Usually, the fast Fourier transform (FFT) is used in software packages because it shortcuts the computation of the full discrete Fourier transform. See Lyons (2004, chapter 3, chapter 4) for a comprehensive explanation of the Fourier transform.

When an FFT is taken of a signal whose times series sinusoidal waves do not have a completed cycle within the input sample (i.e., the amplitude at the boundaries of the sample is not zero), spectral leakage occurs due to properties of the Fourier equation (the sum of the products of the input sequence and sinusoidal wave correlations is no longer equal to zero). Spectral leakage causes the amplitude value at frequencies other than multiples of the fundamental frequency to 'leak' onto the amplitude values in the actual output bins.

A real-world time-series signal (e.g., an EEG segment time-locked to an event, or epoch) has amplitude at frequencies besides those frequencies which are multiples of the fundamental frequency, and is not likely to have perfect completion of all sinusoidal wave cycles present in the sample. One way to contain the resulting spectral leakage is to use a window, which is a function whose shape, when multiplied by the time-series signal, reduces the amplitude of a time-series signal near its sample boundaries. Application of a windowing function reduces but does not eliminate spectral leakage. Common windows employed are the Hanning and Hamming windows, which are shaped respectively like an elongated cosine arch and that same cosine arch raised on a pedestal.

**Detrending** If a trend in the spectrum (or any other kind of dataset) is noticed which is unrelated to the experiment, the data may be detrended, i.e., the slope of the trend may be subtracted from all conditions and analysis performed on the resulting data. In the spectrum of an EEG, physiological noise in lower frequencies can cause a negative trend in the data which may obscure experiment-related amplitude effects.

**Peak analysis** Once a spectral analysis of a signal has been taken, there must be some criteria for determining which peaks are meaningful to the analysis. An initial helpful step is to standardize the power (amplitude squared) data, resulting in a mean power of zero and standard deviation of one, to eliminate irrelevant differences in power baseline across participants or trials. Then (per-trial, per-participant) averaged (standardized, detrended) spectra in each condition may be evaluated for peaks, ideally in a hypothesized location of the spectrum, and then any found peaks tested for significance. Criteria for peak significance may vary across paradigms: One such method is described by Aiken and Picton (2008), wherein background frequency power is defined as an average of eight frequency bins above and eight frequency bins below the peaks, and for a targeted peak to be considered significant, its power must be significantly greater than the power of the background frequencies, e.g. with a related-samples test. In this method, the resolution (which can be achieved with zero-padding) must ensure that the background frequency power of the lowest-frequency peak remains well above the distortion-range bounded by a high-pass filter (or below the distortion range of a low-pass filter), and that background frequency power of any neighboring hypothesized frequencies does not overlap. An example of this method is depicted in Figure 5.2.

#### Excursion: Audio file analysis

Specific frequency information present in stimuli will help to ascertain whether participants entrained to the meter in melodies and sentences, depending on whether the same meter

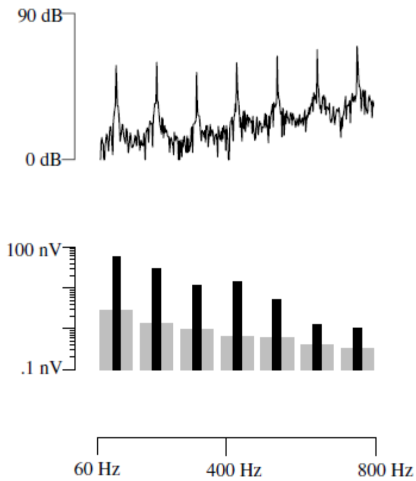



*Figure 5.2: Peak analysis of EEG (bottom) in response to stimuli (top). Significance tests compare amplitude at a target peak frequency (black) to averaged amplitude of background frequencies (gray). Peak analysis of EEG (bottom) in response to stimuli (top). Significance tests compare amplitude at a target peak frequency (black) to averaged amplitude of background frequencies (gray). (Adapted from Aiken & Picton, 2008).*

frequencies are also found in EEG. The following two paragraphs describe steps in spectral analysis specific to the type of audio file used in stimulus presentation.

**Resampling** When comparing the frequency spectrum of two qualitatively different signals, e.g. audio files and EEG, optimally the two spectra should have the same fundamental frequency. This requires a resampling of at least one of the signals. Audio recordings usually have sampling rates on the order of tens of thousands (e.g. 44,100 Hz), compared to EEG sampling rates on the order of hundreds or thousands (e.g. 500 Hz), thus the audio signal must undergo resampling. This process includes several considerations. First, since resampling can only occur in even multiples of the original and new sampling rates, the original signal must be upsampled to a multiple of the desired final sampling rate (e.g., upsampling 44,100 Hz signal by a factor of 5 to achieve 220,500 Hz sampling rate). Second, before downsampling, a filter must be applied to avoid aliasing. Aliasing is the false identification of power at frequencies above the Nyquist frequency, or one-half of the sampling frequency, as power belonging to lower-frequency components. A rule of thumb is to filter at 1/3 the desired new sampling rate (e.g., 160 Hz low-pass filter for a desired 500 Hz sampling rate). Finally the signal may be downsampled by an integer factor to reach a desired sampling rate (e.g., by a factor of 441 to reach 500 Hz from 220,500 Hz).

**Envelope extraction** Speech and music are complex auditory signals whose exact frequency composition is not relevant to the experiments included here. Instead, the envelope or outline of the signal is much more useful to analyze, which contains the information regarding where in time notes and syllables occurred—essential information to determining what frequencies may manifest metrical levels in the signal. The envelope may be detected by a Hilbert Transform of the signal (for a description of the Hilbert transform, see chapter 9 in Lyons, 2004). The envelope is then treated as the input for spectral analysis.

### 5.1.5 ERP analysis

The field of psycholinguistics has used ERP analysis to study various facets of syntactic processes for many decades, in both language and music domains. Such richly nuanced literature (see Chapter 3) makes the use of ERPs ideal in the current paradigm to see whether entrainment based on metrical context influences syntactic processing across domains. The specific processing steps in ERP analysis conducted in the present investigation are described below.

**Baseline correction** After selecting trial epochs with a zero-time-point or trigger ('time-lock') at critical points in stimulus presentation, whose length includes the typical latency of all hypothesized components, establishing a baseline window is a standard way to set a reference amplitude value for measuring the hypothesized ERP components. Baseline correction normalizes the amplitude of potentials in the compared conditions to each other such that the mean in the window is equal to 0 microvolts. A longer baseline window (e.g. 200 ms) avoids residual component amplitude differences based on small fluctuations in noise structure (see Handy, 2005).

**Averaging** ERP components of interest, as mentioned above, are small relative to noise in the signal and therefore an average of many trials (e.g., 40) per condition is necessary to 'reveal' the components in the signal (see Handy 2005, for a thorough discussion of the ERP averaging process). 'Single-subject average' refers to a single participant's averaged trials in a particular condition, and 'grand average' refers to the average of single-subject averages. Here, grand averages are used to quantify effects and single-subject averages are used for correlations with other measures (e.g., scores representing various cognitive factors).

**Effect time windows** Effects are defined as the significant difference between (components in) two conditions (see 5.1.1), thus EEG time-series data needs to be bounded and its values averaged in some time window which best represents the effect. When experimental parameters closely mimic previous experiments that found the hypothesized-for effects, effect windows can be defined a priori, meaning that a window is already part of the hypothesis. Another method more suitable for experiments with novel parameters is to test the difference between two conditions in 'running' blocks of small successive time windows (e.g. 50 ms), statistically comparing the two conditions in each window at each electrode with multiple-comparison correction. An effective version of this method is a cluster-based permutation analysis. The cluster-based permutation approach conducts the running statistical test (e.g, a *t*-test) with a correction for both multiple time windows and multiple electrodes (e.g., Monte Carlo), and determines the windows where effects are significant as well as which electrodes contain the significant effect (see Maris & Oostenveld, 2007 for further reading).

## 5.2 Diagnostic tests

As addressed in the introduction, individual differences in cognitive ability play a role in meter and syntax perception in music and language domains (see Chapter 4). With this in mind, diagnostic tests whose scores represent aspects of working memory, musical and linguistic processing, and temporal perception were included in the empirical approach. Below is a description of the individual tests used here, what they are assumed to measure, and the motivation to include them in the present studies.

### 5.2.1 General procedure

Computerized versions of the all diagnostic tests were auditorily administered to participants at a comfortable volume via headphones. Tests which contained words used recordings spoken by respective professionally trained, female native German speakers. Instructions were either written or spoken aloud by the test administrator. Correct-hand response was counterbalanced across participants (but kept consistent for tests within-participant), and the order of diagnostic test presentation was pseudo-randomized.

### 5.2.2 Digit span

**Cognitive function** Digit span is widely accepted as a measure of working memory capacity. The digit span task, described in detail below, involves a participant repeating a series of digits in either forwards or backwards fashion; according to the Baddeley and Hitch working memory model (1974, see Section 4.3) the cognitive process involved in the digit span task contains a strong verbal component due to the vocalization of responses and thus activation of the phonological loop. The mental reversing of the presented order in backward span involves additional cognitive processes that may be linked to visuospatial abilities or general intelligence, and a volume of literature has shown different behavior of the two measurements when statistically compared with third cognitive measures, thus forward and backward digit span scores are worth recording and analyzing separately (for a review see Ramsay & Reynolds, 1995). Here, the tasks are included to obtain measures of general working memory capacity among participants.

**Testing procedure** Forward and backward auditory digit span were used from the Wechlser Intelligence Test (Tewes, 1994). For forward digit span, participants listened to a series of digits, starting with three digits, and had to repeat the digits back in the order in which they were presented. The test continued up to nine consecutive digits, each 'level' of digits consisting of two trials, one of which participants must correctly repeat in order to proceed to the next level. The backward digit span differed in that the digit series ranged from two to eight total digits, and that participants were instructed to repeat the digits in reverse-order of presentation. The final score in each task was based on the highest level that a participant completed.

### 5.2.3 Musical Ear Test

**Cognitive function** This test was designed to index perceptual musical abilities, separately in melody and rhythm subtests. The melody subtest focused on identifying contour and pitch changes and the rhythm subtest employed primarily beat detection and interval discrimination. The original publication boasts the ability to distinguish musicianship skills among amateur and professional musicians; among musicians the scores were correlated with participants' practice time and pitch- and rhythm reproduction task scores (Wallentin et al., 2010). This test was included to index melody discrimination, necessary for Studies 1 and 3, and rhythm aptitude, which pervades all studies as a potential factor in entrainment.

**Testing procedure** The rhythm and melody subtests each consisted of 52 trials. Participants were instructed to listen to melody or rhythm pairs and determine whether the two are the same ('yes' or 'no' response button). Half of the trials were 'yes' and 'no' correct-response, presented in randomized order throughout the subtests. Participants had one second to respond. Melodies were midi-piano files and ranged in musical complexity; difficulty ranged from easy (three long notes on a diatonic scale, where one note in the second melody was a large interval different from the first melody) to extremely difficult (fast-paced atonal sequences, where a syncopated note is a half-tone different in the second melody). Authors reported the melody test to specifically address pitch and contour discrimination. The rhythms were beat sequences presented with midi wood-block sounds; difficulty again spanned from easy sequences (rhythms had few beats with large temporal interval differences between the sequences) to extremely hard (fast, syncopated rhythms with as little as 50 ms difference in temporal intervals between series). Authors reported subdivision of beats into triplets to be a key feature in difficulty. A faint metronome was played in the background of all melody and rhythm trials, which established the downbeat. Two scores—the total correct answers per subtest, up to 52—were assigned per participant.

### 5.2.4 Non-word repetition

**Cognitive function** Non-word repetition, first published as a task by Mottier (1951), measures phonological working memory capacity and the ability to discriminate phonemes. This test was included as a diagnostic tool for this thesis as it has particular relevance to Study 1, where the language task requires discrimination of subtle phoneme differences between singular and plural verb conjugation.

**Testing procedure** This test presented pseudo-words, which participants were instructed to repeat. The pseudo-words, each with a consonant-vowel structure, e.g., 're-la', were presented in blocks of six, starting with two syllables, and each subsequent block contained a set of pseudo-words with one-higher syllable count. Participants continued until no longer able to perform the task (a rare participant reached the eight-syllable block). A block was considered completed if over half of the pseudo-words were repeated correctly (at least four of six). The final score per participant was assigned based on highest level block that he/she completed.

### 5.2.5 Modified listening span

**Cognitive function** The modified listening span was designed to index the trade-off between storage and processing in the working memory system, specifically as it pertains to language comprehension (Daneman & Carpenter, 1980), and correlates with test scores of reading comprehension. Here, the modified listening span was particularly relevant to Study 3, where participants were instructed to answer questions about sentences with a syntactic ambiguity.

**Testing procedure** In this task, adapted to German from the English version by Daneman and Carpenter (1980), participants heard a series of sentences and were instructed to repeat the last word of each sentence, in the order of presentation, once a cross appeared on the screen after the last sentence. In addition, participants were instructed to evaluate each sentence for global accuracy (e.g. 'All coats are brown' is a globally false statement) and answer 'true' or 'false' with a timed button response after each sentence. After a training with three sentences, the task began with three blocks of three sentences, and continued up to three blocks of seven sentences or until the participant could no longer perform the task. A level was considered complete if two of the three blocks were answered correctly, and the final score per participant was determined by the highest level he/she completed.

### 5.2.6 Anisochrony detection

**Cognitive function** This test estimates participants' temporal discrimination threshold. Attention to incoming information is linked to temporal predictability of that information, as modeled by the DAT (see Section 2.1.1). Thus temporal discrimination of the listener may impact absorption of information in a constant auditory stream such as the sentences and melodies heard by participants in the three studies here, and any differences among participants may account for inter-individual variability in syntax perception or entrainment to meter.

**Testing procedure** The anisochrony detection test was taken from the Battery for the Assessment of Auditory and Sensorimotor Timing Abilities (BAASTA, Dalla Bella et al., 2016). Participants listened to a series of five 1047 Hz (C6) tones (duration = 150 ms), the first three of which had an interval of 600 ms (onset to onset), the fourth of which was either displaced (earlier) or not. The fifth tone was always 1200 ms after the third; any local shift of the fourth tone resulted in both a shortened interval before the fourth tone and a lengthened interval afterward. Participants were asked to answer yes or no, whether the series was regular (i.e. no displacement of the fourth tone). The displacement was calculated by the maximum likelihood procedure, which used a psychometric curve to estimate the participant's discrimination threshold. The algorithm assigned the fourth tone an interval estimated either above or below the discrimination threshold, and then based on whether the participant answered correctly or not, adjusted the interval to be more difficult (after a correct answer) or easier (after a false answer). The maximum likelihood procedure presented the next stimulus interval at a discrimination threshold at some midpoint of the previous outer-boundary estimates, arriving at the participant's estimated individual threshold after sixteen trials.

The sixteen-trial blocks were discarded if one of the two following conditions were met: (1) more than 30% of catch trials (randomly inserted 'regular' trials where the fourth tone had no interval-deviation from the others) were answered incorrectly (perceived as though the fourth tone had a different interval), or (2) the threshold estimate changed by more than 15% over the last 10 (non-catch) trials, meaning the algorithm was not converging to an appropriate mid-point threshold estimate (for example, if a participant answered a few trials mistakenly, which would 'throw off' the convergence of the estimate).

Per participant, the sixteen-trial threshold calculation was repeated three times, and finally an average of the three values was used as a final threshold value. If any blocks were eliminated due to the above reasons, an average of the remaining two blocks or the single remaining value was used for further analysis.

## 5.3 Paradigm

The aims of the current investigation are to determine whether participants entrain to meter in both music and language, whether entrainment to meter (if any) could affect syntactic processing across domains, and to seek out what role individual differences might play in the process. Such aims require stimuli with parameters strictly nuanced to manifest as similarly as possible in music and language domains, and a multi-faceted empirical approach. The cross-domain-comparison parameters, stimuli items, experimental designs across the three studies, and finally general hypotheses are described below. Terminology introduced here will serve all three empirical studies.

### 5.3.1 Cross-domain comparisons

#### Meter

Primary considerations for metrical comparability in music and language were firstly an establishment of hierarchy in language, and secondly a dissociation between meter and timing in music (see Sections 4.2.1 and 4.2.2). Regarding the first consideration, temporal predictability created by nested metrical levels enhances entrainment, as shown in numerous music- or beat-based paradigms (e.g. Jones, 2008; Jones, Moynihan, MacKenzie, & Puente, 2002): a hierarchical meter was therefore necessary in language in order to best compare entrainment across domains (or put another way, the sentence-meter necessarily required poetic or oratory form in a kind of imposed 'musical' organization of beats). Pertaining to the second consideration, the meter in melodies required independence from deviations in timing to allow for optimal comparison with natural speech, thus recordings of human performance were used; otherwise, error-correction of natural non-periodic speech timing could constrain entrainment to language meter but never be taxed in periodic note presentation (such as computer generated MIDI files). The recording and processing of items is described below in Section 5.3.2.

*Hierarchy* is the name designated here to refer to different metrical-level conditions in factorial design of experiments. In stimuli, the *individual* level is the lowest metrical level and represents the beat contained in all eighth notes and syllables. The *group* is the next-highest metrical level, which is bounded by naturally accented notes (based on time-

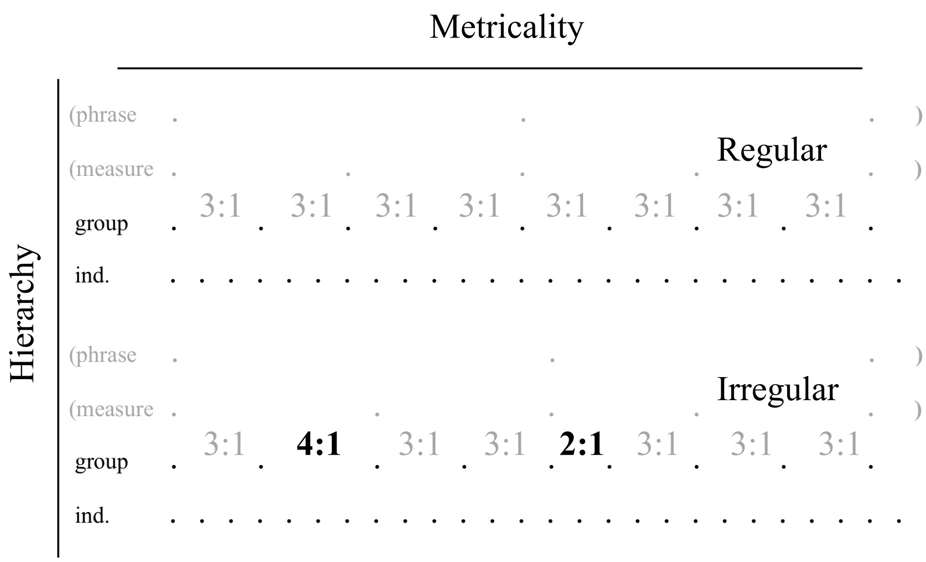



*Figure 5.3: Cross-domain meter comparison: Alternating nested beat groups.* ***Hierarchy****: Dots represent all notes and syllables at the individual level and only weighted notes and syllables at the group level. Higher metrical levels exist (e.g., measure or phrase) but are outside the scope of study.* ***Metricality****: Regular is comprised of even groups of three beats (3:1 individual to group ratio, in gray) while irregular displaces group boundaries to create alternating groups of two, three, and four beats (4:1 and 2:1 individual to group ratio, in bold).*

signature) or naturally weighted syllables, and is comprised of either two, three or four notes or syllables. The group level is organized in a manner which is either *regular* or *irregular*; in the regular condition, the groups always consist of three beats, and in the irregular condition, groups alternatively consist of two, three, or four beats. *Metricality* is the name devised to refer to the group-level regular and irregular conditions. Metrical levels necessarily exist above the group level (at higher *measure* and *phrase* levels), but measurement and discussion of these levels is outside the scope of the dissertation, since temporal predictability established at the group level may enhance or hinder entrainment. Figure 5.3 shows the metrical parameters.

### Syntax

The parameters for syntactic content necessitated comparability of type and location of syntactic manipulation. Melodies and sentences were given a complex syntactic structure with alternate endings (playing upon the ambiguity of relative keys and relative clause attachment), one 'more natural' and one 'less natural' sounding (but neither incorrect); where the syntactic processing was investigated, preferences for more-natural (*preferred*) and less-natural (*non-preferred*) alternate endings were determined on a per-participant basis. The

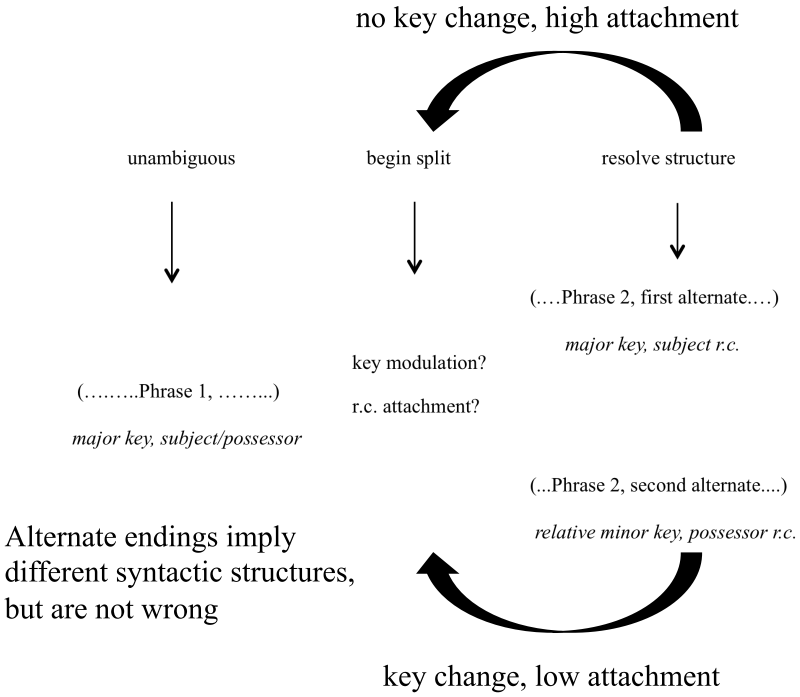



*Figure 5.4: Cross-domain syntax comparison: Alternate syntactic structures.* ***Syntax****: The first phrase of the melody or sentence unambiguously establishes a major key or subject and possessor. The second phrase begins the possible structural split in the first note or syllable: in one alternate the major key remains the same or the relative clause (r.c.) attaches to the subject, and in a second alternate the key modulates to the relative minor and the r.c. attaches to the possessor. The final structure is determined by tonic cadence or verb conjugation near the end of phrase 2. Preferred and non-preferred conditions are assigned in Study 3.*

choice to avoid error–detection and instead focus on complexity of syntactic structure was intended to highlight individual differences in comprehension-related processes as opposed to executive processes. And, in an attempt to create similar processing timelines and costs, the ‘split’ between the two structures and the initial detection of the alternate ending occurred in the same location in melodies and sentences. Determination of the final syntactic structure occurred in the span of one eighth note or one syllable (individual beat) with the added benefit that a subtle change would highlight individual differences in the ability to detect and integrate tonal and phonemic information. Figure 5.4 roughly depicts the alternate syntactic structures, described in detail in Section 5.3.2 below.

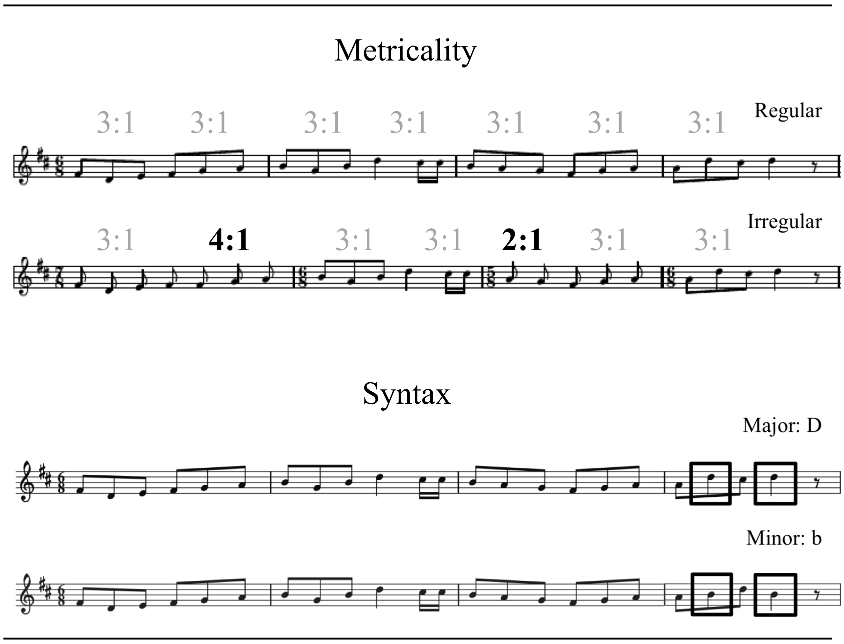



*Figure 5.5:* ***Metricality****: the regular condition has all measures written in 6/8 (strong-weak-weak), and the irregular condition has all melodies written in an alternating 7/8-6/8-5/8-6/8.* ***Syntax****: each melody is two phrases across four measures, the first phrase clearly in a major key, the second phrase ending alternatively in major or relative minor, resolved consistently in anti-penultimate and final notes.*

### 5.3.2 Materials

#### Music stimuli

**Composition** Sixty melodies were composed by Peter Adams. The melodies were then edited by the author to create regular (6/8) and irregular (7/8/5/8 variant) conditions. One note or syllable was added or subtracted appropriately in the melodic contour of a 6/8 regular melody to create its 7/8/5/8 irregular variant. Each melody was comprised of two phrases. In the first phrase, two measures clearly outlined a major key, and the two alternate syntactic structures began their possible division with the first note of the second phrase. The final syntactic structure (whether the key remained constant or modulated to the relative key in the second phrase) was first detectable at the antepenultimate note, which was consistently the tonic of the melody, followed by repetition of the tonic via a final cadence. Example melodies are shown in Figure 5.5. See Section 3.1.1 for a review of relative keys.

In order to avoid any influence of preference of one key over the other (as opposed to the intended preference of major over minor), the melodies were then transposed to four additional keys from the key of C, such that twelve melodies each were in C/a, G/e, F/d, Eb/c, and D/b. The keys were verified by a probabilistic key-finding algorithm (Temperley & Marvin, 2008). The critical note establishing the major-key or minor-key second phrase was always 'buried' in the contour of the melody to assure that its perception was purely syntactic and not due to sensory factors (e.g., not a note 'sticking out' in the contour). A repetition of tonic note after penultimate eighth note was appended on all melodies to avoid any wrap-up effects in EEG experiments (e.g. as reported in Friederici & Frisch, 2000). Items are listed in Appendix A.1.

**Recording and digitization** Melodies were recorded on a Yamaha Clavinova CLP 150 (Yamaha Corporation, Hamamatsu, Shizuoka, Japan) electric MIDI keyboard (44.1 kHz sampling rate) by a conservatory-trained pianist using Finale 2008 (Boulder, CO, USA). She was instructed to clearly enunciate the meter in respective melodies. She was originally given a 600 bpm metronome to which to synchronize, as was the speaker, but then requested the metronome be changed to thrice the speed (one beat per each note, instead of only the metrically strong notes). The melodies were post-processed with Audacity 1.3 and Praat 5.2, to normalize the amplitude (intensity) to 70 dB (the same as the sentences). In the case of a mistake (for example a slipped finger which pressed two keys at once) found during the audio file editing process, the correct note from a similar metrical position of a different melody replaced the incorrect note.

**Cross splicing** The pianist was instructed to not influence the interpretation of the melody with respect to its major or minor ending; however, melodies were cross-spliced to ensure that the syntactic ending (major vs. minor) could not be predicted by any expressive difference in piano performance. In post-hoc digital processing, within each melody pair the measure preceding the ambiguity resolution was switched across syntactic conditions. The cross-splice design is shown in Figure 5.6.

#### Music Fillers

The above-described music stimuli were dispersed with a 50% ratio among filler melodies. Filler melodies were designed with several intentions geared toward the EEG experimental design (Studies 2 and 3). Metricality conditions intended to boost entrainment effects for regular stimuli and minimize any familiarization or learning of the meter in irregular stimuli. Thus regular fillers were composed in the same metrical pattern as regular stimuli, and irregular fillers were composed in metrical patterns different from the stimuli. Syntax conditions intended to allow a comparable tonality change between the first and second phrases, allow-

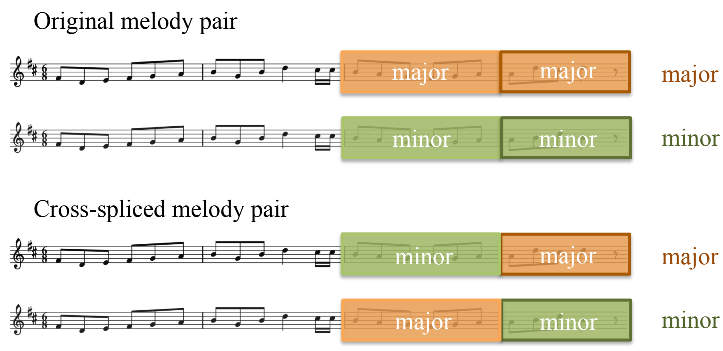



*Figure 5.6: Expressive control: The penultimate measure across syntactic conditions was cross-spliced. The final cross-spliced melody pairs contained the original minor penultimate measure and major final measure or the original major penultimate measure and the minor final measure.*

ing a task to equally apply to stimuli and fillers. Thus half of the filler melodies remained in the same key while half modulated to a new key, one whole step above the starting key (an example from D major to E major is in Figure 5.7). Filler melodies were recorded with the same pianist as the stimuli, with like instructions and subsequent digital processing. See Figure 5.7 for an example, and Appendix A.1 for a complete list of fillers.

### Language stimuli

**Composition** Sixty sentences with four variations each were composed. The four variations were comprised of two pairs, one with regular beat groups and one with irregular beat groups (group boundaries were defined by the natural weight of syllables). Each sentence was comprised of two phrases (clauses). In the first clause, plural verb conjugation clearly established a sentence subject (plural noun) that had a genitive possessor (feminine singular). The two alternate syntactic structures began their possible division with the first word of the second clause (ambiguous 'who'). The final syntactic structure (whether the second clause was a subject relative clause or a possessor relative clause) was first detectable at the final verb conjugation (either singular or plural). Example sentences are shown in Figure 5.8. See Section 3.1.2 for a review of relative clauses.

The lexical content of the sentences was strictly controlled. Two prepositions started each sentence (ten combinations), followed by a verb conjugated for the first plural noun

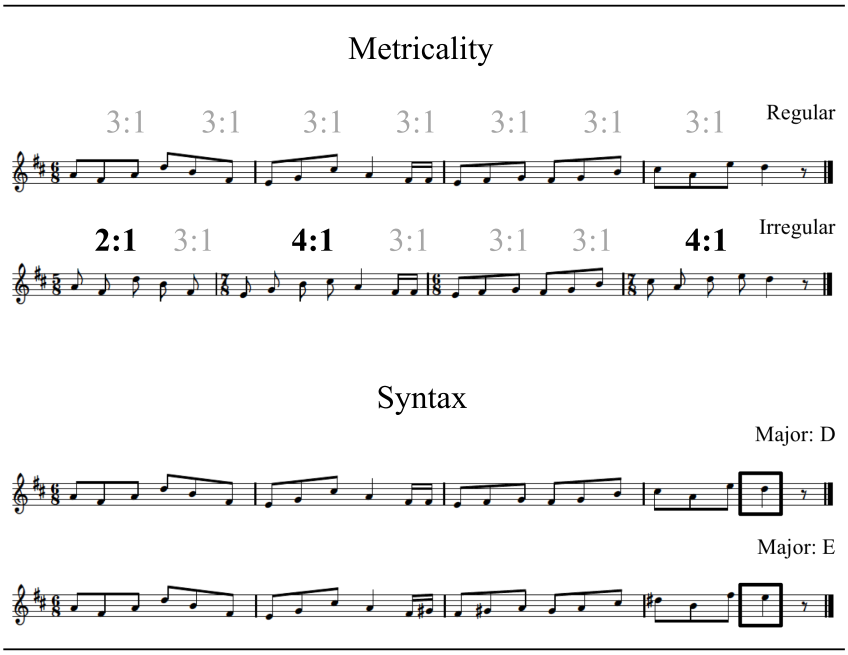



*Figure 5.7: Example filler melodies.* ***Metricality****: Filler items were also composed in regular and irregular meter. The regular fillers retained a ratio of 3:1 individual beats per group, identical to the regular stimuli. The irregular fillers were composed such that the 4:1 or 2:1 beat ratio occurred in locations different than in the stimuli.* ***Syntax****: after the first phrase, the following pickup into the new phrase could either remain in the same key, or be a pivot note into a new key.*

## Metricality

3:1 3:1 3:1 3:1 3:1 3:1 3:1 — Regular

***Da*** hinten ***plau***dern die ***Brü***der der ***Let***tin, die ***Kap***stadt vor ***kur***zem be***such***te/n und ***moch***te/n.

*Back there the brothers of the Latvian woman, who recently visited and liked Cape Town, are chatting.*

3:1 **4:1** 3:1 3:1 **2:1** 3:1 3:1 — Irregular

***Da*** hinten ***fach***simpeln die ***Brü***der der ***Let***tin, die ***Paarl*** vor ***kur***zem be***such***te/n und ***moch***te/n.

*Back there the brothers of the Latvian woman, who recently visited and liked Paarl, are talking shop.*

## Syntax

Subject relative clause

Da hinten plaudern **die Brüder** der Lettin, **die** Kapstadt vor kurzem besuch**ten** und moch**ten**.

*Back there the brothers of the Latvian woman, who recently visited (plural) and liked Cape Town, are chatting.*

Possessor relative clause

Da hinten plaudern die Brüder **der Lettin**, **die** Kapstadt vor kurzem besuch**te** und moch**te**.

*Back there the brothers of the Latvian woman, who recently visited (feminine sing.) and liked Cape Town, are chatting.*

*Figure 5.8: Language stimuli design. **Metricality**: Regular sentences have even groups of three syllables, and irregular sentences have alternating groups of two, three, or four syllables. The different conditions were established by using semantic equivalents with different syllable counts, for example the verb 'plaudern' has two syllables whereas the verb 'fachsimpeln' has three syllables–both words approximate the meaning to talk. **Syntax**: Alternate endings are subject or possessor relative clauses, determined by the final verb conjugation.*

(five verbs total, paired with homonyms to create two or three syllable options to fit the respective metricality condition), then an article with plural noun (15 animate words of relation: aunts, friends, etc.) to set up the first possibility for the ambiguous reflexive pronoun, next a genitive article and singular noun (15 animate words of a nationality, title or profession: heiress, customer, Englishwoman etc.), providing the second option for the ambiguous reflexive pronoun. In the second clause, the first item was always ‘die’/*who*, followed by a city (15 locations, either well-known international cities or cities within Germany, which were paired with a similar class of city to create a one or two syllable variation: e.g., London and Leeds), followed next by the verb ‘besuchen’/*to visit* (conjugated for either singular or plural), followed by the conjunction ‘und’/*and* and a final verb, ‘mögen’/*to like* (conjugated to the same number as *to visit*), appended to sentences in order to avoid EEG wrap-up effects. Lexical items were pseudo-randomized to create 60 sentence versions, and then a second set was reproduced with the partner three-syllable verb or one-syllable city in the first and second clauses, respectively. A native German speaker reviewed the sentences for grammaticality and plausibility. The lexical items for the sentences were controlled for frequency using the Leipziger Wortschatz and Celex word databases. Items are listed in Appendix B.1.

**Recording and digitization** Sentences were recorded with a Rode NT55 microphone (Silverwater, Australia) at 16-bit resolution and a sampling rate of 44.1 kHz using Cool Edit Pro 2.0 (Sibiu, Romania). A trained female native German speaker spoke the sentences in a soundproof room. She was instructed not to assign any special prosodic intonation to either of the nouns to which the ambiguous relative pronoun referred. To keep the rate of speaking steady, she entrained to a 600 bpm metronome during the session, with the instruction to time the strong syllables to the downbeat. After the sentences were recorded, they were digitally processed using Praat 5.2. To ensure that stimuli across domains retained comparable dynamics, the amplitude profile of the sentences was normalized to 70 dB (the same as melodies).

**Cross splicing** Although the speaker was instructed not to give any prosodically disambiguating intonation to the subject of the relative clause, sentences were post-processed to rule out any subconscious cuing by the speaker. The resulting cross-splice plan exchanged critical words in the noun phrase from the first clause (Figure 5.9). After the cross-splice, a sentence pair would have either both nouns from a ‘subject’ position or both nouns from the ‘possessor’ position. Importantly, within a syntactic condition, exactly one (best) version of the final phrase, including the critical word ’besuchten/besuchte,’ was inserted into every sentence to eliminate any slight differences in pronunciation which may influence an EEG

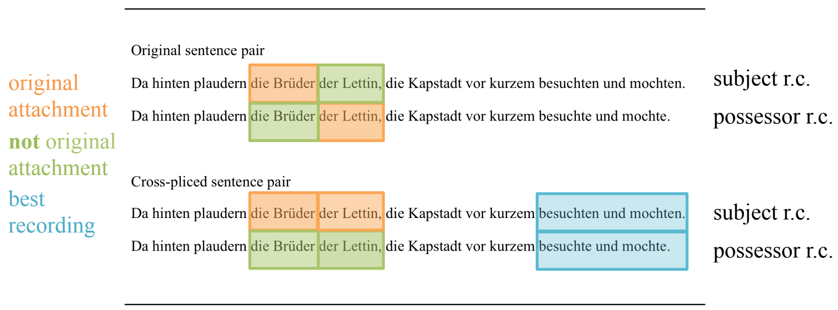



*Figure 5.9: Sentence cross-splice design. R.c. = Relative clause. Across each sentence pair, either the first or second of the NPs were cross-spliced so that either (1) only the NP that was intended to be the subject of the relative clause is present, or (2) only the NP which was not intended to be the relative-clause subject is present. Whether the first or second verb was cross-spliced was counter-balanced across items. For the final verbs, the best-spoken sample of plural and singular verbs (over all recorded sentences) was inserted into each sentence of the respective condition.*

response. If any words were of lesser quality due to noise in the recording, the same word with a clearer recording was substituted, but only from a sentence of the same syntactic condition.

### Language Fillers

The above-described stimuli were distributed with 50% occurrence. Filler melodies were designed with several intentions geared toward the EEG experimental design (Studies 2 and 3). Metricality conditions intended to boost entrainment effects for regular stimuli and minimize any familiarization or learning of the meter in irregular stimuli. Thus regular fillers were composed in the same metrical pattern as regular stimuli, and irregular fillers were composed in metrical patterns different from the stimuli. Syntax conditions, like the stimuli, contained an ambiguity surrounding the word ’die’/*who*. Instead of an ambiguous relative clause attachment, however, this time the ambiguity surrounded the case marking of ’die’/*who* being ambiguously nominative or accusative. Fillers were ambiguous in attempt to prevent participants from developing strategies pertaining to the ambiguity of the target sentences. A similar lexical pool was used for the content words, with the same metric, semantic and frequency constraints. The filler sentences were recorded with the same speaker during the same session as the stimulus sentences, with like instructions and subsequent digital processing. See Figure 5.10 for an example, and Appendix B.1 for a complete list of fillers.

Metricality

3:1 3:1 3:1 3:1 3:1 3:1 3:1 Regular

***Drau***ßen ver***weilt*** die Be***am***tin aus ***Bonn***, die den ***Dich***ter vor ***kur***zem er***kann***te und ***grüß***te.

**2:1** 3:1 **4:1** 3:1 3:1 3:1 **4:1** Irregular

***Drau***ßen ***war***tet die ***Rich***terin aus ***Mainz***, die den ***Dich***ter vor ***kur***zem er***kann***te und be***grüß***te.

Syntax

Subject relative clause

Draußen verweilt die Beamtin aus Bonn, **die den** Dichter vor kurzem erkannte und grüßte.

Object relative clause

Draußen verweilt die Beamtin aus Bonn, **die der** Dichter vor kurzem erkannte und grüßte.

*Figure 5.10: Example filler sentence.* ***Metricality****: Filler items were also composed in regular and irregular meter. The regular fillers retained a ratio of 3:1 individual beats per group, identical to the regular stimulus condition. This was to promote entrainment across regular blocks in the EEG experiment. The irregular fillers were composed such that the 4:1 or 2:1 beat ratio occurred in locations differing from the irregular stimuli.* ***Syntax****: The second phrase was ambiguously attached, with either a subject relative clause, meaning the pronoun was the agent, or object relative clause, meaning the pronoun received the action.*

### 5.3.3 Design of experiments

The empirical part of this dissertation is comprised of three studies, each with two experiments. In Study 1, Experiment 1a evaluates the diagnostic scores from all participants by means of a factor analysis, which statistically assesses which psychological processes are cognitively linked (e.g., if two or more scores 'load' onto one factor). Experiment 1b validates the stimuli, to assure that the syntactic structures are discriminable (via the task), while doubly serving to assess behaviorally whether groups of musicians and non-musicians entrain to the regular-meter melodies and sentences in a 2 x 2 design with factors Metricality (regular, irregular) by Domain (music, language). Study 2 evaluates the frequency spectrum of both stimuli (audio files) and EEG for evidence of frequencies representing the group or individual level of the metrical hierarchy. Experiment 2a assesses which peaks (if any) represent levels in the metrical hierarchy of stimuli, and Experiment 2b assesses whether peaks in EEG reflect entrainment to hierarchical levels in stimuli. The amplitude of found peaks is compared in a 2 x 2 x 2 design of factors Metricality (regular, irregular) by Hierarchy (group, individual) by Domain (music, language). Study 3 investigates the comparability of syntactic integration in both domains in an ERP analysis, and whether entrainment impacts syntactic integration comparably across domains. Experiments 3a and 3c compare amplitudes of found ERP effects respective in music and language domains, both in a 2 x 2 design with factors Metricality (regular, irregular) by Syntax (preferred, non-preferred). Experiment 3b pilots German relative clause attachment preference for the purposes of coding the preferred syntax conditions in Experiment 3c. The 'individual differences' approach to this dissertation on the one hand manifests in stimuli parameters which highlight individual differences in meter and syntactic comprehension, and on the other hand the correlation of participants' cognitive factor scores (from Experiment 1a) with experimental results across all three studies.

### 5.3.4 General hypotheses

It is hypothesized that participants entrain to regular more than irregular meter across all experiments—in both music and language domains—due to its greater temporal predictability reinforced by nested metrical levels (Jones, 1976, 2008; Large & Jones, 1999; see Section 2.1.1). Entrainment should facilitate efficiency among neural networks due to their increased synchronization (e.g., Singer, 2013), reflected by faster reaction time and higher accuracy in behavioral measures (Study 1) and higher peaks to entrained frequencies in EEG (Study 2).

The complex syntactic structures are expected to produce a P600 component in both domains (Study 3), whose amplitude is significantly greater where participants are confronted with the structure they find less natural or prefer less (e.g., Gouvea et al., 2010; Patel et al., 1998). An interaction between meter and syntax is also hypothesized in both domains, reflected by reduced amplitude of the P600 effect indicating ease of processing cost in entrained condition (Roncaglia-Denissen et al., 2013). Finally, cognitive factor scores representing the diagnostic test battery are expected to correlate with behavioral and EEG measures to reveal any relationships between those cognitive factors and entrainment to meter or syntactic integration.

# Chapter 6

# Study 1: Individual differences and detection of critical notes and syllables

The primary goal of Study 1 was to see whether cognitive entrainment to metrical context impacts detection of notes and syllables critical to syntactic structure, and in what way (if any) individual differences in cognitive processing play a role. This goal was addressed in two experiments: Experiment 1a intended to establish a stable index of individual differences relevant to performance in experiments throughout the dissertation—scores from diagnostic tasks assessing musical expertise, temporal perception, and verbal working memory—and Experiment 1b assigned participants the detection of subtle syntactic events in the context of either regular or irregular meter to see whether cognitive entrainment to regular meter facilitated performance. Experiment 1b further served to validate the stimuli (i.e., to determine whether the non-error syntax manipulations across domains were at all detectable).

## 6.1 Experiment 1a: Collection and exploratory factor analysis of diagnostic scores

### 6.1.1 Introduction

Chapter 4 outlined relative roles of musical expertise, temporal perception and working memory in aspects of meter and syntax processing in music and language domains. Musical expertise was associated with improved detection of subtle harmonic and morphosyntactic information (Fitzroy & Sanders, 2013; Koelsch et al., 2002), improved syntactic integration of monophonic harmonic structure (Besson & Faïta, 1995) and linguistic long-distance dependencies (Brod & Opitz, 2012), and higher verbal working memory capacity (Chan et al., 1998; Ho et al., 2003). Temporal perception played a role in both the initial perception of notes and syllables and the processing of syntactic structure (Schmuckler & Boltz, 1994; Strait et al., 2012; White-Schwoch & Kraus, 2013), and auditory working memory was associated with improved encoding of syllables (Tzounopoulos & Kraus, 2009), general music syntax perception (Fiveash & Pammer, 2014), and integration of ambiguous language syntactic structures (Bornkessel et al., 2004).

Seven diagnostic tests were selected in order to capture participants' abilities associated with musical expertise, temporal discrimination and verbal working memory. Exploratory factor analysis (Spearman, 1904) was subsequently applied to the seven scores in order to evaluate whether a single cognitive resource underlies more than one diagnostic test. (This approach streamlines representation of individual differences, thereby circumventing problems of multiple comparisons when assessing individual differences among these participants in subsequent experiments.) The musical expertise of each participant was also assessed according to age onset of musical training and current amount of practice hours. These measures are among reported quantization of musical expertise, and associated with anatomical and physiological changes in musicians that may account for improvement of cognitive functioning such as those listed above (Ellis et al., 2012; Tzounopoulos & Kraus, 2009). As the factor analysis was exploratory, no specific hypothesis was generated for correlations between any resulting factor scores and age onset or current practice hours.

### 6.1.2 Methods

#### Participants

Participants were 88 native German speakers, 50 female, with a mean age of 25 (*SD* 7) years, and spanned a range of musicianship (from self-reported non-musical to music conservatory graduates). Participants were students (university, vocational, or conservatory) and professionals from the Leipzig area who were listed in the Max Planck Institute participant database. Of the participants, 30 were self-reported non-musicians and 58 were self-reported musicians.

#### Materials

Diagnostic tests included the forward and backward digit spans (Tewes, 1994), non-word repetition (Mottier, 1951), modified listening span (Daneman & Carpenter, 1980), Musical Ear Test (Wallentin et al., 2010), and anisochrony detection threshold (Dalla Bella et al., 2016). A full description of each test is listed in Section 5.2.

### Procedure

All tests were administered via headphones at a comfortable volume. The total amount of time needed to complete all seven tests was approximately 1.5 hours, which was conducted in one or two sessions, dependent upon participant availability. Some participants already had test scores on record, in which case their previously-recorded data was used. Correct-response hand (left or right) was kept constant for all tests, and this as well as the order of completed tests was counterbalanced across participants. Self-reported musicians were additionally asked at what age they commenced music training and how many hours per week they spent practicing during the time of experimentation. Participants received €7 per hour for their participation. The experiment was approved by the local ethics committee of the University of Leipzig. A list of participants musical instrument, age of musical training onset, and weekly practice time can be found in Appendix C.

### Analysis and results

**Factor analysis of diagnostic scores** An exploratory factor analysis was performed on the seven diagnostic scores with SPSS Version 20 (IBM, Armonk, NY). This statistical procedure searches for mutual statistical dependence among multiple test scores, with the assumption that a common underlying cognitive resource may justifiably be responsible for multiple scores if those scores 'cluster together' or 'load' onto a factor.

Exploratory factor analysis does not require that data be normally distributed; however, extreme outliers may not be present in the data in order for the procedure to be accurately conducted. The data were therefore screened for extreme outliers among standardized diagnostic scores: extreme-outlier cases were flagged in a box plot using SPSS Explore procedure, which means that the case had a value less than or equal to the first quartile minus 3 times the interquartile range, or greater than the third quartile plus 3 times the interquartile range. Two out-of-range values were found in anisochrony detection, resulting in exclusion of those participants. The minimum amount of data for factor analysis was satisfied (at least 10 participants per individual item input, Field, 2005), with a final sample size of 86 providing a ratio of over 12 cases per variable.

The seven diagnostic test scores were appropriately examined, using several well-recognized criteria for passing the assumptions of exploratory factor analysis (Field, 2005). First, it was observed that six of the seven test scores correlated at least .3 with at least one other item, suggesting reasonable factorability. Second, the Kaiser-Meyer-Olkin measure of sampling adequacy was .67, above the commonly recommended value of .6, and Bartlett's test of sphericity was significant ($\chi^2(21) = 84.14$, $p < .001$). The diagonals of the anti-image correlation matrix were also all over .5, and the determinant was greater than 0.00001 (.358), indicating no problems with multicollinearity. Finally, the communalities were all above .3 (see Table 6.1), further confirming that each item shared some common variance with other items. Given these overall indicators, factor analysis was deemed to be suitable with all seven test scores.

Principal component analysis was used because the primary purpose was to identify and compute composite scores for any factors underlying the seven diagnostic scores. Initial eigenvalues indicated that the first two factors explained 33% and 19% of the variance respectively. The third through seventh factors had eigenvalues below one, and were therefore not further considered. The two factor solution, which explained 52% of the variance, was accepted because of theoretical support for clustering of verbal WM and pitch and temporal discrimination on separate factors, as well as the insufficient number of primary loadings and difficulty of interpreting subsequent factors.

An oblimin rotation provided the best defined factor structure, as there was no reason to believe that the two factors should be wholly independent from one another. The pattern matrix was consulted to interpret the unique contribution of each test score to its factor; all test scores had primary loadings over .5, and cross-loadings were all below .2. The pattern matrix is presented in Table 6.1.

The factor labels were assigned according to the common cognitive functions responsible for their respective tasks: Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD). Consistency of the test scores per factor was examined using Cronbach's $\alpha$, in order to assure that the various scales of each test (though standardized) were comparable. The $\alpha$ was moderate among VWM tests (.66 for four items), and poor for TPD (.55 for three 3 items). No increases in $\alpha$ for either factor components could have been achieved by eliminating any diagnostic tests. As the $\alpha$ scores were lower than .7, an alternative method was sought to test the stability of the factor loadings. A jackknife procedure was employed, whereby each participant was removed and the factor loadings re-estimated with the remaining 85 participants. For all 86 reiterations of the factor analysis, the two factors remained constant and retained their individual diagnostic test components (although the loading-order varied slightly within factors). Since the two factors remained consistent during all jackknife iterations, and all other assumptions of the factor analysis were met, the internal consistency of the two-factor solution was deemed sufficient.

Composite scores output from SPSS represented a standardized factor score per participant. Higher scores indicated greater aptitude in the cognitive functions underlying VWM and TPD. Descriptive statistics are presented in Table 6.2. The skewness and kurtosis were well within a tolerable range for assuming a normal distribution, and both factors were normally distributed (VWM: Shapiro-Wilk $W$ (86) =.99, $p$ =.42; TPD: $W$ (86) = .98, $p$ = .34). Although an oblimin rotation was used, only a small correlation existed between the

*Table 6.1: Factor loadings and communalities based on a principal components analysis with oblimin rotation for seven diagnostic test scores. VWM = Verbal Working Memory, TPD = Time and Pitch Discrimination. Factor loadings < .2 are suppressed.*

| | VWM | TPD | Communality |
|---|---|---|---|
| Backward digit span | .80 | | .62 |
| Forward digit span | .79 | | .60 |
| Modified listening span | .59 | | .41 |
| Non-word repetition | .58 | | .42 |
| MET melody subtest | | .74 | .59 |
| Anisochrony detection | | .70 | .55 |
| MET rhythm subtest | | .68 | .47 |

*Table 6.2: Descriptive statistics for standardized scores of the two factors. VWM = Verbal Working Memory, TPD = Time and Pitch Discrimination.*

| | No. of loaded items | Min. | Max. | Skewness | Kurtosis | Cronbach's $\alpha$ |
|---|---|---|---|---|---|---|
| VWM | 4 | -2.36 | 2.02 | -0.14 | -0.61 | .66 |
| TPD | 3 | -2.67 | 1.94 | -0.29 | -0.31 | |

composite scores (.19).

**Expertise (correlations)** The diagnostic scores were then analyzed in terms of expertise. Non-musicians were defined as having no history of extra-curricular musical training, and musicians were defined as practicing an instrument or voice for at least one hour per week at the time of testing. Based on these criteria, two self-reported non-musicians with extra-curricular childhood musical training were discarded from the 86 participants with factor scores, along with four self-reported musicians who no longer were actively playing their instrument. Final group sizes were 26 non-musicians (mean years of age = 26.3, *SD* = 3.1, 11 female) and 54 musicians (mean years of age = 23.4, *SD* = 2.5, 31 female). Among musicians, average age onset of musical training was 11.1 years (*SD* = 4.8) and average number of practice hours per week was 5.38 (*SD* = 4.25). Onset age of training was considered as the youngest age of regular training on any instrument regardless of whether musicians later switched primary instruments, also in the case of two musicians who reported a break of two years or less in active practice/education during late adolescence. Practice hours per week were cumulative in the case of musicians who played multiple instruments.

Based on findings reported in the literature (Rammsayer & Altenmüller, 2006; Tzounopoulos & Kraus, 2009; Wallentin et al., 2010), a *t*-test was performed with the hypothesis that musicians would have higher VWM and TPD than non-musicians. The results partially

*Table 6.3: Musical expertise and the two factors: independent-samples t-test. VWM = Verbal Working Memory, TPD = Time and Pitch Discrimination. Musicians had a significantly higher TPD mean. [a]n = 26. [b]n = 54. Significant ($\alpha$ < .05) effects in bold. Absolute Cohen's d reported.*

| | Mean (SD) | | | | |
|---|---|---|---|---|---|
| | Non-musicians[a] | Musicians[b] | *t*(78) | *p* | *d* |
| VWM | -0.11 (0.99) | .02 (1.02) | 0.56 | .58 | 0.13 |
| TPD | -0.70 (0.83) | .38 (0.83) | 5.39 | **<.001** | 1.30 |

*Table 6.4: Musical expertise and individual diagnostics: independent-samples Mann-Whitney U Test. VWM = Verbal Working Memory, TPD = Time and Pitch Discrimination. Impact of expertise on individual diagnostic scores upheld impact on factor scores. [a]n = 26. [b]n = 54. Mann-Whitney's U, Bonferroni-corrected $\alpha$ significance threshold is .0071. Significant effects in bold. Absolute effect sizes reported.*

| | Median | | | | |
|---|---|---|---|---|---|
| | Non-musicians[a] | Musicians[b] | *U* | *p* | *r* |
| VWM | | | | | |
| Backward digit span | 01.07.00 | 01.06.00 | 675.0 | .78 | .03 |
| Forward digit span | 01.07.00 | 01.07.00 | 693.0 | .92 | .01 |
| Modified listening span | 01.04.00 | 01.04.00 | 625.5 | .39 | .10 |
| Non-word repetition | 01.05.75 | 01.06.00 | 592.0 | .23 | .13 |
| TPD | | | | | |
| MET Melody subtest | 32.00 | 39.50 | 290.5 | **<.001** | .47 |
| Anisochrony detection | 01.10.88 | 01.08.77 | 427.0 | **.005** | .32 |
| MET Rhythm subtest | 33.50 | 38.00 | 393.5 | **.001** | .36 |

supported the hypothesis: musical expertise affected TPD but not VWM; musicians had overall better TPD but not better VWM than non-musicians (Table 6.3).

As particularly auditory/verbal working memory has in the past been reported to be affected by musical expertise (Chan et al., 1998; Tzounopoulos & Kraus, 2009), independent-samples tests were individually administered to each diagnostic score to see whether any effect was masked due to relative loading on a factor. Using the Bonferroni method, $p$ significance threshold was adjusted to .0071 from .05. Six of the seven tests were non-normally distributed within-group (Shapiro-Wilk $W$ test (musicians): backward digit span $W = .90$, $p < .001$; forward digit span $W = .90$, $p < .001$; modified listening span $W = .87$, $p < .001$; non-word repetition $W = .89$, $p < .001$; MET melody subtest $W = .92$, $p = .002$; MET rhythm subtest $W = .95$, $p = .016$); therefore non-parametric independent-samples Mann-Whitney $U$ Test was employed.

Results of all seven Mann-Whitney U tests reflect findings of the $t$-tests with factor

*Table 6.5: Correlation of factor scores and aspects of musical expertise. VWM = Verbal Working Memory, TPD = Time and Pitch Discrimination. N = 54 musicians. Hours of weekly practice correlated with TPD. Spearmans's ρ, p < .05*. Significant correlation in bold.*

| | Age at onset of musical training | Practice hours per week |
|---|---|---|
| VWM | -.04 | .02 |
| TPD | .01 | **.32*** |

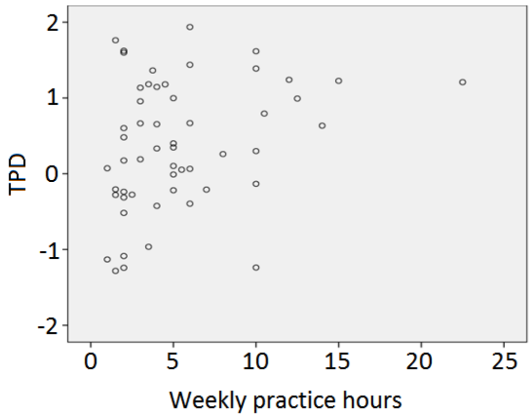



*Figure 6.1: Correlation of Time and Pitch Discrimination and weekly practice hours (musicians = 54).*

scores (Table 6.4): diagnostic scores which loaded onto VWM showed no effect of expertise whereas diagnostic scores which loaded onto TPD all showed effects of expertise. This finding supported the theoretical assumption that single cognitive resources were responsible for diagnostic tests which loaded onto like factors, and validated the further use of factor scores instead of individual diagnostic scores.

Within the musician group, age of acquisition and practice hours were then correlated with the two factor scores. Neither age of acquisition ($W$ = .82, $p$ = .006) nor practice hours ($W$ = .94, $p$ < .001) were normally distributed; non-parametric correlations are listed in Table 6.5. Results showed a significant correlation between hours of practice and TPD (Spearman's $\rho$ = .32, $p$ = .018). One outlying participant who practiced 22.5 hours per week strengthened the correlation but did not drive it entirely (Figure 6.1); without this participant the correlation was still significant (Spearman's $\rho$ = .30, $p$ = .03).

### 6.1.3 Results summary

These analyses indicated that two distinct factors were underlying the seven diagnostic tests collected in this study. Factors were named Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD), appropriate to the cognitive functions associated with tasks in the respective diagnostic tests found in each factor. Although Cronbach's $\alpha$ was slightly low especially in TPD, a jackknife procedure verified the stability of the factor loadings (constant for 86 iterations of "– $n$"). Standardized composite factor scores were normally distributed, thus the data were suited for parametric statistical analyses.

Musical expertise was associated with improved TPD but did not show any statistical relationship to VWM, which was verified by comparison of group means in each individual diagnostic test. Within musicians, the number of practice hours correlated positively with TPD but not VWM, whereas musical-training age onset showed no correlation with either factor.

### 6.1.4 Discussion

With the assumption that diagnostic scores index musical ability, temporal perception and verbal working memory, results of the exploratory factor analysis were expected to indicate whether cognitive factors underlying these tasks operate in conjunction or functionally discretely from one another. The separation of diagnostic scores into two factors is logical in this light: All of the tasks in VWM indeed involved verbal working memory, and various working memory tasks which included the phonological loop (Baddeley, 2000) were found to activate the same network in a large meta-analysis of working memory functional networks (Nee et al., 2013). The tasks clustered onto TPD each required discrimination of music-related properties: temporal interval, melodic contour, and rhythm patterns. Since many participants had musical training, it is plausible that a single cognitive resource across all participants would be recruited for the subset of tasks which involved musical abilities. It would be particularly unusual if the MET rhythm subtest and anisochrony detection task did not share some underlying cognitive resource, as interval discrimination is heavily emphasized in anisochrony detection task and also exclusively tested in some items of the MET rhythm subtest.

Verbal working memory was not affected by musical expertise (neither the factor score nor any individual diagnostic score), which contradicted previous findings in the literature (Chan et al., 1998; Ho et al., 2003; Tzounopoulos & Kraus, 2009). This may be due to the criteria for musicians including participants who started musical training in up through late adolescence and without regard to a history of consistent practice. Thus perhaps the musicians as a group missed some critical period for plasticity (e.g. Pantev et al., 1998; Schlaug, 2001) which would have offered better functional networks associated with verbal working memory[1].

[1] Although the current experiment employed looser criteria regarding training age-of-onset and uninterrupted weekly practice than some previous studies (Lotze, Scheler, Tan, Braun, & Birbaumer, 2003; Strait et al., 2012; Tzounopoulos & Kraus, 2009), the classification of musicians and non-musicians is nevertheless still

Musicians had a higher TPD mean score than non-musicians, in line with previous findings that musical expertise impacts melodic contour, rhythm- and real-time perception of temporal information (Rammsayer & Altenmüller, 2006; Wallentin et al., 2010, see footnote 1). The correlation with weekly practice hours suggests that perhaps these abilities are sensitive to short-term fluctuations in training or exposure. The perception of pitch contour and temporal discrimination are both necessary to evaluating complex sounds at the timescale of speech; indeed improved auditory brain stem encoding of speech is associated with increased exposure to music in the short-term (Tierney et al., 2013). This resource may therefore be associated with improved detection of pitches and syllables, addressed in the next experiment.

## 6.2 Experiment 1b: Metrical context and the perception of critical notes and syllables

### 6.2.1 Introduction

This experiment addressed the cognitive aspect of the current neurocognitive entrainment approach, which is described fully in Section 2.1. Briefly restated, cognitive entrainment refers to entrained attention such as that described by the dynamic attending theory (DAT, e.g., Large & Jones, 1999) and the attentional bounce hypothesis, (Pitt & Samuel, 1990). According to the DAT, listeners may utilize the temporal predictability in a stable meter, thus sharpening an internal attentional pulse, particularly when levels in the metrical hierarchy are coupled in simple integer rations such as 3:1 (as is the 6/8 meter employed in this paradigm; Large & Palmer, 2002). Within the DAT framework, this latter statement describes future-oriented attending, as opposed to analytic attending, which is rather based on attention to local regularities as opposed to hierarchical temporal ones (Jones & Boltz, 1989). In line with this, the attentional bounce hypothesis predicts that attention is guided by regularly placed strong syllables in language.

The detection of notes and syllables, the current task, is an essential early step in syntactic comprehension. In addition to containing information salient to metric and semantic processing, pitches and phonemes are also crucial building-blocks to syntactic structure. Perception of harmonic and morphosyntactic information carried by notes and syllables

valid. Both MET melody- and rhythm subtests were shown to be impacted by expertise when comparing non-musicians to mostly amateur musicians (including a handful of professional-track conservatory students). This extends findings of the initial MET validation study (Wallentin et al., 2010), where MET scores were only impacted by expertise when comparing professional musicians to non-musicians; amateur musician scores were not significantly different from either group. Thus the musician group still reflected superior performance in musical ability exceeding typical amateur standards, validating current group criteria as representative of musician performance.

may be enhanced when attention is entrained: previous evidence indicated that in music, future-oriented attending improves pitch discrimination compared to analytic attending (Boltz, 1993), and in language, the attentional oscillator when primed by a regular beat can improve phoneme discrimination (Cason & Schön, 2012). Thus in line with the current theoretical approach, regular melodies and sentences were hypothesized to promote cognitive entrainment over irregular melodies and sentences. This ultimately could impact ease of syntactic comprehension of the materials (addressed in Study 3).

The impact of entrainment on syntax-related perceptual processes in music and language has until now not been addressed within a single paradigm. The current experiment expected to replicate previous findings while extending the cross-domain comparability in three principle aspects: metrical context, salience of the syntactic event, and impact of individual differences.

First, meter and timing often lack empirical 'equal footing' in music and language. As outlined in Sections 2.4.1, 2.4.2 and 4.2.1, music-meter experiments have typically investigated meter perception using rigid (computerized) temporal profiles of hierarchical metrical structures (e.g. Ellis & Jones, 2010; Schmuckler & Boltz, 1994, but for an exception see Tierney & Kraus, 2015), while language-meter studies have typically employed meter defined by relative adjacent accent/weight of syllables spoken with natural human fluctuation in timing (e.g. Magne et al., 2007; Schmidt-Kassow & Kotz, 2009a). The current paradigm established a 6/8 or 6/8-variant meter in both melodies and sentences for metrical comparability, and both the pianist-performed melodies and spoken sentences contain natural human fluctuation in timing.

Second, the current paradigm attempted comparability with respect to the syntactic salience of notes and syllables; the type of detection required from participants was of a syntactic event (a pitch or phoneme) which structurally influenced the overall melody or sentence but was not a violation of syntactic structure per se. In music, the critical pitch indicated whether a key modulated to the relative minor or remained the same, and in language, the critical phoneme was a verb inflection which indicated whether a relative clause referred to the plural subject or singular genitive possessor of the main sentence (Figure 6.2; explanations of relative keys and relative clause attachment may be found respectively in Sections 3.1.1 and 3.1.2). Particularly in language, this subtle difference in morphosyntax that indicated number is ecologically valid, as in everyday life confusion may arise when ambiguous linguistic information is falsely interpreted.

Third, entrainment to meter and perception of critical syntactic information have been linked variously to auditory working memory, temporal perception and musical ability/expertise (see Chapter 4; Kraus et al., 2009; Marie et al., 2011; Tzounopoulos & Kraus, 2009). Thus the per-participant factor scores VWM and TPD established in Experiment 1a, as well as categorical musician and non-musician grouping, allowed unique insight into cognitive resources or experienced training which may have impacted the utilization of meter for entrainment and/or detection of subtle syntactic events.

#### Hypotheses

1. Participants will cognitively entrain to the regular meter in both sentences and melodies (similar to Lidji, Palmer, Peretz, & Morningstar, 2011), reflected in accuracy and reaction time.

2. Due to superior pre-attentive detection of morphosyntactic and harmonic information (Fitzroy & Sanders, 2013; Koelsch et al., 2002), musicians should react faster than non-musicians to the critical note and syllable.

3. Since cognitive entrainment has been demonstrated in participants with little musical experience (Ellis & Jones, 2010; Jones & Boltz, 1989), musical expertise should not influence entrainment to regular meter in either domain.

4. Musicians should outperform non-musicians in music (but not language), as the critical note is rendered more salient by theoretical knowledge of music syntax (Besson & Faïta, 1995; James et al., 2012).

5. As both verbal working memory and pitch and timing information are relevant to melody and sentence perception (Fiveash & Pammer, 2014; Kraus et al., 2009), both factor scores from Experiment 1a were expected to correlate with performance in both domains

### 6.2.2 Methods

#### Participants

Thirty participants who took part in Experiment 1a were pre-selected for melodic discrimination aptitude (participants must have had scores 27 out of 52 on the MET melody subtest, to assure that they could distinguish between the musical syntactic-conditions). Participants were split into groups with little to no musical experience, and those with musical training/expertise. Non-musicians ($n$ = 15, 7 females), aged between 22 and 31 years ($M$ = 26.3 years, $SD$ = 2.94), received little to no formal musical training beyond obligatory school courses (on average less than one year of training). One participant had 9 years of keyboard training, but had not practiced during the last 9 years before participating in the current experiment. Musicians ($n$ = 15, 8 females), aged between 20 and 23 years ($M$ = 21.1 years, $SD$ = 0.88) had an average of 13.5 years (range = 9–18 years) of musical training.

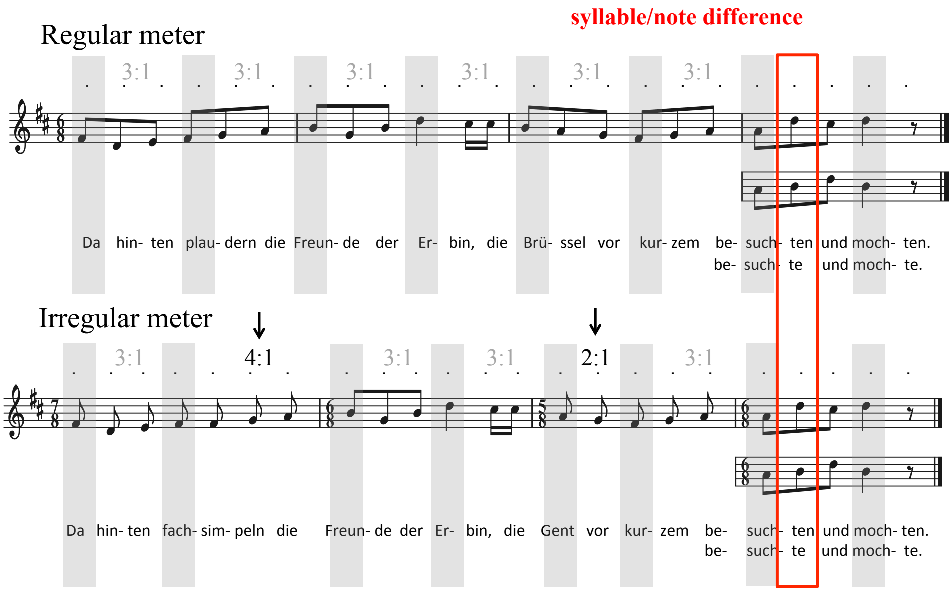



*Figure 6.2: **Meter**: Regular meter was a 6/8 or strong-weak-weak beat pattern. Irregular meter was an alternating 7/8, 6/8, 5/8, 6/8 meter or strong beats with varying three, two or one weak beats in-between. Gray boxes highlight strong beats. Stressed syllables in bold italics. **Syntax**: Critical note (in square) confirmed major or minor final key. Critical syllable (in square) contained the verb inflection which indicated whether the relative clause referred to plural subject or singular object of sentence.*

### Materials

Melody- and sentences pairs (120 total) each had two possible alternate syntactic endings: melodies resolved according to a major key or relative minor key, and sentences contained final verb inflections which assigned a relative clause to either the subject or genitive possessor of the main sentence. Of the pairs, half were comprised of a regular meter (6/8 or strong-weak-weak repeated syllables) and the other half were comprised of an irregular meter (the repeating pattern jittered). See Figure 6.2 for a depiction of the design, Section 5.3.2 for full stimuli description and appendices A.1 and B.1 for full list of melodies and sentences.

### Procedure

The experiment was divided into a language session and a music session, the order counterbalanced among participants and administered on different days. Sentences and melodies were presented at a comfortable volume via headphones. In both sessions, items were pseudo randomized such that no one condition occurred more than three consecutive times.

#### Task

Participants were presented with two melodies or sentences per trial, and asked to compare the content of the second to the first, pressing the corresponding button for ‘same’ or ‘different’ content. Trials were designed to be 50% same/different ratio. The order of melody or sentence type (major vs. minor key, plural vs. singular verb) was counterbalanced across trials. The button press ‘same’ and ‘different’ was counterbalanced among participants for right and left hands. Participants completed a short training to make sure they understood and could perform the task before starting the experiment.

### 6.2.3 Analysis and results

The mean reaction time (per condition per subject) for correct trials was calculated. Responses outside of two standard deviations from the mean were discarded, and the accuracy and reaction-time means adjusted accordingly. Participants whose accuracy was below 50% were discarded, as well as participants who did not follow task instructions. Seven participants were excluded with these criteria.

After exclusion criteria, the remaining 23 participants’ accuracy scores were converted to percentage scores, and a one-sample *t*-test performed to determine whether accuracy was significantly above chance-level (50%) in each variable. In all conditions, discrimination was above chance-level (Music, Regular: $t(1,22) = 9.2$, $p < .001$; Language, Regular $t(1,22) = 26.4$, $p < .001$; Music, Irregular $t(1,22) = 6.9$, $p < .001$; Language, Irregular $t(1,22) = 26.5$, $p < .001$), indicating that the pitch and morphosyntactic cues to respective harmonic modulation or relative clause attachment were discernible. Thus the stimuli were validated for subsequent experiments investigating syntactic integration processes.

#### Mixed-design ANOVAs

Analyses of variance were performed separately on accuracy and reaction times, with within-subjects factors Metricality (regular meter, irregular meter) and Domain (music, language), and between-subject factor Expertise (musician, non-musician). Accuracy values were log-transformed in order to stabilize variance between groups, as the variance for the condition music-regular was non-homogeneous across musicians and non-musicians (significant Levene’s Test of Equality of Error Variances, $F(1,21) = 6.43$, $p = .019$). Table 6.6 lists the results of an ANOVA on the transformed accuracy data and reaction time.

#### ANOVA main results

The regular meter context yielded more accurate scores and faster reaction times (main effect Metricality), partially supporting the hypothesis. Music responses were globally less

*Table 6.6: ANOVA of accuracy and reaction time. Significant (α < .05) effects in bold.*

| Effect | $F$(1,21) | partial $\eta^2$ | $p$ |
|---|---|---|---|
| Accuracy | | | |
| Metricality | 01.04.34 | .171 | **.050** |
| Domain | 103.40 | .831 | **<.001** |
| Expertise | 01.02.72 | .115 | .114 |
| Metricality x Expertise | 01.01.86 | .081 | .187 |
| Domain x Expertise | 6576 | .238 | **.018** |
| Metricality x Domain | 4.39 | .173 | **.049** |
| Metricality x Domain x Expertise | 01.01.29 | .058 | .270 |
| Reaction time | | | |
| Metricality | 01.06.76 | .244 | **.017** |
| Domain | 22.30 | .515 | **<.001** |
| Expertise | 01.04.48 | .176 | **.046** |
| Metricality x Expertise | 0.67 | .031 | .423 |
| Domain x Expertise | 0.02 | .001 | .888 |
| Metricality x Domain | 3 | .000 | .955 |
| Metricality x Domain x Expertise | 01.07.16 | .254 | **.014** |

accurate and slower compared with language responses (main effect Domain), which was not specified in the hypothesis. Reaction time was globally faster in musicians compared to non-musicians, supporting the hypothesis.

### ANOVA interactions

The regular meter context did not affect performance equally across domains, contrary to the hypothesis. In accuracy, the music domain was impacted by meter context more than the language domain (Metricality x Domain). The Metricality x Domain interaction was interpreted by splitting the 3-way ANOVA into two 2-way ANOVAs, one music and one language, and the effect Metricality was compared across domains. This showed that the Metricality effect reached significance in the music domain ($F(1,21) = 5.20$, $p = .033$, partial $\eta^2 = .198$), showing better performance in regular meter, but was not significant in the language domain ($F(1,21) = .015$, $p = .093$, partial $\eta^2 = .001$).

The hypothesis that expertise would not influence entrainment to regular meter context was partially supported. While the hypothesis was supported in accuracy, reaction times showed that metrical context influenced musicians in the music domain and non-musicians in the language domain. Accuracy: Interactions Metricality x Expertise and Metricality x Domain x Expertise were non-significant. Reaction time: The Metricality x Domain x Expertise interaction was split into two 2-way ANOVAs per domain, each with factors Meter and Expertise: In music, only musicians had a Metricality effect (faster RT in the regular-meter condition, see Figure 6.3). Metricality x Expertise was significant ($F(1,11) = 4.35$, $p = .049$, partial $\eta^2 = .172$). Significance of the Metricality effect (in musicians but not non-musicians) was confirmed by further splitting the comparison by expertise, and performing a paired-samples *t*-test within group. Among musicians, music regular RT was significantly faster than music irregular RT ($t(11) = -2.23$, $p = .048$) whereas among non-musicians there was no difference in RT ($t(10) = 0.58$, $p = .58$).

In language, only non-musicians had a Metricality effect (faster RT for regular sentence discrimination, see Figure 6.4). The Metricality x Expertise interaction was also significant ($F(1,11) = 4.9$, $p = .038$, partial $\eta^2 = .188$), but the effect was in the opposite direction when compared to the music domain. Significance of the Metricality effect (in non-musicians but not musicians) was confirmed by further splitting the comparison by expertise, and performing a paired-samples *t*-test within group. Among non-musicians, language regular RT was significantly faster than language irregular RT ($t(10) = -2.95$, $p = .014$) whereas among musicians there was no difference in RT ($t(11) = -0.12$, $p = .91$).

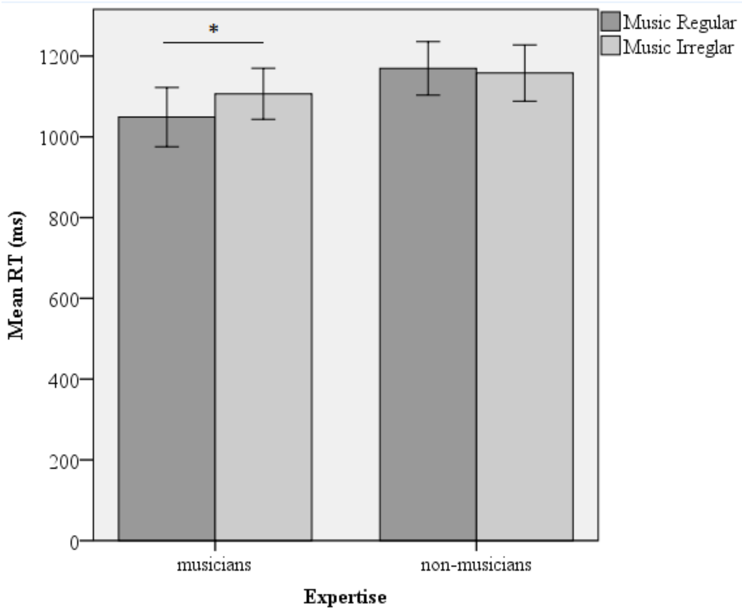



*Figure 6.3: Music RT: Expertise x Metricality ANOVA interaction. Musicians, but not non-musicians, have a faster RT in regular- compared to irregular-meter melody discrimination. Error bars ± 2 SEM. p < .05**

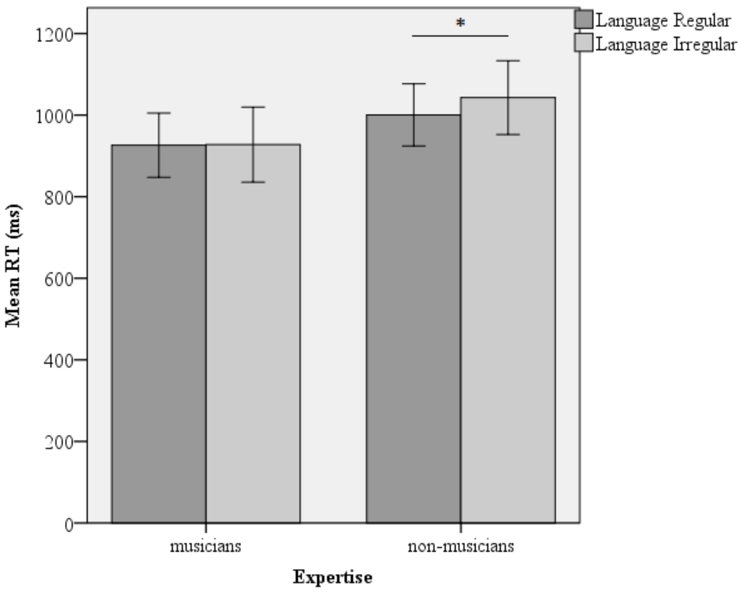



*Figure 6.4: Language RT: Expertise x Meter ANOVA interaction. Non-musicians, but not musicians, had a faster RT in regular- compared to irregular-meter sentence discrimination. Error bars ± 2 SEM. p < .05**

Musicians were globally more accurate than non-musicians in the music task, in ac-

*Table 6.7: Correlation of expertise and diagnostic factors with accuracy and reaction time. VWM = Verbal Working Memory, TPD = Time and Pitch Discrimination. Pearson's R, two-tailed correlations. Exp. = Expertise. L = language, M = music, R = regular meter, I = irregular meter. Significant correlations in bold. p < .01**, p < .05*, p < .1⁺*

| | | Accuracy | | | | RT | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Expertise | MR | MI | LR | LI | MR | MI | LR | LI |
| Expertise | | **.53**** | .25 | -.06 | -.14 | **-.47*** | -.23 | -.28 | -.36$^{+}$ |
| VWM | -.159 | .38$^{+}$ | .37$^{+}$ | .26 | .13 | -.07 | .04 | -.20 | -.23 |
| TPD | **-.64**** | **.72**** | .38$^{+}$ | .14 | -.02 | -.16 | -.03 | -.38$^{+}$ | **-.41*** |

cordance with the hypothesis. The Domain x Expertise interaction, when split into 2-way ANOVAs per domain, showed that expertise only had an effect in the music domain; musicians were globally better than non-musicians at discriminating melodies ($F(2,20) = 4.09$, $p = .032$, partial $\eta^2 = .29$). Expertise played no role in the language domain ($F(2,20) = 0.15$, $p = .86$, partial $\eta^2 = .015$).

#### Correlations

Verbal Working Memory had no significant correlations with either accuracy or reaction time, and TPD correlated with music regular- and language irregular performance beyond expertise, partially supporting the hypothesis. Expertise in turn only correlated with music regular performance (Table 6.7). The result for Experiment 1a—that musical expertise (dichotomous musician or non-musician classification) impacted TPD but not VWM—was replicated in Experiment 1b.

### 6.2.4 Results summary

Results of the ANOVAs are summarized in Figure 6.5.

#### Domain

The Domain main effect indicated that performance was globally better in language than in music, in both reaction time and accuracy measures across all participants.

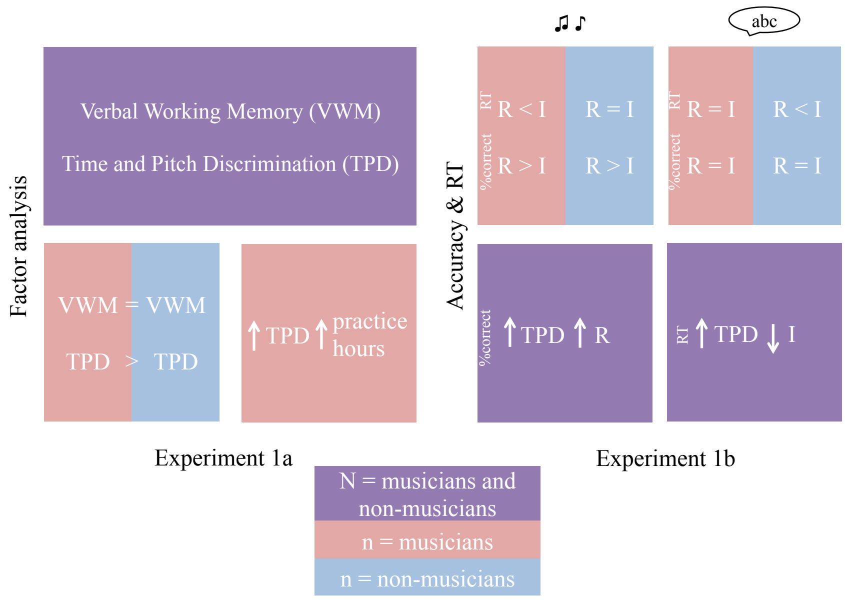



*Figure 6.5: Left panel: Experiment 1a factor loadings and correlations. Right panel: Experiment 1b behavioral results. $p < .05$ *.*

### Metricality

In accordance with the DAT (better attention to more predictable signal), regular metrical context offered perceptual advantage over irregular metrical context in discriminating critical notes. In contrast to the sweeping music result, regular metrical context had no global impact across all participants in discriminating critical syllables (but influenced non-musician RT).

### Individual differences

Musicians were globally better than non-musicians in music accuracy and music RT, and in language RT. Regarding the impact of metrical context, musical expertise only impacted results in the language domain: while no difference between musicians and non-musicians in music was displayed, non-musicians profited from regular meter in language RT whereas musicians demonstrated no impact of metrical context in either accuracy or RT. Furthermore, expertise correlated with TPD and music regular accuracy and reaction time. While VWM did not lead to any significant correlations, TPD correlated with music regular accuracy and language irregular reaction time. Results are discussed in the next section.

### 6.2.5 Discussion

The current experiment focused on detecting subtle changes in notes and syllables—a crucial early step in determining larger scale syntactic structure in music and language domains—and how this perception was affected by cognitive entrainment to regular vs. irregular metrical contexts. In order to account for the role of individual differences, participants were grouped according to musical expertise and possessed each a VWM and TPD score (see Experiment 1a). Melody and sentence pairs differing only in harmonic or morphosyntactic nuance were discriminated; behavioral measures collected were accuracy and reaction time.

#### Detection of syntax-critical events in both domains

The task of participants was to detect whether melody and sentence pairs were 'same' or 'different,' thus they had to perceive subtle pitch differences or verb inflections which defined alternate syntactic structures, both of which were grammatically correct. This detection of structural difference proved to be easier for participants (irrespective of musical expertise) in the language compared to the music domain, in both accuracy and reaction time. This may be due to the greater everyday salience of linguistic verb inflection compared to tonal key modulation, for both non-musicians and musicians, and is in accordance with previous findings that phoneme perception is salient to native speakers at the pre-attentive level (Näätänen et al., 1997).

#### The role of metrical context

The current experiment largely supported the cognitive aspect of the current neurocognitive entrainment theoretical framework (Section 2.3), in that the regular meter context yielded better accuracy and faster reaction times, which indicated greater cognitive entrainment (or entrained attention) in these conditions. Un-hypothesized-for influences of domain and expertise, however, were found regarding entrainment to regular- compared with irregular meter.

Results from the music domain followed the hypothesis: the regular meter encouraged better critical-note detection (i.e. higher accuracy) than irregular meter. This finding extended music-domain evidence of cognitive entrainment to be based on metrical regularity, with human-performed temporal fluctuations; previous experiments reported evidence of entrainment to be based largely on temporal predictability (Jones & Boltz, 1989; Large & Palmer, 2002, although "human" temporal asynchronies were addressed in their model).

In reaction time, only musicians showed faster response in the regular-meter condition in music. Although both groups were hypothesized to reflect cognitive entrainment to regular meter in their reaction time, the result is understandable when one considers that a longer response time reflects more cognitive resources involved (Sternberg, 1969): the non-musicians, not usually faced with musical tasks, may have had to recruit a more general type of cognitive resource in order to complete their melody task, such that the 'benefit' of regularity due to entrainment did not surface among multiple engaged processes.

Non-musicians as a group discriminated verb inflections more quickly when presented in a regular-meter context, in accordance with the hypothesis that cognitive entrainment to the regular meter would be enhanced. This is in line with previous findings (in non-musicians) that entrainment can take place in the language domain (Lidji et al., 2011) and that entrainment to regular meter can influence linguistic processes (e.g. Cason & Schön, 2012; Magne et al., 2007; Schmidt-Kassow & Kotz, 2008). The finding moreover supports the attentional bounce hypothesis (Pitt & Samuel, 1990), which stated that temporal predictability among strong syllables (in this case, provided by more regular metrical structure) would enhance phonemic perception.

It now remains to be clarified why then the musicians, who should have better access to language meter due to generally enhanced perception of syllables (Marie et al., 2011; Tzounopoulos & Kraus, 2009), did not indicate entrainment to regular- over irregular meter. An explanation presents itself here which draws both on the concept of 'regularity' in language meter and on results from factor score correlations.

The current study pioneered meter in language stimuli which was comparable on a one-to-one basis to meter in musical stimuli, thus the definition of 'irregular meter' was rather a musical than linguistic one. While necessary to keep meter across domains comparable, the repeating two, three, or four syllable groups were, however, more regular than completely unstructured linguistic meter, containing for example more organization than present in spontaneous speech.

One aspect of the DAT—the backbone of the current concept of cognitive entrainment—is that the internal attentional oscillator is constantly self-correcting, allowing entrainment to take place despite real-time errors in timing (Large & Jones, 1999); it may be that musicians were more apt in this correction process, thus they were able to entrain to irregular sentences based on semi-regular organization of the meter. Indeed 'irregular' syllable groupings were based on three syllables within phrase and 12 syllables across phrase, give or take respectively one or two syllables. The production of these beat groups by the speaker may have been temporally predictable enough for musicians to utilize, if their error-correction were wide enough to encompass the (predictable) addition or subtraction of syllables.

An oscillator model of time perception (like the DAT) inherently suggests a link between interval discrimination (measured by the anisochrony detection test) and phase or period correction, as the same attentional oscillator which detects timing intervals self-corrects to adjust for timing errors (Large & Jones, 1999, see Section 2.1.1). Thus the cognitive resource behind TPD, which included anisochrony detection, may well also be linked with correction of timing errors. In line with this, musicians as a group possessed higher TPD scores (and TPD correlated with language irregular reaction time), further supporting the notion that musicians may have entrained to both metrical contexts in the language domain.

### Individual differences

**Musical expertise** It was appropriate that musicians reacted faster to all detection conditions (as hypothesized), as they had more immediate access to the auditory information compared to non-musicians, who needed more time to make their decisions. This result is in line with previous evidence, which reported that musicians have improved pre-attentive detection of auditory musical and linguistic events compared to non-musicians. ERP evidence has shown this to be the case in detection of syntactic violations in both domains (Fitzroy & Sanders, 2013; Koelsch et al., 2002), and auditory brainstem responses to speech signals showed musicians to be more accurate with regard to representing time and frequency information (Bidelman & Krishnan, 2010; Kraus et al., 2009).

The detection of a key change between melody pairs was easier for musicians compared to non-musicians, reflected by greater music accuracy among musicians. This was as hypothesized: Compared to non-musicians, musicians can perceive subtle musical events at the pre-attentive level whereas non-musicians do not (Koelsch, Schroeger, & Tervaniemi, 1999). Furthermore the key change was a subtle harmonic difference existing only in a melody (and not with full chord accompaniment), thus theoretical knowledge aided musicians in supplementing sensory cues as reported in previous findings (Besson & Faïta, 1995; James et al., 2012).

**Cognitive factors** While VWM (Factor 1) revealed no correlations with expertise or any of the experimental outcomes, TPD (Factor 2) showed itself to be involved in music regular accuracy, language irregular reaction time, and expertise.

Literature was presented in Chapter 4 that shows the involvement of working memory in retention/comparison of both full melodies and sentences (Fiveash & Pammer, 2014), therefore it is surprising that Verbal Working Memory had no significant correlation with Experiment 1b performance. Furthermore, where marginal correlations occurred they were with melody discrimination and not with sentence discrimination (which undoubtedly involves the phonological loop and therefore verbal working memory). It may be that somehow the cluster of diagnostic scores on the factor masked the participation of some sub-component of a cognitive resource, however this seems unlikely considering that in Experiment 1a, the first usage of the factor scores in statistical analysis was symmetrical to results of the test performed on all seven diagnostic scores. One potential explanation may relate to participant strategy in listening to the stimuli. The sentences were spoken at a fast pace (5 Hz syllable rate) and highly metrically structured in both conditions and not 'typical' spoken speech—perhaps there was a trade-off of resources performing the task, which moved the utility of Verbal Working Memory beneath some other resource, e.g. TPD or one not captured by these diagnostic tests.

TPD was shown to be the individual characteristic most involved in nuanced detection of syntactic events, compared with Verbal WM or expertise. The correlation of TPD with both melody and sentence performance suggested that this cognitive resource is domain general, which is in line with studies indicating that the perception of pitches is influenced by temporal processing, temporal accuracy of encoding is associated with improved perception of syllables, pitch perception transfers to language comprehension, and that rhythm ability transfers to language processing (Boltz, 1993; Kraus et al., 2009; Moritz, Yampolsky, Papadelis, Thomson, & Wolf, 2013).

Moreover, the resource underlying TPD may be involved in the correction of the internal attending oscillator described by the DAT. Evidence supporting this assertion included the TPD correlation with music regular accuracy and language irregular reaction time. Music regular accuracy was the condition to which everyone entrained and therefore the involvement of the resource would have played a role across all participants, taxing the timing error-correction via natural temporal fluctuations in melodies. Time and Pitch Discrimination was improved in musicians compared to non-musicians (described by the correlation with expertise, and see Experiment 1a), and the language irregular reaction time correlation may reflect that this resource was specifically involved in musicians entraining to the semi-regular language meter (see above). The significance of TPD will be further addressed throughout this dissertation, in light of results from Study 2 and Study 3.

## 6.3 Conclusion

This study addressed whether metrical context influences the detection of subtle syntactic events in melodies and sentences, and how individual differences play a role. Experiment 1a demonstrated two cognitive factors underlying seven diagnostic tests, named Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD). Experiment 1b saw a facilitative impact of regular metrical context on the detection of subtle syntactic events in melodies and sentences, indicating cognitive entrainment in these conditions, though in language the hypothesized impact was restricted to non-musicians.

### 6.3.1 Individual differences

Study 1 established metrics for individual-differences analysis which were successfully applied to other empirical results, thus they will also be employed in further studies in this dissertation.

Verbal Working Memory was not impacted by expertise, nor did it demonstrate a relationship to the detection of syntactic events in melodies or sentences. Although behavioral measures did not show involvement of this resource, EEG measures of entrainment (Study 2) or syntactic integration (Study 3) may draw on it.

Time and Pitch Discrimination was indicated to be a domain-general resource, and perhaps related to error-correction abilities of the internal attentional oscillator described by the DAT. Musicians had a higher TPD than non-musicians, but the factor correlated with performance independently from and beyond expertise; the brain resource corresponding to the tasks represented by TPD (melodic-, temporal- and rhythm discrimination) is therefore perhaps fine-tuned by musical expertise, but present in the non-musician healthy population as a natural aptitude which contributes to task performance. This underscores the importance of understanding individual differences when testing the impact of musical training on other aspects of psychological processing: simple categorical grouping into musicians and non-musicians may miss the influence of some third variable that may be involved in cognitive processing strategies across the whole population.

### 6.3.2 Metrical context and the detection of critical notes and syllables

The critical syntactic events in melodies and sentences were perceivable despite being non-violations, thus the stimuli intended to investigate syntactic expectancy in alternating metrical contexts were validated and will be employed further in Study 2 and Study 3.

Meter influenced the perception of syntax-critical notes in the hypothesized direction (namely, that performance was better when meter was regular) among all participants. This supports the presence of cognitive entrainment in music, according to components of the DAT: the regular meter context, conducive to entrainment or future-oriented attending, offers improved attention and consequently better detection of syntactic events compared to the irregular meter context, which spurs rather analytic attending (Jones, 1976; Jones & Boltz, 1989). Importantly, this finding extends the utility of future-oriented dynamic attending in syntactic processing from reliance on temporal regularity in music (Schmuckler & Boltz, 1994) to reliance on metrical regularity in the presence of temporal jitter.

Entrainment to linguistic meter manifested differently in musicians compared to non-musicians: While non-musicians showed the expected increased cognitive entrainment to regular meter, it is speculated here that musicians, due to improved temporal processing (indexed by TPD), can entrain even to 'irregular' linguistic meter, as long as it contains hierarchical patterns that are 'more regular' than spontaneous speech, as in the present stimuli. This suggests that the current neurocognitive entrainment framework has the potential to explain language perception as well as music perception, but that perhaps individual factors (related to or trained by musical expertise) alter the salience of metrical features in language.

The following study turns to the neural entrainment aspect of the current neurocognitive entrainment approach, exploring neural entrainment to meter in language and music among musicians, with frequency analysis of stimuli and musicians' EEG response to stimuli.

# Chapter 7

# Study 2: Spectral analysis of stimuli and EEG data

The goal of Study 2 was to evaluate melody and sentence stimuli for frequencies that represent different levels in the metrical structure, and then to look for signs of neural entrainment to these metrical levels in the EEG response of musicians. To this end, two experiments were conducted: Experiment 2a examined the frequency content of the stimuli and Experiment 2b evaluated the frequency content of the EEG recorded while musicians listened to those stimuli.

## 7.1 Experiment 2a: Stimuli spectral analysis

### 7.1.1 Introduction

As introduced in Chapter 2, meter—a hierarchical organization of strong and weak beats—exists in both music and language. Meter in music is established by a combination of temporal and melodic elements such as dynamic accents or contour change (Hannon et al., 2004), while the establishment of meter in language differs across disciplines: in poetry, meter is defined by the organization of weighted syllables in a given line of text (Fabb, 2001); in metrical theory, the meter is formed by flexible strong-weak patterns in speech (Prince, 1983); in prosody, meter is comprised of phonemically- and phonologically defined perceptual units (Nespor & Vogel, 1986).

Temporal predictability created by nested metrical levels in music enhances entrainment (e.g. Jones, 2008; Jones et al., 2002, see Chapter 2), and the predictable organization of stressed syllables has been proposed to guide attention in speech (Pitt & Samuel, 1990). Since similar metrical constructions afford cross-domain comparability with respect to a participant's entrainment to an auditory stimulus, the poetic definition of language meter was adopted in this dissertation due to its closeness to musical meter (e.g., the two have converged for centuries in song). An additional choice to facilitate comparable cross-domain, entrainment-relevant circumstances was to use human- (as opposed to computer-) generated melodies and sentences. This allowed for metrical construction to be independent of deviations in timing, avoiding that error-correction of natural non-periodic speech timing would constrain entrainment to language meter but not to 'perfectly timed' computer-generated music or vice versa.

Particularly since the stimuli were a naturally recorded signal, identification of metrical content, or the frequencies at which the nested metrical levels were occurring, was key in validating the stimuli. Identification of meter-representative peaks was also necessary for establishing hypotheses for EEG, or which frequencies to 'tag' in the EEG to determine whether participants entrained to these metrical frequencies. Therefore the aim of Experiment 2a was to identify the hierarchical metrical levels in the melody and sentence stimuli as they surfaced in the frequency domain. As multiple metrical levels in language poetry have been successfully identified in the frequency domain using an analysis of the speech envelope (Leonge, 2012), the amplitude envelope of both melodies and sentences was used here as the input to the spectral analysis.

#### Hypotheses

1. Frequency peaks which represent metrical structure should be present in the spectrum of both melody- and sentence amplitude envelopes: frequencies should surface that represent the weighted-only notes/syllables, or group level, as well as frequencies representing individual note/syllable, or individual level.

2. Amplitude of meter-representative frequencies should have most prominent energy in regular conditions, since weighted elements should align more across trials.

3. Spectra should be similar across domains if meter and speed are consistent.

### 7.1.2 Methods

#### Materials

Stimuli were 240 melodies and 240 sentences. The stimuli's meter (Figure 7.1) was a specifically designed hierarchy with a bottom *individual* level (representing eighth notes and syllables) and a higher *group* level (representing the naturally accented notes in melodic contour or naturally weighted spoken syllables). Stimuli were divided into *regular* and *irregular* metricality conditions (120 melodies and 120 sentences per condition), defined at the group level in the hierarchy. Regular items contained a predictably weighted note or

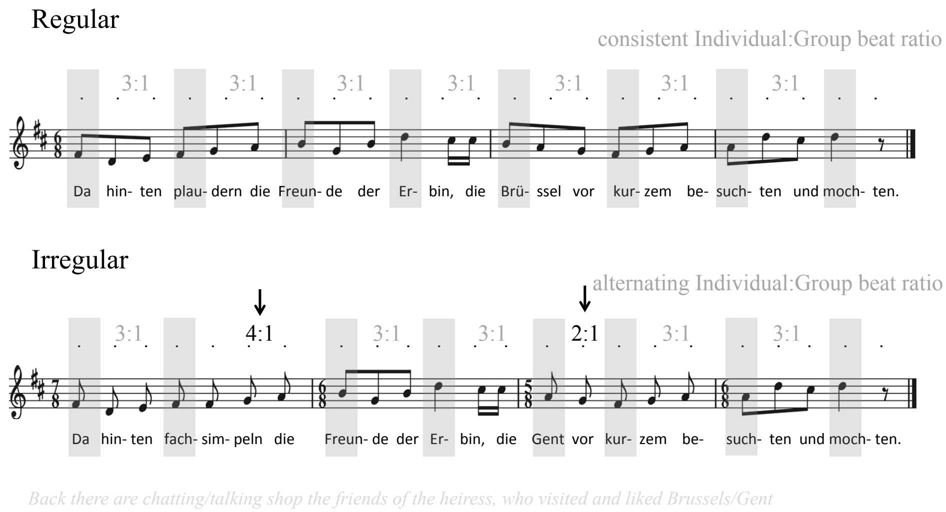



*Figure 7.1: Design. **Hierarchy**: Dots represent the beats in individual notes and syllables. The group level in the metrical hierarchy is comprised of weighted notes and syllables (shaded gray), and the individual level is comprised of all notes and syllables. **Metricality**: Regular is comprised of consistent groups of three beats while irregular displaces group boundaries (arrows) to create alternating groups of two, three, and four beats.*

syllable every three beats, and irregular items contained a less predictable weighted note or syllable every two, three or four beats. Thus the four stimulus conditions were: music regular, music irregular, language regular, and language irregular. Items in each condition contained group and individual metrical levels. A full description of stimuli and how they were recorded can be found in Section 5.3.2.

#### Procedure

Melodies and sentences were recorded in separate sessions. Melodies were played by a conservatory-trained pianist and recorded on an electric MIDI keyboard (44.1 kHz sampling rate). Sentences were spoken by a trained female native German speaker and recorded at 16-bit resolution and a sampling rate of 44.1 kHz. All items were post-processed with Audacity 1.3 and Praat 5.2, to normalize the peak amplitude (intensity) to 70 dB. The pianist and speaker synchronized their productions to a metronome to ensure temporal stability to the beat. Although performers listened to a metronome during recording, the sound of the metronome was not included in melody or sentence recordings. Section 5.3.2 provides more details about the recording process.

*Table 7.1: Meter-representative frequencies found in stimuli spectra per condition*

| Condition | Determined Group frequency (Hz) | Determined Individual frequency (Hz) |
|---|---|---|
| Music Regular | 1.7 | 5.1 |
| Music Irregular | 1.9 | 5.1 |
| Language Regular | 1.8 | 4.9 |
| Language Irregular | 1.8 | 5.3 |

### 7.1.3 Analysis and results

**Stimuli preprocessing**

Analysis was tailored for optimal future comparison to the frequency spectrum in EEG recordings, which were sampled at 500 Hz with a 0.4 Hz high-pass filter applied offline (see Experiment 2b). Audio recordings of the stimuli had an original sampling rate of 44100 Hz, therefore all stimuli items were up-sampled by a factor of five, low-pass filtered (143 Hz cutoff, zero-phase 10th-order Butterworth IIR) to avoid aliasing, and then down-sampled by a factor of 441 to reach 500 Hz sampling rate. Stimuli were finally high-pass filtered (0.4 Hz cutoff, zero-phase order 10th-order Butterworth IIR).

**Stimuli peak analysis**

The envelope was extracted using a Hilbert transform, and then an FFT was performed on the envelope of each stimulus item (500 Hz sampling rate, Hanning window applied to reduce spectral leakage, zero-padded to obtain a 0.0625 frequency resolution), and frequency-power output was standardized (mean: zero, standard deviation: one). Per condition, standardized power spectra were averaged over the 120 items. The standardized, averaged power spectra were evaluated for frequency peaks which would represent group and individual levels in the metrical hierarchy.

The metronome used by the pianist and speaker during recording facilitated a 1.7 Hz group frequency (as primary frequency for regular metricality, and as an average of group frequencies for irregular metricality) and a 5 Hz individual frequency (for both metricality conditions), thus the spectrum of each stimulus condition was evaluated for a peak between 1.2 and 2.2 Hz (hypothesized range for the group frequency) and between 4.5 and 5.5 Hz (hypothesized range for the individual frequency). The found group and individual frequencies are shown in Table 7.1. Spectra are displayed in Figure 7.2.

Once peaks were established, Wilcoxon signed rank tests compared power at peak frequencies to averaged power of the surrounding 1 Hz (16 averaged frequency bins, or 0.5 Hz in either direction), or ‘background,’ to see whether the power at the peak was significantly different from the background frequencies (Aiken & Picton, 2008). Results are listed in

*Table 7.2: Peak-frequency- vs. background-frequency power: related-samples Wilcoxon signed rank. Significant effects in bold (.05 $\alpha$ ) indicate a peak-power significantly greater than background power. Absolute effect size reported.*

| | Peak power median | Background power median | *Z* | *p* | *r* |
|---|---|---|---|---|---|
| Group | | | | | |
| MR | -0.14 | -0.18 | 8.39 | **<.001** | .82 |
| MI | -0.06 | -0.16 | 9.46 | **<.001** | .86 |
| LR | 0.06 | -0.14 | 9.48 | **<.001** | .87 |
| LI | -0.03 | -0.14 | 8.06 | **<.001** | .74 |
| Individual | | | | | |
| MR | 0.14 | -0.08 | 8.29 | **<.001** | .86 |
| MI | 0.19 | -0.08 | 7.76 | **<.001** | .71 |
| LR | -0.23 | -0.22 | 1.69 | .09 | .15 |
| LI | -0.15 | -0.18 | 6.13 | **<.001** | .56 |

*Table 7.3: 2 x 2 x 2 ANOVA: Stimuli peak power at meter-representative frequencies. GLM performed on log-transformed data. Significant effects in bold.*

| Comparison | *F(1,119)* | *p* | partial $\eta^2$ |
|---|---|---|---|
| Domain | 121.20 | **<.001** | .51 |
| Metricality | 0.07 | .80 | .00 |
| Hierarchy | 4.35 | **.039** | .04 |
| Domain x Metricality | 22.44 | **<.001** | .16 |
| Domain x Hierarchy | 273.63 | **<.001** | .70 |
| Metricality x Hierarchy | 0.23 | .63 | .00 |
| Domain x Metricality x Hierarchy | 10.91 | **.001** | .08 |

Table 7.2. Non-parametric tests were chosen because the power data were not normally distributed: Only three distributions were normal: music regular background, language regular background, and language irregular background. Otherwise all Shapiro-Wilk *W* statistics were ca. .9 and all $p$'s $< .01$.

Power at peaks was then compared across conditions in a 3-way ANOVA, factors Domain (music, language) x Metricality (regular, irregular) x Hierarchy (group, individual). Main effects of Domain and Hierarchy were found, as well as interactions Domain x Metricality, Domain x Hierarchy, and Domain x Metricality x Hierarchy (Table 7.3).

**Peak power main effects** For the main effects (Figure 7.3), power at peaks in music was generally greater than power at peaks in language, and power at group peaks was generally greater than power at individual peaks.

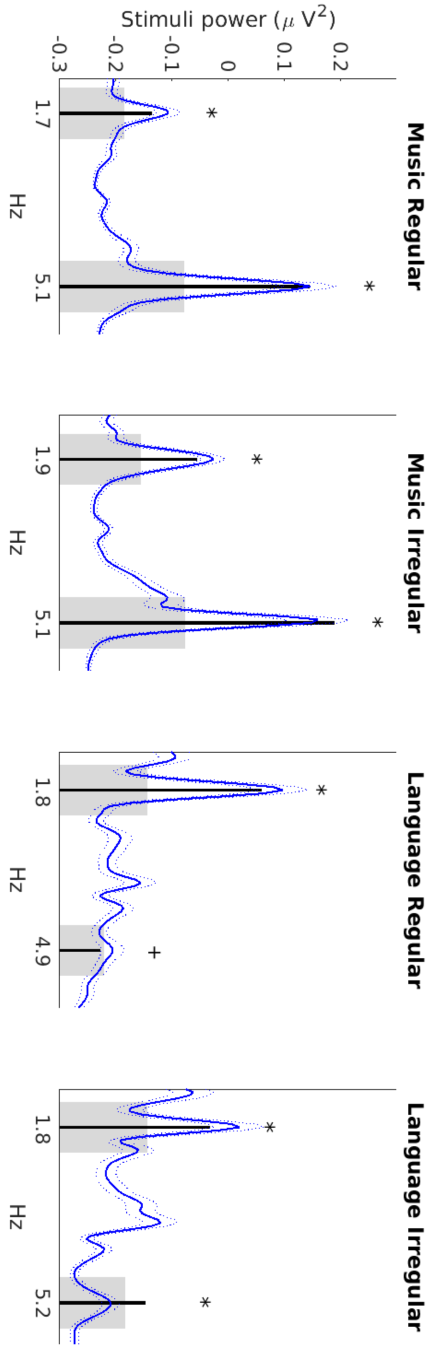



*Figure 7.2: Plotted average (solid) and standard error curves (dotted) of the frequency spectra for all four stimulus conditions. Wilcoxon signed rank related samples compared the median values of the peak frequencies (black lines) to the surrounding 1 Hz background (gray boxes) to determine peak significance ($p < .05^*$; $p < .1^+$). The x-axis indicates the frequencies (Hz) at which peaks represented group (left) and individual (right) metrical levels.*

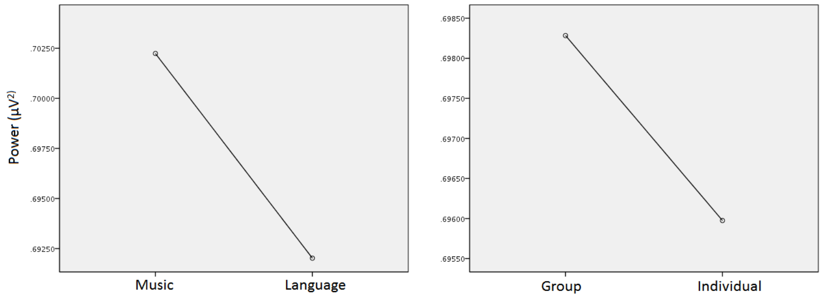



*Figure 7.3: Stimuli ANOVAs main effects.*

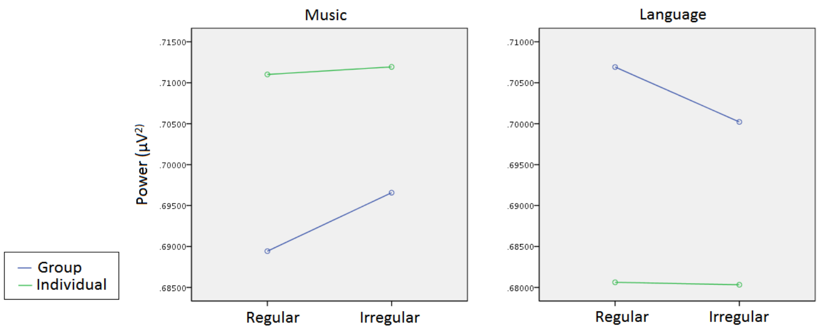



*Figure 7.4: Stimuli ANOVAs interactions.*

**Peak power interactions** To resolve the interactions of Domain x Metricality and Domain x Hierarchy, the 3-way ANOVA was split into two (music and language) 2-way ANOVAs, with factors Metricality x Hierarchy. Interactions are displayed in Figure 7.4.

**Domain x Metricality** While the metrical context impacted peak power in both domains, the interaction was in opposite directions: 1-way ANOVAs revealed that music regular had less power than music irregular ($F$(1,479) = 5.82; $p$ = .016), whereas language regular had greater power than language irregular ($F$(1,479) = 4.54; $p$ < .034).

**Domain x Hierarchy** While group and individual peaks had significantly different peak power in both domains, the interaction was in opposite directions: 1-way ANOVAs revealed that music group had less power than music individual ($F$(1,479) = 161.50; $p$ < .001), whereas language group had greater power than language individual ($F$(1,479) = 330.50; $p$ < .001).

**Domain x Metricality x Hierarchy** The Domain x Metricality x Hierarchy interaction echoed the above Domain x Metricality and Domain x Hierarchy interactions, showing that power at music group peaks was less than music individual peaks and that power at language group peaks was greater than at language individual peaks. The interaction was resolved by first splitting the 3-way ANOVA into respective music and language 2-way (Metricality x Hierarchy) ANOVAs. The music Metricality x Hierarchy interaction was significant ($F(1,119) = 6.18$; $p = .014$, partial $\eta^2 = .05$), subsequent 1-way ANOVAs showing that the music regular group peak power was less than the music irregular group peak power ($F(1,239) = 40.43$, $p < .001$) and that there was no difference between music regular individual and music irregular individual ($F(1,239) = 0.12$, $p = .73$). The language Metricality x Hierarchy interaction was significant ($F(1,119) = 5.82$; $p = .017$, partial $\eta^2 = .05$), subsequent 1-way ANOVAs showing that language regular group peak power was greater than language irregular group peak power ($F(1,239) = 8.52$; $p = .004$) and that there was no difference between language regular individual and language irregular individual ($F(1,239) = 0.09$; $p = .77$). Power means are displayed in Figure 7.4.

### 7.1.4 Discussion

#### Results summary

Music and language stimuli showed very clear peaks at frequencies representing both group and individual metrical levels. In music, both regular and irregular metricality conditions displayed significant peaks at the group and individual level. In language, both regular and irregular sentences had significant peaks at the group level, regular sentences had a marginally significant individual peak, and irregular sentences had a significant individual peak.

Intuitively, only at the group level did metrical regularity impact the peak power (as the metricality condition was defined at the group level). In music, regular melodies had a lower group peak, and in language, regular sentences had a higher group peak. At the individual level, peaks were generally larger in melodies than in sentences.

Within domain, the group peak power was less in regular compared to irregular melodies, and greater in regular compared to irregular sentences. Across domains, a bias for less energy at group compared to individual peaks was found in music, but more energy at group compared to individual peaks was found in language.

#### Stimuli peak power

The presence of low-frequency range peaks which represent levels in the metrical hierarchy validates the creation of the stimuli; using a metronome to control tempo in pianist and speaker successfully yielded a comparable, measurable beat in melodies and sentences.

**Group metrical level across domains** In melodies, the lower group peak in regular compared to irregular conditions was counter-hypothesis; this result may have been due to a particular playing strategy employed by the pianist, who made many more errors for the irregular compared to regular melodies during recording. The pianist was observed to strike harder at weighted notes (although not instructed to do so), simply as a way to engrain a motoric pattern for the unaccustomed structure (she was trained in the classical genre, which has fewer deviant metrical structures than for example jazz). This was not necessary for the regular melodies, which she easily played; thus greater peak at the irregular group frequency could have been due to greater attack velocity on irregular group-level notes.

The group peak was greater in regular compared to irregular sentences, as hypothesized: due to the definition of the metrical hierachy (weighted syllables comprised the group level), syllables at the group level were spoken with greater volume. Thus in regular sentences, where the weighted syllable was consistently every three beats, the averaged spectra showed a higher power at this single frequency, whereas in irregular sentences, the repeating syllable count was spread over two, three, or four syllables and therefore the volume spread over different frequencies.

**Individual metrical level across domains** Comparing the individual frequency peak power across domains, it is no surprise that individual (eighth note) peaks were stronger in music than in language. The pianist likely treated the individual level as a tactus, or a unit of time-keeping in music carefully maintained by conscientious musicians (such as our conservatory-trained pianist; see London, 2012). Although melodies contained a mixture of rhythmic values at the individual level, for example dotted eighths and quarter notes, and not exclusively eighth-notes, those varying lengths were by convention multiples or precise subdivisions, which would preserve the energy at the frequency of the individual level.

Syllables, on the other hand, varied in length depending on their consonant (and to a lesser extent, vowel) structure, and although 5 Hz is a repeatedly published frequency for syllable speech-rate (e.g. Greenberg et al., 2003), it makes sense that the frequency power would still be greater in music. Regarding why the regular sentence, individual frequency did not reach significance, where the irregular sentence, individual did reach significance, perhaps the speaker used less volume on weak syllables in the regular metricality condition because they were always weak: (e.g., ONE two three), whereas the shifted syllable weight in the irregular metricality condition (e.g., ONE two three FOUR, ONE two ONE) caused more attention to each syllable and therefore individual syllables were spoken on average with more volume.

**Implications for entrainment**

Meter-representative peaks found in melodies and sentences created an opportunity to measure entrainment with EEG: if the same peaks, or harmonics of the same peaks, are found ('tagged') in frequency-spectra of EEG, neural entrainment to the stimuli can be inferred. Three points with respect to entrainment emerge here: First, the greater music-individual compared to greater language-group peaks could mean that with respect to entrainment, in music the individual level is more salient while in language, the group level is the more salient metrical level. Second, although the group-level peak was smaller in regular as opposed to irregular melodies, the DAT (hence, the current neurocognitive entrainment approach, see Section 2.3) would still predict greater entrainment to regular as opposed to irregular melodies because the occurrence of group-level weighted notes in regular melodies was more temporally predicable (e.g. approximately every 600 ms) compared to the weighted notes in irregular melodies (e.g. approximately every 400 ms, 600 ms, or 800 ms). Finally, the fact that there was a group-level peak in irregular sentences at all points to there being at least some regularity in the metrical grid above a syllable-level, more than for example in everyday prose, in the current stimuli. This could imply the possiblity of (some) entrainment to irregular sentences, considering that the irregularly weighted syllable, though temporally on par with the irregularly weighted note, still adds much more temporal predictability than typically found in conversational speech.

In the current experiment, the relative height of the peaks may be traced to stimuli creation and recording, yet the role that peak height plays in terms of salience to actual entrainment is yet to be determined. The way musicians' low-frequency neural oscillations were entrained by these found peaks will be explored in the following EEG experiment.

## 7.2 Experiment 2b: EEG spectral analysis

### 7.2.1 Introduction

As presented in Chapter 2, temporally predictable streams of information are conducive to neurocognitive entrainment, and an auditory signal with a convergence of temporally predictable strong beats (e.g., nested metrical levels) should be more conducive to entrainment than a signal with less temporally predictable beats (e.g., Jones, 1976; Large & Jones, 1999). Complimentary to this, neural resonance theory proposes that the perception of meter is embodied by neural oscillations entrained to meter: perceived metrical levels correspond to the frequencies at which neurons "resonate" in response to a metrical stimulus (e.g., Large, 2008). Neural oscillations representative of meter perception may be found at stimulus frequencies and at harmonics and sub-harmonics of stimulus frequencies (e.g., 2:1

ratio; Tierney & Kraus, 2015).

Recent work capitalizing on the resemblance of neural oscillatory behavior and meter perception has shown that frequency tagging in EEG can reflect frequencies at which beats are perceived, both real and imaginary, and that the neurons also oscillate at harmonics and sub-harmonics of perceived meter frequencies (reviewed in Nozaradan, 2014). Accordingly, in music, entrainment to a drum beat super-imposed on the natural timing of musical meter has shown that EEG peaks and their 2:1 harmonic are representative of entrainment to the meter (Tierney & Kraus, 2015). In language, frequency tagging has shown entrainment to isochronous artificial syllables (Buiatti et al., 2009; Ding et al., 2016), but evidence of entrainment to nested metrical levels remains elusive.

Individual differences among musicians entraining to a piano melodies have recently been reported (Doelling & Poeppel, 2015), and entrainment to current stimuli was suggested in Experiment 1b to be indexed by cognitive factor Time and Pitch Discrimination (TPD). Thus here, individual differences among musicians entraining to meter-representative peaks may be reflected by TPD scores.

The stimuli evaluated in the previous experiment were comprised of conditions with more- and less- temporally predictable metrical downbeats (regular and irregular metricality conditions), and meter-representative peaks were found in the low-frequency range. Thus the current melody- and sentence stimuli are ideal media to test for cross-domain auditory entrainment in an EEG frequency-spectrum paradigm. The aim of Experiment 2b was to identify peaks in EEG spectra that would provide evidence of neural entrainment to the meter in current melody and sentence stimuli.

#### Hypotheses

1. Participants will neurally entrain to both melodies and sentences, to a greater extent in regular compared to irregular metricality conditions (e.g., Large, 2008; Tierney & Kraus, 2015; see Section 2.3; significant peaks will be found in the regular-condition EEG frequency spectra in both domains and be larger than peaks found in irregular conditions).

2. Based on peaks in Experiment 2a stimuli analysis, entrainment to individual frequencies will be greater in melodies, entrainment to group frequencies will be greater in sentences (larger peak height found at music-individual and language-group target frequencies).

3. The cognitive factor Time-Pitch Discrimination (obtained in Study 1) may play a role in musicians' capacity for entrainment (positive correlation with peak power).

### 7.2.2 Methods

#### Participants

Twenty-eight musicians (14 female) from Experiment 1 were recruited for this study and paid €7 per hour for their participation. Mean age was 24.8 (*SD* 2.5) years, mean age of musical training onset was 13.9 (*SD* 4.7) years, and mean weekly practice hours was 5.3 (*SD* 4.9). Participants were all right-handed native German speakers with sound hearing and a healthy neural history. No participants were pianists, to avoid motor-related responses to recorded piano melodies. Participants provided informed consent. The experiment was approved by the local ethics committee of the University of Leipzig. A list of participants musical instrument, age of musical training onset, and weekly practice time can be found in appendix C.

#### Materials

Stimuli for the EEG experiment were the recorded 240 melodies and 240 sentences analyzed in Experiment 2a (see Section 7.1.2). Filler items (additional 240 melodies and 240 sentences) had similar metrical composition as stimuli, i.e. half of them were regular and half irregular. Categorization of regular and irregular metricality conditions was defined at the group level. The filler regular condition was identical to the stimuli regular condition (weighted note or syllable every three beats), while the filler irregular-condition alternated the position of weighted beats. Thus fillers emphasized the predictability of regular-condition stimuli and prevented participants simply learning the stimuli's irregular-condition pattern. A full description of fillers may be found in sections 5.3.2 and 5.3.2). As part of Study 1, seven diagnostic tests were given to participants in a pseudo-randomized order: Forward and backward digit span (Tewes, 1994), modified listening span (Daneman & Carpenter, 1980), non-word repetition (Mottier, 1951), anisochrony detection (Dalla Bella et al., 2016), Musical Ear Test (MET) melody- and rhythm subtests (Wallentin et al., 2010). A full description of the tests may be found in Section 5.2.

#### Procedure

The melodies and sentences were aurally presented to participants while EEG was recorded in separate music and speech sessions. The sessions occurred on different days and the order of the sessions was counterbalanced across participants. Within sessions, Regular and Irregular items were presented in discrete blocks of approximately 6 minutes (60 items per block, 50% fillers).

Items were presented via Sony MDR-XD100 stereo headphones (Sony Corporation, Tokyo, Japan) at a comfortable volume while participants fixated on a white cross appeared on a black screen. They were instructed to blink between trials or during short self-paced breaks between the 6-minute blocks. Items were separated by a 400 ms inter-stimulus interval (ISI). Participants were asked to mentally answer a question regarding the syntactic content of each melody and sentence. After 10% of the items, a prompt indicated to provide their answer with a button press, in order to ensure their attention. Preparation and testing took approximately 90 minutes.

#### EEG recording

EEG was recorded with 64 Ag/AgCl electrodes placed in an elastic cap (Electro Cap Inc., Eaton, OH, USA) according to the extended 10-20 system (see Sharbrough et al., 1991) using a 24 bit Brainvision QuickAmp 72 amplifier (Brain Products GmbH, Gilching, Germany). Sampling rate was 500 Hz. Impedances were kept below 5 Ω throughout the experiment. Eye movements were monitored by bipolar horizontal and vertical electrooculograms (EOG) recorded from electrodes placed beneath the canthus of both eyes as well as above and below the right eye. Additionally, electrodes were placed on the left and right mastoid bones (M1 and M2) and a ground was placed on the sternum. The reference during recording was M1.

Diagnostic tests were collected separately in a session after EEG. Computerized, German versions of the diagnostic tests were presented at a comfortable volume via headphones in a pseudo-randomized order. Left- or right-correct responses, where applicable, were kept consistent within participant and were counter-balanced across participants.

### 7.2.3 Analysis and results

#### EEG preprocessing

In all participants, no more than 33% of the 120 trials were removed per condition. The average number of trials per condition that entered the analysis was 104 $\pm$ 37. With EEProbe (ANT neuro, Enschede, NL), EEG data were re-referenced to linked mastoids, 0.4 Hz high-pass filtered to remove excessive drifts (-3dB cutoff, fir, 3333 points, Blackman window applied) and then epochs with large artifacts were automatically removed whenever the standard deviation of the EOG exceeded 20 $\mu$V within a 200 ms gliding search window. Remaining visible artifacts were removed manually with the EEProbe graphic user interface. Further preprocessing and analysis was performed with FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011) in MATLAB 8.0 (The MathWorks, Inc., Natick, Massachusetts, United States). Independent component analysis was used to remove blinks and eye-movements, and five frontal electrodes, which were a source of noise in several participants, were removed (FPz, FP1, FP2, AF3, AF7).

### EEG peak analysis

Grand average ERP epochs (180 ms baseline plus 180 ms post stimulus-onset) were averaged across the four experimental conditions, and cluster-based permutations determined the electrodes where an auditory N1 was significant (compared to baseline, determined with cluster-based permutations). The 46 electrodes where the auditory N1 was significant were designated the region of interest (Figure 7.5).

At each electrode in the ROI, the time-domain epoch (per trial) used for frequency analysis was four seconds in duration, based on two considerations: First, auditory-stimulus-onset ERP effects, which extended to 460 ms on average, were avoided (thus the N1 which was used to define the ROI was not included in the signal to be analyzed). Second, the epoch avoided the ISI of averaged trials in any one condition, assuring that the stimuli were being heard throughout the entire duration of the epoch. The length of the shortest averaged stimuli condition—language regular (4.46 seconds)—was used to bound the analyzed EEG epoch for all conditions; the final four-second epoch (Figure 7.5) used for the spectral analysis began at 460 ms (after the P2) and continued until 4.46 seconds (shortest average stimuli length).

An FFT was performed over all ROI electrodes on the four-second epochs per trial (500 Hz sampling rate, Hanning window applied to reduce spectral leakage, zero-padded to obtain a 0.0625 frequency resolution), and frequency-power output was standardized (mean: zero, standard deviation: one). Per condition, standardized power spectra were averaged over trials and participants, and data were detrended after a negative slope (i.e., $1/f$ noise) was observed in all conditions. Peak analysis was performed on one spectrum per condition, which represented standardized, detrended data averaged over trials, participants, and ROI electrodes.

**Meter-representative frequencies** Per condition, power at hypothesized 'target' frequencies (determined from Experiment 2a, see above Table 7.1) was compared to averaged surrounding 1 Hz 'background' power with a related-samples test (Aiken & Picton, 2008). Since several conditions had non-normally distributed target frequency power (Shapiro-Wilk tests: music irregular group, $W = .89$, $p = .018$; music irregular individual, $W = .90$, $p = .023$; language irregular individual, $W = .88$, $p = .009$) and noise (Shapiro-Wilk tests: music irregular group, $W = .87$, $p = .005$; music irregular individual, $W = .81$, $p = .001$; language irregular individual, $W = .80$, p <.001), non-parametric Wilcoxon related samples signed rank tests were used (Table 7.4).

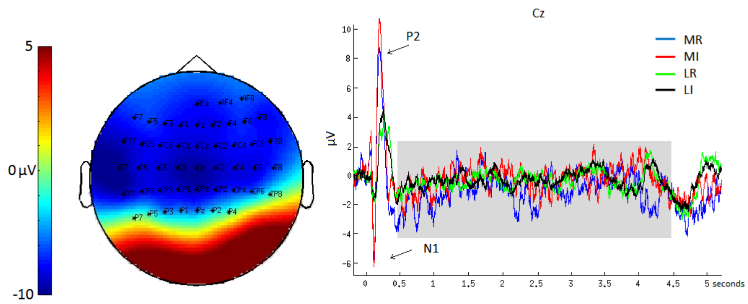



*Figure 7.5: Region of interest and time-domain epoch. The headplot shows electrodes with a significant auditory N1, used for the ROI in all conditions. Electrode Cz shows the epoch used for the spectral analyses is shaded in gray. The N1 and P2 components are indicated with arrows.*

Music regular and music irregular conditions contained peaks at target frequencies that were significantly greater than background power, which indicated that participants entrained to the stimuli at these frequencies. Both significant peaks represented the individual level in the metrical hierarchy (Figure 7.6).

No significant peaks were found at any of the group metrical levels, however; post-hoc analysis revealed significant peaks at 2:1 harmonics of the group frequencies in music-regular and language-regular conditions (Table 7.4), the very conditions predicted to have group-level entrainment. The music-regular peak was additionally a 2:3 subharmonic of the individual-level stimulus frequency, but considering that simpler ratios represent more stable entrainment (frequencies coupled in 1:1 ratio being the most stable, followed by 1:2 or 2:1; Large & Kolen, 1994), it is more likely that the music-regular peak was a 2:1 ratio of the group peak than a 2:3 of the individual. Peaks within 1 Hz range of the exact 2:1 ratio of the group peak were also evaluated in music-irregular and language-irregular conditions, but were not found to be significant. Evaluated peak frequencies are listed in Figure 7.6.

To mimic the ANOVA conducted on stimuli spectra, target-frequency-power was used in a 2 x 2 x 2 ANOVA with factors Domain (music, language), Hierarchy (group, individual) and Metricality (regular, irregular). The 3-way ANOVA reflected that music individual peaks were entrained most prominently: as a result, music was cumulatively more entrained than language (Domain: $F(1,22) = 14.39$, p =.001, partial $\eta^2 = .40$), and individual was more entrained than group ($F(1,22) = 4.35$, $p = .049$, partial $\eta^2 = .17$; marginal interaction of Domain x Hierarchy: $F(1,22) = 3.20$, $p = .087$, partial $\eta^2 = .13$). No other main effects or interactions reached the level of statistical significance (Metricality main effect: $F(1,22) = 0.01$, $p = .92$, partial $\eta^2 = .001$; Domain x Metricality: $F(1,22) = 2.12$, $p = .16$, partial $\eta^2 = .09$; Hierarchy x Metricality: $F(1,22) = 0.001$, $p = .97$, partial $\eta^2 = 0$; Domain x Hierarchy x Metricality: $F(1,22) = 0.01$, $p = .94$, partial $\eta^2 = 0$).

In order to elucidate the relationship of the group-harmonic peaks to the individual

*Table 7.4: Target- vs. background-frequency-power: related-samples Wilcoxon signed rank test. Significant effects in bold (p < .05), where target frequency power was greater than background frequency power (indicating a peak). In red, significant or near-significant findings where background frequency power was greater than target frequency power (no peak).*

| | Power median | | | | |
|---|---|---|---|---|---|
| | Target | Background | *Z* | *p* | *r* |
| Group | | | | | |
| MR | -1.00 | -0.98 | -1.80 | .07 | .38 |
| MI | -1.00 | -0.99 | -1.16 | .25 | .24 |
| LR | -1.02 | -1.00 | -2.25 | .02 | .46 |
| LI | -1.01 | -1.00 | -2.71 | .007 | .56 |
| Harmonic (2 x Group) | | | | | |
| MR | -1.00 | -1.01 | 3.38 | **<.001** | .70 |
| MI | -1.02 | -1.01 | 0.00 | 1.0 | .00 |
| LR | -1.01 | -1.02 | 2.04 | **.04** | .43 |
| LI | -1.00 | -1.01 | 1.58 | .11 | .33 |
| Individual | | | | | |
| MR | -0.97 | -0.99 | 3.35 | **.001** | .70 |
| MI | -0.98 | -1.00 | 3.07 | **.002** | .64 |
| LR | -1.00 | -1.00 | -1.25 | .21 | .26 |
| LI | -1.00 | -1.00 | -1.09 | .27 | .23 |

*Table 7.5: 2 x 2 x 2 ANOVA: EEG peak power with group-harmonic and individual peak values. GLM performed on log-transformed data. Significant effects in bold.*

| Comparison | $F(1,22)$ | *p* | partial $\eta^2$ |
|---|---|---|---|
| Domain | 4.87 | **.038** | .18 |
| Metricality | 2.09 | .16 | .09 |
| Hierarchy | 0.62 | .44 | .03 |
| Domain x Metricality | 2.13 | .16 | .09 |
| Domain x Hierarchy | 13.45 | **.001** | .38 |
| Metricality x Hierarchy | 3.65 | .07 | .14 |
| Domain x Metricality x Hierarchy | 0.01 | .92 | .001 |

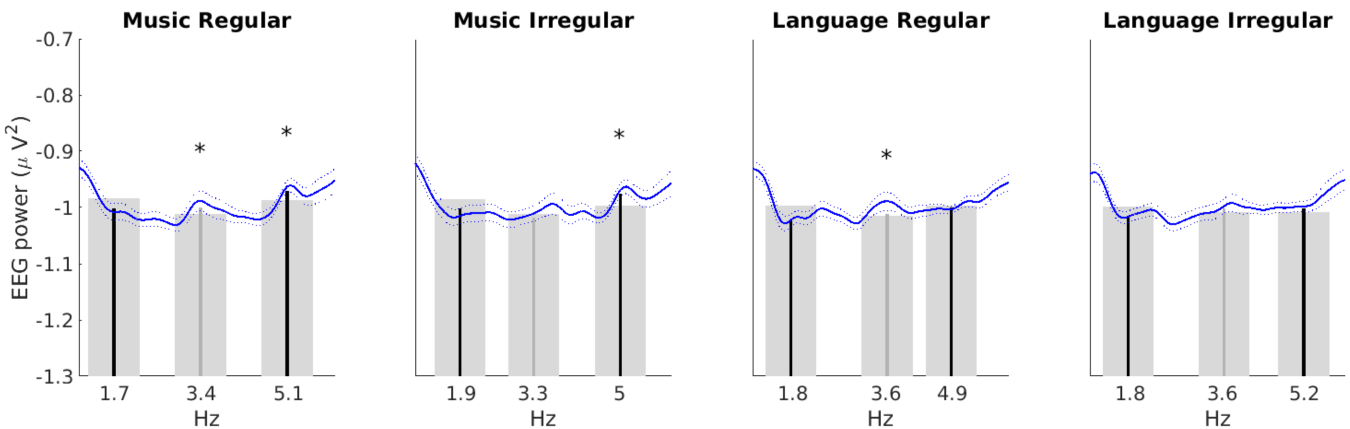



*Figure 7.6: EEG spectra. Each spectrum represents an average of 46 electrodes in the region of interest. Plotted average (solid) and standard error curves (dotted) of the frequency spectra for all four EEG conditions. Wilcoxon signed rank related samples compared the median values of the peak frequencies (lines) to the surrounding 1 Hz background (boxes) to determine peak significance ($p < .05$ *). The x-axis indicates the frequencies (Hz) at which peaks represented group (left), group-harmonic (center) and individual (right) metrical levels.*

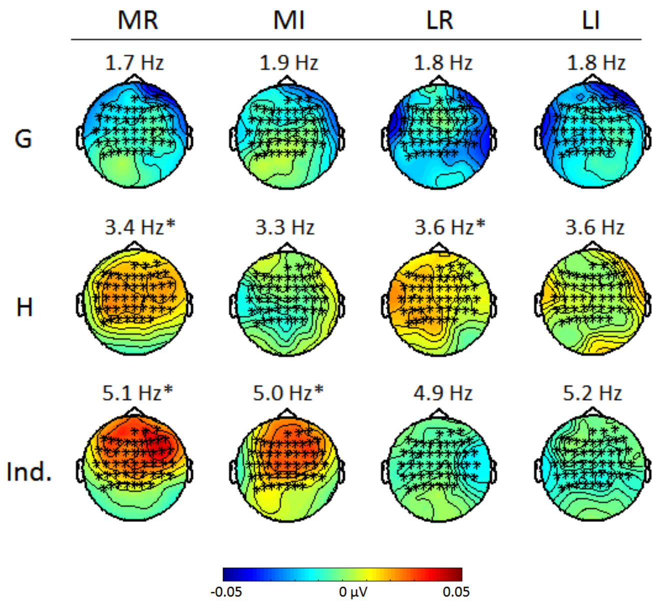



*Figure 7.7: Topography of target frequency power – background frequency power. M = Music, L = Language, R = Regular, I = Irregular, G= Group, H = Group-Harmonic, Ind. = Individual. Scalp topographies show 'peak power – background power' at group and individual target frequencies, ROI electrodes indicated with asterisks within plots. Plots representing significant peaks are indicated with asterisks (Hz*).*

*Table 7.6: Experiment 2b peak height correlations with factor scores. M = Music. L = Language. R = Regular. I = Irregular. H = Group-Harmonic. Ind. = Individual. No significant correlations (p < .05).*

| | Pearson's R (*p*) | |
|---|---|---|
| | VWM | TPD |
| MR-H | -.03 (.95) | .15 (.99) |
| LR-H | -.09 (.70) | .25 (.25) |
| MR-Ind. | -.36 (.09) | -.20 (.36) |
| MI-Ind. | -.12 (.58) | -.16 (.46) |

peaks, one further 2 x 2 x 2 ANOVA was conducted with the harmonic and individual frequency power. Since the power values were not normally distributed (Shapiro-Wilk $W$ =.79, $p$ <.001), log-transformed values were entered into the ANOVA. Results are listed in Table 7.5. The 3-way ANOVA with factors Domain (music, language), Metricality (regular, irregular) and Hierarchy (group-harmonic, individual) showed the relationship of the EEG group-harmonic peaks to the EEG individual peaks. Results were similar to the previous 3-way ANOVA, in that several effects were driven by the large music individual peaks. Music peaks were cumulatively larger than language peaks (Domain main effect), and in particular, music individual peaks were larger than language individual peaks: a Domain x Hierarchy interaction was resolved with 1-way ANOVAs, which showed that in music, group-harmonic peaks were less than individual peaks ($F(1,91) = 6.13$, $p = .015$) while in language, there was no significant difference between group-harmonic and individual peaks ($F(1,91) = 0.38$, $p = .54$). The large music individual peaks likely also drove a marginal interaction, that individual peaks were larger than group-harmonic peaks in the irregular conditions: a marginal Metricality x Hierarchy effect was resolved with 1-way ANOVAs, which showed that group-harmonic irregular peaks were slightly less than individual irregular peaks ($F(1,91) = 3.03$, p = .09) while there was no difference between group-harmonic regular and individual regular peaks ($F(1,91) = .18$, $p = .68$). There were no other main effects or interactions.

**EEG factor score correlations** Each participant had two cognitive factor scores from Study 1 of this dissertation, which had been obtained from a factor analysis of diagnostic tests on a larger participant sample. The cognitive factors, Verbal Working Memory (VWM; clustered diagnostic values for forward and backward digit span, modified listening span, non-word repetition) and Time and Pitch Discrimination (TPD; clustered diagnostic values for anisochrony detection, MET melody, and MET rhythm), were used here to see whether peak height corresponded with these measures of individual difference.

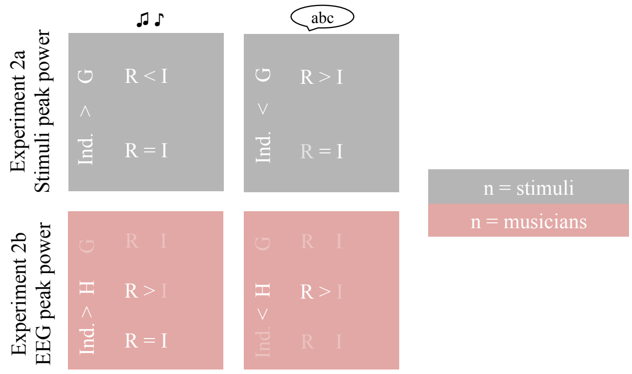



*Figure 7.8: Summary of the Study 2 spectral analyses. Ind. = Individual, G = Group, H = Group-Harmonic, R = Regular, I = Irregular. Translucent letters represent non-significant effects. See text for details.*

To see if the cognitive factors impacted the strength of entrainment, the peak height (peak power — background power) for the four significant peaks (music regular harmonic, language regular harmonic, music regular individual, music irregular individual) was correlated with each participant's Time and Pitch Discrimination and Verbal Working Memory cognitive factor scores (Table 7.6). No correlations were significant.

### 7.2.4 Discussion

#### Results summary

Brain data were evaluated in a fronto-central ROI, using a method similar to stimuli analysis. Neural entrainment to the group level in the metrical hierarchy occurred, represented by the 2:1 harmonic of the group in the hypothesized music-regular and language-regular conditions. Moreover, there was no significant difference between the harmonic resonance peak power in music compared to language, indicating that where metrical hierarchy in the stimulus is comparable, entrainment to metrical hierarchy is comparable. A different pattern emerged at the individual level. Clear music-individual peaks indicated entrainment to the eighth note beat, independent of metricality, whereas no language-individual peaks emerged to represent entrainment to the syllable. Cognitive factors did not play a role in the height of significant peaks. No other significant peaks in the EEG frequency spectra represented metrical levels in the stimuli. Results are summarized in Figure 7.8.

#### Entrainment to stimuli

**Group metrical level** The exact 2:1 harmonic of the group level, moreover in its hypothesized regular conditions, was found in both domains. The presence of harmonic peaks indicates that coupling to the signal was strong enough to produce the closest harmonic, since the amplitude of resonance peaks is proportional to the coupling strength of the originally entrained frequency, from simple to complex ratios (e.g., original 1:1 frequency entrainment would have the highest amplitude, and next highest would be 2:1 resonance, then 3:1 resonance, and so on; see Large & Kolen, 1994). The 2:1 harmonic of the regular-group conditions is reminiscent of a recent EEG finding where the 2:1 harmonic of an entrained 'on-beat' frequency indicated enhanced entrainment compared to the entrainment of an 'off-beat' frequency, where no harmonic was produced (Tierney & Kraus, 2015). Thus the current regular metricality condition, with predictable metrical organization, can be said to have enhanced neural entrainment compared to the less predictable irregular metricality.

The two group-harmonic peaks were statistically indistinguishable in power, despite the fact that in the stimulus, the language-group peaks were higher than the music-group peaks: thus, represented by comparable attenuated resonance peaks, neural entrainment to the upper metrical level was comparable across domains even though the physical prominence of the upper metrical level was different in the two domains. Nozaradan et al. (2012) showed this too, that the brain up-regulated salient metrical frequencies beyond their comparative amplitude presence in the stimulus. Here, the comparable music and language group-harmonic EEG peaks strengthen the notion that metrical organization plays a salient, orienting role for neural oscillations in the delta range, regardless of its domain, consistent with the current proposed account of neurocognitive entrainment (see Chapter 2).

In light of current literature (e.g., Large, 2008; Large & Snyder, 2009; Tierney & Kraus, 2015), the equivocal group-harmonic peaks in the current data indicate neural entrainment to the meter, as well as comparable perception of the upper metrical level in melodies and sentences (despite differences in entrainment to the individual level). Unfortunately, the weak spectral information below 2 Hz obscured potential findings of a weaker entrainment to the irregular metrical conditions that produced no harmonic; presumably, the musicians also perceived the irregular metrical levels, which according to neural resonance theory should be reflected by neural oscillations at the frequency of the perceived level. While current findings indicate stronger entrainment to regular melodies and sentences compared to irregular ones, it does not necessarily exclude a less-strongly coupled entrainment to irregular melodies or sentences, since the 1.7 Hz range fell below the threshold of analysis.

An explanation for the presence of a 2:1 harmonic in the absence of evidence of its original frequency entrainment is that the lower stimulus frequency dropped below the threshold that the frequency analysis was capable of measuring due to shortness of stimuli, while the 2:1 harmonic, like the individual-level frequency, remained above that threshold. Since harmonics of entrained frequencies have been previously reported in the literature for both perception of metrical levels in tone sequences (Nozaradan et al., 2011) and superimposed drum beats (Tierney & Kraus, 2015), and perception of a tactus in musical excerpts (Doelling & Poeppel, 2015), this current result is a reliable indication of neural entrainment.

**Individual metrical level** Current results at the individual frequencies in EEG response to melodies and sentences indicated that musicians neurally entrained to the amplitude envelope of the eighth note but not to the envelope of the syllable. The music data support claims of previous EEG and MEG evidence of entrainment to musical tactus (e.g., Doelling & Poeppel, 2015; Nozaradan et al., 2011); the language data also resemble previous EEG frequency-tagging results that showed a disappearance of entrained syllable rate once syllables were grouped into perceptual units (Buiatti et al., 2009).

In the current EEG experiment, musicians clearly entrained to the individual level of melodies, regardless of metricality condition, consistent with previous paradigms showing entrainment to a steady pulse (Nozaradan et al., 2011, 2012) or musical tactus (Doelling & Poeppel, 2015). The power at EEG individual peaks was not significantly different between metricality conditions, which should logically be the case; music-individual peaks in stimulus analysis did not differ, nor was metricality hypothesized to affect the individual level (but rather the group level, because the weighted note or syllable was either regularly or irregularly positioned).

In language, the lack of an individual frequency peak, which represents the individual syllable, is in line with previous evidence from a word-segmentation study (Buiatti et al., 2009). Buiatti et al. (2009) showed entrainment to randomly sequenced artificial syllables, presented isochronously, as demonstrated by a spectral peak at the syllable presentation rate. Interestingly, once the syllables were grouped into pseudo-words according to transitional probability and pauses, the spectral peak at the syllable rate disappeared and instead a peak emerged that represented the beginning of syllable groups (learned psuedowords). Since in Buiatti et al. (2009) the syllables were isochronously presented, the current language-individual result can be seen as a replicated perceptual phenomenon that is related to salient boundaries among syllables, and not to the precise timing of syllable onsets. Thus comparing the perception of musical meter to linguistic meter, it would seem that while predictable grouping of metrical constituents is comparably entrained across domains, nevertheless the lowest metrical level at the individual note or syllable rate has comparably different salience in music as compared to language .

In light of previous evidence, the current result may therefore be interpreted such that in music, the individual level remains salient in addition to higher metrical levels, while in language the salience of the individual level is minimal compared to higher metrical levels. This discrepancy between the domains is an interesting finding on its own, since it shows dissociation in the way the metrical levels in music are processed compared to the way metrical levels in language are processed, and would be an interesting topic for future research. For now, the issue will be further examined from the aspect of how the neural entrainment to the group metrical level impacts syntax processing in Study 3; if meter interacts with syntax processing in both domains, then the greater cognitive salience for entrainment lies with metrical organization (since entrainment to group occurred in both domains) and not the beats represented by individual notes and syllables.

The current frequency tagging method is useful in this case to compare meter perception across domains, and while the individual condition results indicate that the brain attended the syllable qualitatively differently than musical eighth note (and boundaries of syllable groups), the finding is not conclusive evidence that the syllable was not entrained. Indeed, naturally spoken sentences have previously been used to show entrainment to the syllable rate with individual trial classification based on phase information (Luo & Poeppel, 2007, described in Section 2.4.2). It is possible that trial classification is a more robust method to measure entrainment to syllables in sentences, and that syllable entrainment did occur, however; considering the necessary diverse lexical content of our stimuli for a valid syntax experiment (see Section 5.3 and Study 3), such a trial classification approach (with one sentence repeated per condition, compared to our 60 different sentences per condition) was computationally inappropriate for the current study. Thus the current interpretation is limited to how the brain entrains to metrical levels in music compared to language.

**Group compared to Individual** The ANOVA that compared the group-harmonic peak power to the individual peak power showed that, contrary to visual inspection, there was no difference in peak power when comparing the height of group-harmonic to individual peaks in either music or language. This result may be misleading, and must not be definitively interpreted such that entrainment was not different across metrical levels. According to the model proposed along with neural resonance theory (Large & Snyder, 2009), the harmonics of entrained frequencies are attenuated compared to the original entrained frequency; the original 1:1 entrained peak has the highest amplitude, and 2:1 or 1:2 ratio harmonics have comparably attenuated peaks, followed by 3:1 or 1:3 ratio harmonics with further attenuated peaks, and so on, with attenuation increasing as the ratio of harmonic to original becomes more complex. Thus, the original entrained group frequency (which fell below the current threshold of analysis) should have a higher amplitude than its 2:1 resonance in both music and language conditions. Therefore, it is not known from the current study whether the participants entrained more to the individual or to the group metrical levels, since the original group frequency power could still be significantly greater than the individual frequency power in either domain. Future studies need to clarify this, for example by presenting melodies with group metrical levels at higher frequencies.

**Individual differences**

The cognitive factors Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD), obtained in Study 1, did not demonstrate any statistical relationship with the peak power of regular-group-harmonic frequencies or music-individual frequencies, despite suggestions from the literature that especially the resources named in TPD—temporal, rhythmic, and melodic discrimination—are resources that should aid entrainment to meter (see Chapter 4).

This null-result, while it could be evidence that current cognitive factors are not involved in neural entrainment to musical beat in a lower metrical level, could alternately be explained such that the current experiment was simply unable to measure individual differences among musicians when they entrained to the musical beat.

If the current musicians in fact did not use cognitive resources associated with either VWM or TPD to entrain to music-individual frequencies, this could be a result of training-related strategy. Such a phenomenon has been reported previously, when comparing professional and amateur chess players (Chase & Simon, 1973). Professional players were postulated to have a rote response to basic performance requirements and only recruit active resources for more difficult or novel situations. Thus, like chess players, these musicians may have developed a distinct skill set for entraining to the lower metrical level that by-passes typically associated cognitive resources when listening to an 'easy' eighth note beat.

Alternatively, the cognitive factors, particularly TPD, could have been involved in neural entrainment but simply not present in the results of the current analysis because the scores were not dispersed enough (to demonstrate a linear relationship with EEG power) among these musician-only participants. When the diagnostic scores were originally clustered onto the factors (Experiment 1a), musicians scored altogether higher in TPD than non-musicians, supporting the possibility that musician scores were too clustered in the current analysis.

Another possibility is that TPD influenced the entrainment to the regular-group metrical level in both domains, but that the attenuated resonance frequency, measured by the group-harmonic peak power, did not reflect the relationship. For example, the coupling strength of the entrainment to the group metrical level, which determined the amplitude of the 2:1 resonance (see Large & Kolen, 1994), could have varied among participants in a relationship nonlinear to TPD or VWM, or according to some third variable such as years of musical training.

Concerning this latter possibility, hypothesis-directed testing with additional measures, to assess individual entrainment to higher metrical levels, can be a valuable inclusion to future studies with respect to entrainment to group level in metrical hierarchy of melodies (or sentences), however; it is clear from the current results that the main priority of such further investigation into neural entrainment to higher metrical levels should be to increase the power of the EEG frequency spectrum in the range which would represent the higher metrical level, via longer stimuli and/or larger participant groups than used in the current study.

## 7.3 Conclusion

This study examined the frequency spectra of melody and sentence stimuli and the EEG response of musicians to the stimuli, in an effort to establish which frequencies might represent group and individual metrical levels, whether musicians neurally entrained to the meter, and whether individual differences played a role in neural entrainment. Experiment 2a found hypothesized meter-representative frequencies in melody and sentence spectra, and Experiment 2b demonstrated that, as hypothesized, musicians entrained more to the group metrical level in regular compared to irregular metricality conditions, in both domains. Additionally, entrainment occurred to music-individual (the eighth note) but not to the language-individual (the syllable) conditions. A null result for individual differences in significant peak height (measured with cognitive factors VWM and TPD) was attributed either to the closely clustered ability of musicians to entrain or to the unreliable transfer of individual differences from the originally entrained frequency (which fell below the threshold of analysis) to the attenuated resonance peak.

### 7.3.1 Stimuli spectral analysis

The current stimuli deliver what they promise: two metrical levels, one representing the group level (weighted notes and syllables) and one representing individual level (individual eighth notes and syllables). Thus the paradigm is further validated, that if care is taken to construct analogous metrical structures in music and language, the acoustic signals will manifest comparable metrical levels.

### 7.3.2 EEG spectral analysis

Musicians were able to neurally entrain comparatively more to regularly grouped beats compared to irregularly grouped beats, shown by the presence of the 2:1 harmonic of the regular group metrical level, in both music and language. This cross-domain entrainment to metrical hierarchy is a landmark finding in two aspects. First of all, the current finding is the first to report neural entrainment to hierarchical metrical organization in language, showing that language meter is sufficient to couple neural oscillations although its timing is traditionally less periodic than in music meter. Secondly, the resonance peaks were statistically comparable across domains, despite amplitude differences in melody and sentence group frequencies in stimuli. Thus it would seem that the brain utilizes metrical organization independent of domain, entraining to the hierarchical grouping of metrical constituents in human-produced sound streams.

Now that neural entrainment to the upper metrical level in regular-meter melodies and sentences has been shown, the next study will investigate whether this entrainment has an impact on syntax processing in both domains. A cognitive repercussion of neural entrainment to meter can be determined by the presence or absence of a meter–syntax interaction in the following Study 3 ERP experiments.

### 7.3.3 Individual differences

The individual-differences null result from Experiment 2b may have been a consequence of this dissertation's intention to study syntax perception across domains—musicians are the ideal participants for cross-domain syntax perception, particularly the kind of non-violation used here (because one may presume syntactic representation in music), but they are not the ideal test group when attempting to index individual differences in entrainment to a musical beat. Although an attempt was made to get a 'spread' among musicians by using amateurs with varied practice time, Time and Pitch Discrimination and their musical ability may have been too concentrated on the higher end of performance to measure differences.

Although neither cognitive factor surfaced here in the direct role of neural entrainment to the stimulus, one or both may still play a role in syntactic integration of the current melodies and sentences: individual differences among these musicians with respect to their integration of syntactically non-preferred melody and sentence endings, in regular and irregular metricality contexts, will be explored in Study 3.

# Chapter 8

# Study 3: Does regular meter facilitate syntactic integration across domains? An ERP study of melodies and sentences.

Study 2 has just shown that the melody and sentence stimuli contain comparable metrical levels, and that musicians neurally entrain more to regular than to irregular hierarchical metrical structure in both domains. Study 3 aimed to investigate the impact of that entrainment on the cognitive process of syntactic integration in melodies (Experiment 3a) and sentences (Experiment 3c). A necessary prerequisite to the sentence ERP experiment was to empirically determine native speaker's preferred ending in current sentence stimuli (Experiment 3b). Additionally, and despite that no individual differences were found in entrainment to meter in Study 2, this study will examine whether individual differences play a role in syntactic integration.

## 8.1 Introduction

The introduction laid out a groundwork for how neurocognitive entrainment (Chapter 2) should impact syntactic comprehension (Chapter 3) in the music and language domains. Restated briefly, metrical organization can guide attention (Large & Jones, 1999; Large & Palmer, 2002; Pitt & Samuel, 1990) to important syntactic events (Lerdahl & Jackendoff, 1983; Selkirk, 2011). Moreover, the temporal predictability provided by a perceived meter can facilitate neural synchronization to the signal (e.g., Large & Snyder, 2009), which perpetuates more synchronization in subsequent dynamic neural communication (e.g., Singer,

2013) associated with cognition (e.g., Varela et al., 2001), including hierarchically phase-coupled neural networks in delta, theta and gamma frequency bands (Giraud & Poeppel, 2012; Schroeder & Lakatos, 2009). Consistent with the assumptions of this approach, Study 2 has just shown that musicians entrain more to the regularly composed meter than to the irregularly composed meter in both domains. The current study addressed the second part of the neurocognitive entrainment hypothesis (within-participant), that cognition should be facilitated more in a regular compared to irregular metrical context when processing the same syntactic manipulation (e.g., Cason & Schön, 2012; Roncaglia-Denissen et al., 2013; Rothermich et al., 2012).

The syntactic manipulation used here, preferred versus non-preferred structural resolution to alternate syntactic structures (see Section 5.3), was intended to test the cognitive process of syntactic integration, with two particular motivations. First, a contemporary theory hypothesized syntactic integration mechanisms to be shared in music and language domains (SSIRH; Patel, 2003; see Section 3.1.3), thus syntactic integration was ideal when comparing the impact of entrainment on music and language cognition. Second, the current stimuli were correct in both preferred and non-preferred syntactic structures, in an attempt to avoid error-detection mechanisms that may have confounded cross-domain comparison in previous music-language syntax paradigms (e.g., Patel et al., 1998; see Section 3.1.3). It was therefore expected that the non-preferred syntax condition should elicit similar ERPs in both domains (a P600; e.g., Besson & Faïta, 1995; Osterhout & Holcomb, 1992).

Crucial to the current approach, syntactic integration in both domains should be similarly impacted by the metrical context in which it occurs. While studies addressing the impact of musical meter on syntax processing have been scarce (but see Section 2.4.3, which addresses the confound of timing and meter in music cognition studies), event-related potential studies in the language domain have shown linguistic processes to be facilitated by a regular meter context, marked by a reduced ERP effect in regular compared to irregular metrical contexts[1] (Cason & Schön, 2012; Roncaglia-Denissen et al., 2013; Rothermich et al., 2012). Thus the current study expected to replicate this effect of metrical regularity, in showing a reduced syntax P600 ERP effect in regular compared to irregular metrical contexts.

Study 3 also addresses individual differences in auditory perception. The diagnostic tests used in Study 1, which provided cognitive factor scores Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD) to the current participants, were chosen for their relevance to musical and linguistic meter and syntax perception (see Section 5.2). Considering that syntax ERPs were previously reported to be affected by individual

[1] But processes such as deviance detection have been marked by larger components in a temporally predictable context (e.g., Schwartze et al., 2011)

differences (e.g., Besson & Faïta, 1995; Bornkessel et al., 2004; see Chapter 4), it was hypothesized here that factor scores VWM and TPD should correlate with syntax ERPs in both domains.

One final point addressed by Study 3 was to determine how native German speakers prefer the current sentences to end, with a high- or low-attached relative clause, since reports of auditory modality relative clause attachment preference have been varied in German (as opposed to consistent visual modality reports; Augurzky, 2006; Hemforth et al., 2000; see Section 3.1.2). An exploratory behavioral experiment was presented before the language ERP experiment (with a separate participant group from the ERP experiments) to address the native preference for the current sentence stimuli, which served to inform the language ERP coding for preferred and non-preferred conditions.

This study was the first of its kind to assess within (-musician) -participant integration of non-preferred syntactic structures across domains, taking into account the respective roles of metrical context and individual differences. Two separate experiments addressed syntax ERPs in music (Experiment 3a) and language (Experiment 3c), in regular as opposed to irregular metrical contexts. Native relative clause attachment preference in the current stimuli was addressed in Experiment 3b.

## 8.2 Experiment 3a: ERP responses to melodies with varying metricality and syntax

### 8.2.1 Introduction

A domain-general ERP component called the P600 has been elicited in music with respect to integration of violated harmonic structure (e.g., Featherstone et al., 2013; Patel et al., 1998) and within-key unexpected endings (Besson & Faïta, 1995), and it was expected to be elicited in Experiment 3a when musicians were confronted with syntactically non-preferred (but not incorrect) melodic endings.

Although evidence supporting neural entrainment to musical meter has been established in the music domain (e.g. Nozaradan et al., 2012; Tierney & Kraus, 2015; see Study 2), there are no ERP studies which have investigated whether music perception, in particular music syntactic comprehension, is facilitated by (entrainment to) a regular metrical context (but behavioral paradigms suggest this indeed to be the case: harmonic processing is better in temporally predictable musical passages, e.g. Schmuckler & Boltz, 1994; see Section 4.2.2 of this dissertation). Experiment 3a investigated syntactic integration in the context of regular and irregular metricality conditions, which were shown with the same participants in Study 2 to have comparable entrainment to the individual level but comparatively more entrainment to the higher metrical group level in regular conditions. Thus, any interaction of the metricality and syntax conditions would show music cognition to be impacted by hierarchical metrical organization.

While previous studies have shown the later music-syntactic integration processes, represented by the P600, to be impacted by musical expertise (e.g., a larger amplitude in musicians compared to non-musicians; Besson & Faïta, 1995; Featherstone et al., 2013), no studies have yet addressed how individual differences in cognitive ability among musicians may impact syntax-related processing. Experiment 3a therefore monitored individual differences in Verbal Working Memory (cognitive factor VWM) and Time and Pitch Discrimination (cognitive factor TPD) with respect to music syntax ERP amplitude.

#### Hypotheses

1. A P600 should be elicited in the non-preferred syntax condition.
2. Syntax should interact with metricality (regular-meter context should facilitate syntactic processing, resulting in a reduced P600 in the regular-meter condition).
3. Individual differences should play a role in syntax ERPs (factor scores would correlate with P600 amplitude).

### 8.2.2 Methods

#### Participants

Music EEG data were those recorded in Experiment 2b, that were re-analyzed in this experiment with focus on syntax. Thus, participants were twenty-eight right-handed, monolingual native German speakers from Study 2 (see Section 7.2.2), whose seven diagnostic scores were previously collected (Study 1). Methods relevant to this ERP analysis are listed below.

#### Materials

Melodies (60 templates with four versions each) contained an identical first phrase with an established major key and a second phrase with either the same major key or a modulation to the relative minor key. The final syntactic structure (whether the key remained constant or was modulated to the relative key in the second phrase) was first detectable at the antepenultimate note, which was consistently the tonic of the melody, followed by repetition of the tonic via a final cadence. The two initial syntax conditions were thus major or minor, subsequently coded as *preferred* and *non-preferred* based on a post-hoc questionnaire (see

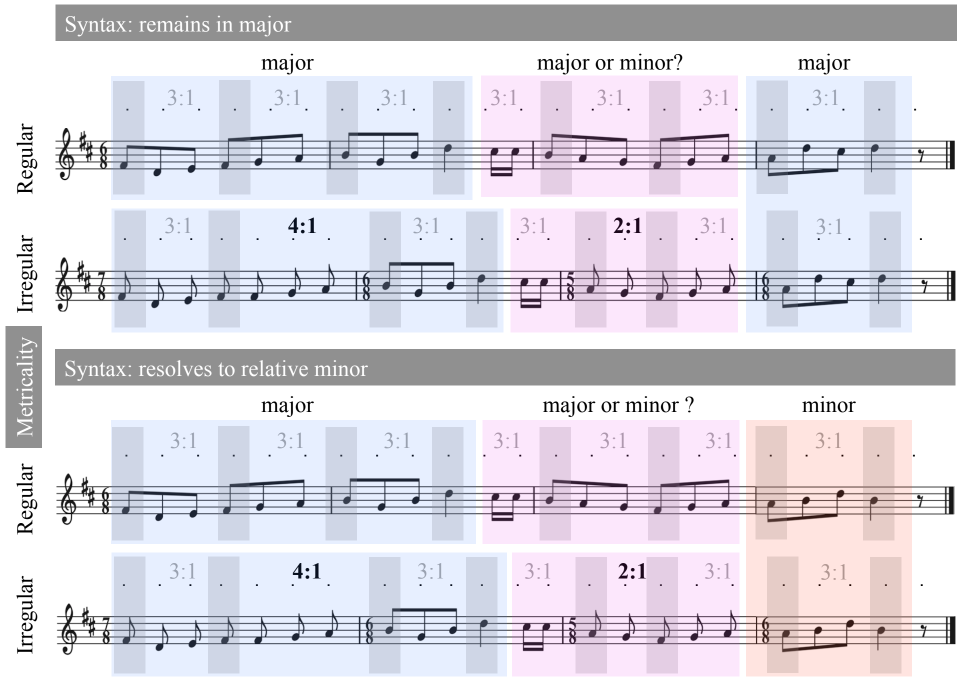



*Figure 8.1: Music stimuli design.* ***Syntax****: Each melody is two phrases across four measures, the first phrase clearly in a major key, the second phrase (at the 16th-note pick-up) ending alternatively in major or relative minor, resolved consistently in anti-penultimate and final notes. Each syntax condition has a regular and irregular* ***Metricality*** *version, defined by the ratio of weighted beats (shaded gray) to tactus beats (dots).*

Section 8.2.3). In each of the two metricality conditions (regular and irregular), melodic contour was built into the two versions of the meter such that no harmonic irregularity was present. Example melodies are shown in Figure 8.1. See Section 5.3.2 for a full stimulus description and Section 3.1.1 for a review of relative keys.

As mentioned in Section 5.3.2, the filler items (240 additional melodies) had metricality conditions similar to the experimental stimuli described above. Syntactic conditions in fillers were not identical to melodies but fit the general description of a first phrase with an established key, and a potential key modulation in the second phrase. The two alternate endings were either (1) a continuation of the major key from the first phrase or (2) a modulation to a different major key one scale degree higher (the supertonic of the original key) that started at the onset of the second phrase (in the third measure). See Figure 5.7 in Section 5.3.2 for an example filler item.

### Procedure

Participants first completed a short computerized training (8 trials), with feedback, where they explicitly had to answer the question “Does the second phrase contain the same key as the first phrase?" Participants had to successfully complete the training before continuing the experiment, correctly answering the question after all condition types (for stimuli and fillers). To avoid artifactual brain response due to motor preparation and execution (e.g. from a button press), instructions during the EEG experiment were to mentally answer the question after each melody, and to answer with a button press when the question was presented visually after 10% of the trials. To obtain behavioral data for each trial, participants were re-invited to a separate post-hoc session at least two days following the music EEG session, wherein they completed the same experiment again, this time with the instruction to answer the question with a button press after each trial (and with no EEG recorded).

In both EEG and behavioral experimental sessions, blocks of 60 items were presented with stimuli and fillers of only one metricality condition (regular or irregular) separated by the question and a self-paced break; the order of blocks was counterbalanced across participants. Items were pseudo-randomized such that no more than four stimuli or fillers appeared successively nor did any syntactic condition appear more than four times successively.

### 8.2.3 Analysis and results

#### Alternate ending preference

Generally the major mode is more commonplace in Western Tonal Music than the minor mode (Parncutt, 2012), and it has been shown that the Western listener has a general preference for major mode as early as in childhood (Nieminen, Istók, Brattico, & Tervaniemi, 2012). In line with these reports, all participants in the EEG study reported in a post-hoc questionnaire that they preferred major endings over minor endings in music stimuli. Therefore, for all participants the syntactic condition with the major final measure was considered the preferred condition, and likewise the syntactic condition with the minor final measure was considered the non-preferred condition.

#### ERPs

Trials were preprocessed as part of Experiment 2b (see Section 7.2.3 for trial rejection criteria). In one participant, one electrode within the regions of interest (described below) was removed and subsequently interpolated. Participants with more than 20 rejected trials per condition (i.e. less than 40 trials) were excluded from further analysis (final group size was 23).

Epochs used for ERP calculations were extracted from the larger and previously preprocessed epochs (Figure 8.2). The trigger or zero-point in the melody was the onset of the first eighth- note, which consistently indicated the final mode (major or minor) across melodies.

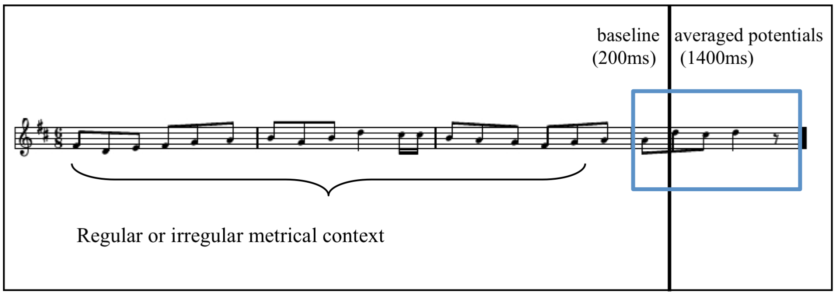



*Figure 8.2: Music ERP epoch. The epoch is marked by the blue box. The vertical black line indicates the trigger.*

A baseline was extended to 200 ms before the trigger. Epochs extended for 1400 ms, which included the end of respective stimulus items and the following inter-stimulus interval (ISI).

For each participant, the epoched time-series data were averaged separately for the preferred (major) and non-preferred (minor) endings as identified in the behavioral post-test (absolute baseline correction applied). The single-subject average per condition was then averaged across the group, creating a grand average for each condition. Grand averages of preferred- and non-preferred syntax conditions (collapsed across metricality type) were used to determine syntax main effects, computed with cluster-based permutations (Oostenveld et al., 2011)). Electrodes found to be significant in the cluster-based permutation tests were used to form midline and lateral regions of interest (ROI): Midline ROIs were anterior (AFz, Fz, FCz, Cz) and posterior (CPz, Pz, POz, Oz). The bi-hemispheric ROIs were sectioned as follows: left anterior (F7, F5, F3, FT7, FC5, FC3), left middle (T7, C5, C3, TP7, CP5, CP3), left posterior (P7, P5, P3, PO7, PO3, O1), right anterior (F4, F6, F8, FC4, FC6, FT8), right middle (C4, C6, T8, CP4, CP6, TP8), right posterior (P4, P6, P8, PO4, PO8, O2). Averaged-over-ROI, averaged-over-time-windows amplitude values were used to conduct (1) a midline ANOVA with factors Syntax (preferred, non-preferred), Metricality (regular, irregular), and AntPost (anterior midline, posterior midline) and (2) a lateral ANOVA with factors Syntax (preferred, non-preferred), Metricality (regular, irregular), Hemisphere (left, right), and AntPost (anterior, middle, posterior). As the amplitude values were not normally distributed for all ROIs, data were log-transformed, and ANOVAs performed on the transformed dataset. All other reported values (mean amplitude, etc.) represent untransformed data.

Cluster-based permutation tests comparing the ERPs evoked by non-preferred and preferred tonal key revealed two significant syntax effects: A negativity between 762 ms and 968 ms, and a positivity between 1050 ms and 1374 ms, both with centro-posterior distribution (see below for statistical analysis). Based on previous literature, these effects will be labeled respectively N500 (e.g., Koelsch et al., 2000) and P600 (e.g., Besson & Faïta, 1995)

*Table 8.1: ANOVA of music ERP amplitudes, midline ROIs. S = Syntax. M = Metricality. A = AntPost. Significant effects in bold.*

| | N500 | | | P600 | | |
|---|---|---|---|---|---|---|
| Effect | $F$(1,22) | partial $\eta^2$ | $p$ | $F$(1,22) | partial $\eta^2$ | $p$ |
| S | 10.10 | .315 | **.004** | 15.20 | .409 | **.001** |
| S x M | 1.36 | .058 | .256 | 4.18 | .160 | **.053** |
| S x A | 0.13 | .006 | .727 | 13.95 | .388 | **.001** |
| S x M x A | 1.42 | .061 | .245 | 2.53 | .103 | .126 |

*Table 8.2: ANOVA of music ERP amplitudes, lateral ROIs. S = Syntax. M = Metricality. H = Hemisphere. A = AntPost. Significant effects ( $p < .05$) in bold. Marginal effects ( $p < .1$) underlined.*

| | N500 | | | P600 | | |
|---|---|---|---|---|---|---|
| Effect | $F$(1,22) | partial $\eta^2$ | $p$ | $F$(1,22) | partial $\eta^2$ | $p$ |
| S | 9.74 | .30 | **.005** | 7.68 | .259 | **.011** |
| S x M | 3.86 | .149 | .062 | 3.28 | .130 | .084 |
| S x H | 0.48 | .021 | .494 | 3.76 | .146 | .066 |
| S x M x H | 0.71 | .031 | .408 | 3.31 | .131 | .082 |
| S x A | 3.73 | .262 | **.041** | 10.07 | .490 | **<.001** |
| S x M x A | 3.60 | .255 | **.045** | .62 | .056 | .549 |
| S x H x A | 0.54 | .049 | .591 | .57 | .051 | .576 |
| S x M x H x A | 0.17 | .016 | .842 | 2.00 | .157 | .167 |

for presentation purposes, despite a later latency; a detailed argument for these labels may be found in Section 8.2.4. Since the hypothesized positivity was approximately 400 ms later than expected, the latency of the positivity was further correlated with measures of musical training to try and systematically account for the difference. However, neither age of musical training onset (Pearson's $R$ =.131, $p$ = .56) nor weekly practice hours (Spearman's $\rho$ = -.042, $p$ = .85) significantly correlated with the latency of the positive effect.

Two ANOVAs were then conducted on the N500 and P600 amplitudes in order to see how metrical regularity of the melody context impacted the syntax ERPs and their respective scalp distributions. Only the results related to the Syntax factor are listed in Tables 8.1 and 8.2 (see Study 2 for an explicit investigation into the Metricality aspect of stimuli and EEG; the same musicians were found to have entrained more to the hierarchical organization (group level) of regular compared to irregular melodies, and equally to the tactus (individual level) in both metricality conditions).

**N500** In the N500 time window (762 ms to 968 ms), a main effect of Syntax was found both in the midline-ROI ANOVA ($F$(1,22) = 10.10, $p$ = .004, partial $\eta^2$ = .315) and lateral-

ROI ANOVA ($F$(1,22) = 9.74, $p$ = .005, partial $\eta^2$ = .300), corroborating the cluster-based permutation findings. The amplitudes of the N500 per ROI are listed in Table 8.3.

**Distribution of the N500** The distribution of the N500 was centro-posterior (Figure 8.3), as indicated by the lateral-ROI ANOVA interaction Syntax x AntPost ($F$(1,22) = 3.73, $p$ = .041, partial $\eta^2$ = .262). This finding was confirmed by separating the analysis into three separate ANOVAs per anterior, middle, and posterior region, where it was shown that the Syntax main effect was significant in the middle and posterior ROIs (middle: $F$(1,22) = 8.05, $p$ = .010, partial $\eta^2$ = .268; posterior: $F$(1,22) = 17.94, p <.001, partial $\eta^2$ = .449), and not in the anterior ROI ($F$(1,22) = 1.99, $p$ = .172, partial $\eta^2$ = .083).

**The impact of Metricality on the N500 amplitude and distribution** The N500 was larger and more broadly distributed in the regular metricality condition compared to the irregular metricality condition. In order to statistically verify how metrical context influenced the size and distribution of the N500 effect, the lateral-ROI ANOVA Syntax x Metricality x Antpost interaction ($F$(1,22) = 3.60, $p$ = .045, partial $\eta^2$ = .255) was split into anterior, middle, and posterior ANOVAs. In those regions where the Syntax x Metricality interaction was significant, ANOVAs were then further separated for regular and irregular metricality conditions. Intuitively, results showed that the Syntax x Metricality interaction only occurred where the Syntax effect was distributed, in middle and posterior ROIs but not in the anterior ROI (middle: $F$(1,22) = 4.62, $p$ = .043, partial $\eta^2$ = .174; posterior: $F$(1,22) = 5.74, $p$ = .025, partial $\eta^2$= .21; anterior: $F$(1,22) = 0.69, $p$ = .415, partial $\eta^2$ = .030). Further analyses indicated that in the regular condition, the Syntax effect remained significant in those middle and posterior regions (middle: ($F$(1,22) = 9.78, $p$ = .005, partial $\eta^2$= .308; posterior: $F$(1,22) = 16.97, p <.001, partial $\eta^2$ = .435), while in the irregular condition, the N500 was moved posterior (middle: $F$(1,22) = 2.10, $p$ = .162, partial $\eta^2$ = .087; posterior: $F$(1,22) = 6.24, $p$ = .020, partial $\eta^2$ = .221) and was significantly reduced in amplitude compared to the regular condition ($F$(1,22) = 5.74, $p$ = .025, partial $\eta^2$= .21). Amplitudes of the N500 per ROI per metricality condition are listed in Table 8.4.

**P600** In the P600 time window (1050 ms to 1372 ms), a main effect of Syntax was found both in the midline-ROI ANOVA ($F$(1,22) = 15.20, $p$ = .001, partial $\eta^2$ = .409) and lateral-ROI ANOVA ($F$(1,22) = 7.68, $p$ = .011, partial $\eta^2$ = .259), supporting the findings of the cluster-based permutation test. The amplitudes of the P600 per ROI are listed in Table 8.3.

**Distribution of the P600** The P600 was central posterior (Figure 8.3), which was demonstrated by a Syntax x AntPost interaction in midline and lateral ANOVAs (midline:

*Table 8.3: Mean N500 and P600 amplitudes ($\mu V$; ± SD) per ROI, elicited by non-preferred syntactic structure (non-preferred minus preferred).*

| ROI | N500 | | P600 | |
|---|---|---|---|---|
| anterior midline | -0.50 | (± 1.19) | 0.28 | (±0.75) |
| posterior midline | -1.21 | (± 1.24) | 0.91 | (±0.77) |
| left anterior | -0.25 | (± 0.96) | -0.14 | (±0.71) |
| right anterior | -0.24 | (± 0.97) | -0.04 | (±0.80) |
| left middle | -0.68 | (± 1.14) | 0.32 | (±0.62) |
| right middle | -0.70 | (± 1.14) | 0.49 | (±0.74) |
| left posterior | -1.14 | (± 1.25) | 0.72 | (±0.81) |
| right posterior | -1.05 | (± 1.18) | 0.10 | (±0.83) |

$F(1,22) = 13.95$, $p = .001$, partial $\eta^2 = .388$; lateral: $F(1,22) = 10.07$, $p < .001$, partial $\eta^2 = .490$). These interactions were resolved by examining the significance of a Syntax effect in anterior and posterior midline ANOVAs and anterior, middle, and posterior lateral ANOVAs. Consistent with the scalp topography, the midline Syntax main effect was significant in the posterior- but not anterior region (posterior: $F(1,22) = 30.72$, $p < .001$, partial $\eta^2 = .583$; anterior: $F(1,22) = 3.24$, $p = .086$, partial $\eta^2 = .128$.), and the lateral Syntax effect was significant in middle and posterior regions but not in the anterior region (middle: $F(1,22) = 9.04$, $p = .006$, partial $\eta^2 = .291$; posterior: $F(1,22) = 28.01$, $p < .001$, partial $\eta^2 = .560$; anterior: $F(1,22) = 0.47$, $p = .49$, partial $\eta^2 = .021$).

**The impact of Metricality on the P600 amplitude** The P600 was reduced in the regular-metricality condition, as hypothesized, shown by the Syntax x Metricality interaction in the midline ANOVA ($F(1,22) = 4.18$, $p = .053$, partial $\eta^2 = .160$). The interaction (shown in Figure 8.4) was resolved by evaluating ANOVAs for a significant Syntax effect separately for regular and irregular conditions: The Syntax main effect was significant in the irregular metricality condition ($F(1,22) = 11.40$, $p = .003$, partial $\eta^2 = .341$) but not in the regular metricality condition ($F(1,22) = 0.765$, $p = .391$, partial $\eta^2 = .034$). The P600 difference waves per metricality condition are shown in Table 8.4.

### Behavior

In a separate behavioral session, at least two days after the music EEG experiment, participants repeated the experiment with the same task (to answer whether the tonal key of the second phrase was the same as the tonal key in the first phrase), with the exception that participants answered with a button press after each item as opposed to after 10% items. This

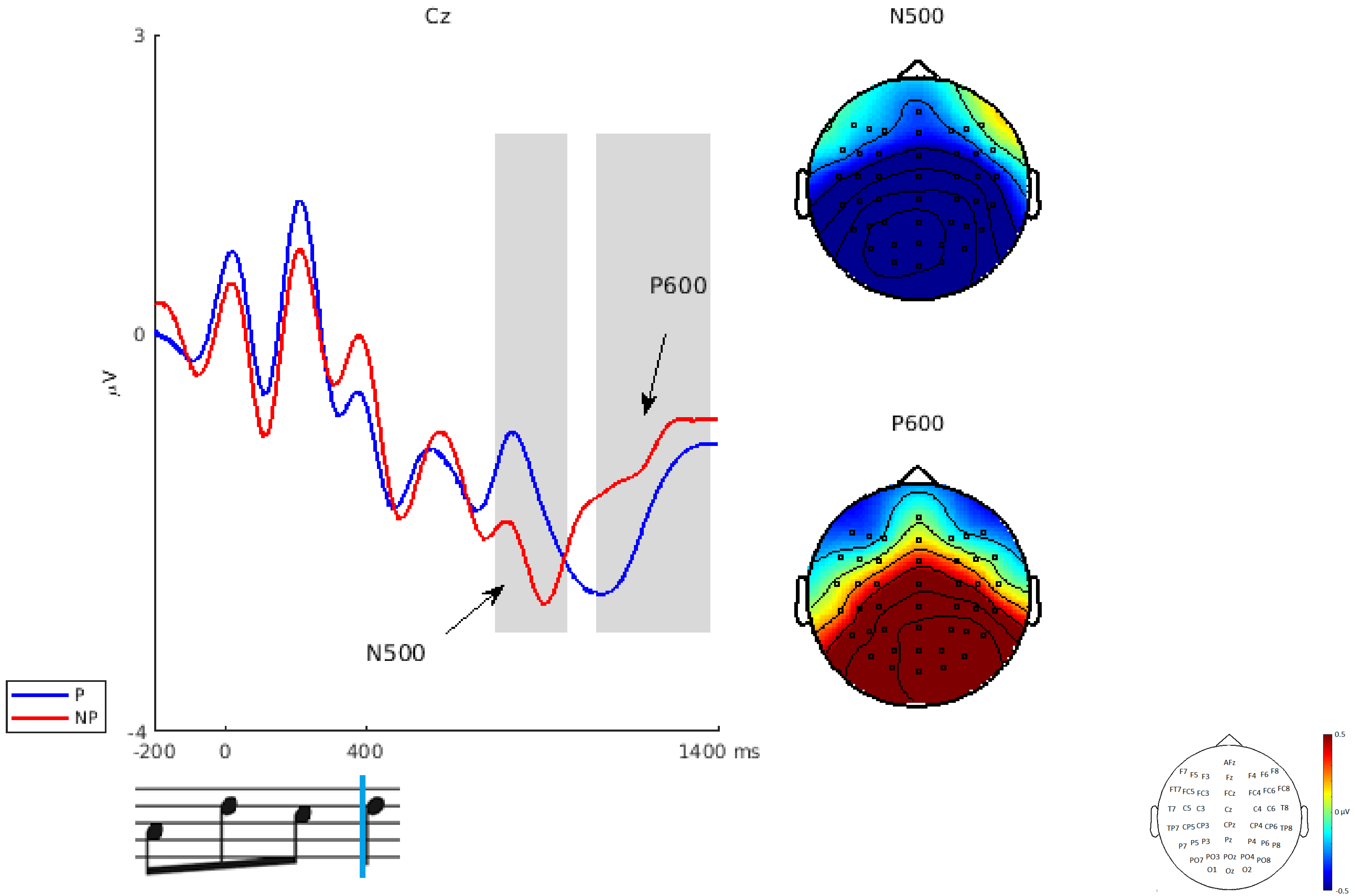



*Figure 8.3: Music ERP potentials and scalp topography. P = Preferred. NP = Non-preferred. Music N500 and P600 effects shaded in gray (left). Both had a centro-posterior distribution (right). While ERPs were time-locked (0 seconds) to the third-from-last note, which consistently indicated the final key, effect latency of the N500 and P600 were consistent with processes that were triggered by the final note (vertical blue line, ca. 400 ms). Scalp topography shows the shaded regions.*

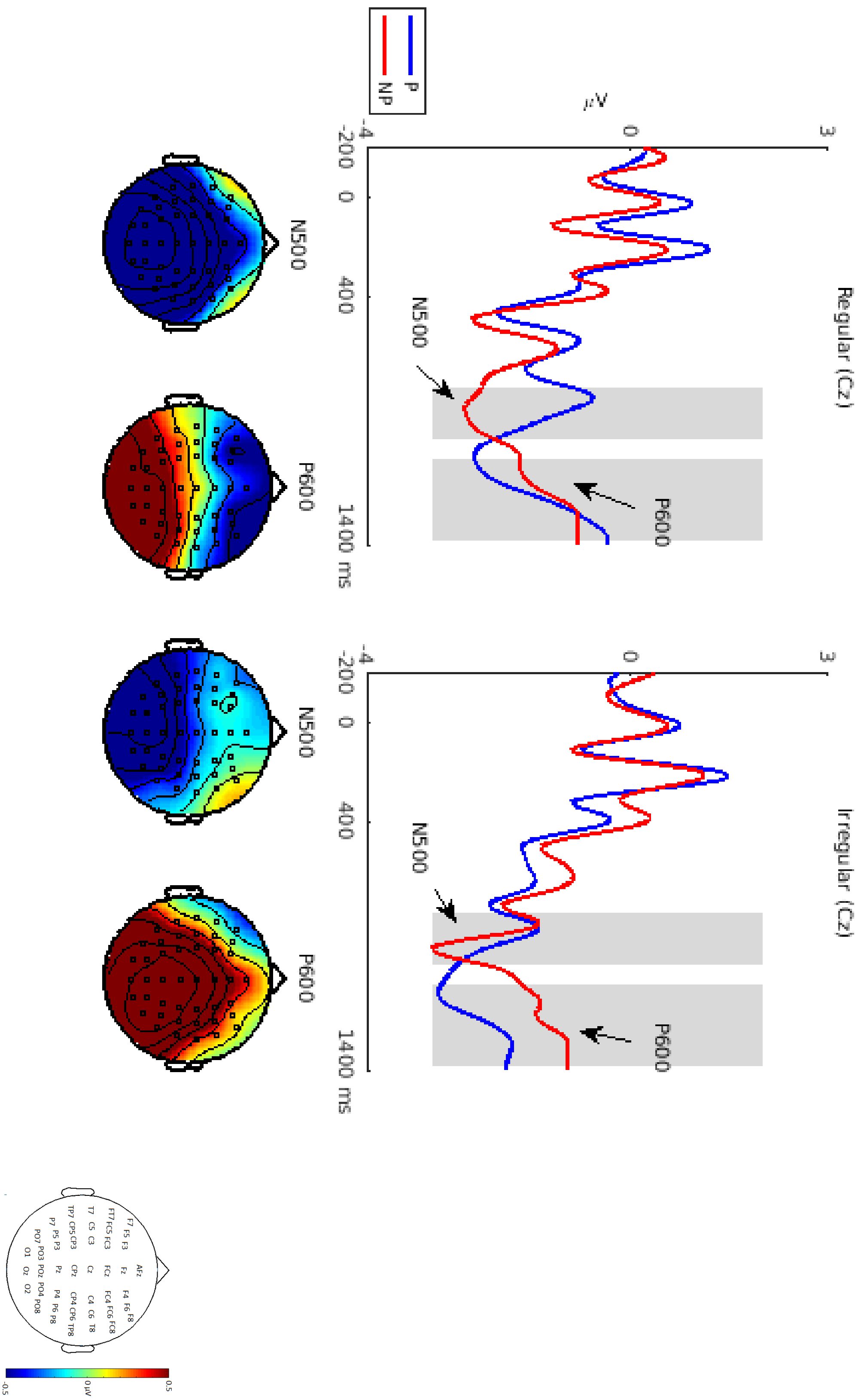



*Figure 8.4: Syntax x Metricality interaction in the N500 and P600. P = Preferred. NP = Non-preferred. Both music ERP effects were influenced by the metrical context. The N500 was larger and more broadly distributed in the regular metricality condition, while the P600 was larger with a larger distribution in the irregular metricality condition. The effect time windows are shaded in gray. Scalp topographies show the shaded regions.*

*Table 8.4: Mean N500 and P600 difference-wave amplitudes (μV; ± SD) per ROI, non-preferred minus preferred syntactic structure per metricality condition.*

| | N500 | | | | P600 | | | |
|---|---|---|---|---|---|---|---|---|
| ROI | Regular | | Irregular | | Regular | | Irregular | |
| anterior midline | -0.81 | (± 1.74) | -0.19 | (± 1.57) | -0.15 | (± 1.07) | 0.71 | (± 1.32) |
| posterior midline | -1.71 | (± 2.00) | -0.70 | (± 1.32) | 0.65 | (± 1.14) | 1.17 | (± 1.36) |
| left anterior | -0.34 | (± 1.23) | -0.16 | (± 1.19) | -0.36 | (± 0.99) | 0.08 | (± 1.08) |
| right anterior | -0.45 | (± 1.27) | -0.03 | (± 1.35) | -0.43 | (± 1.29) | 0.34 | (± 1.05) |
| left middle | -0.97 | (± 1.62) | -0.39 | (± 1.08) | 0.14 | (± 0.90) | 0.49 | (± 0 .96) |
| right middle | -1.12 | (± 1.57) | -0.29 | (± 1.38) | 0.16 | (± 1.17) | 0.82 | (± 1.10) |
| left posterior | -1.54 | (± 1.78) | -0.74 | (± 1.21) | 0.56 | (± 1.02) | 0.87 | (± 1.14) |
| right posterior | -1.52 | (± 1.69) | -0.57 | (± 1.32) | 0.78 | (± 1.09) | 1.21 | (± 1.23) |

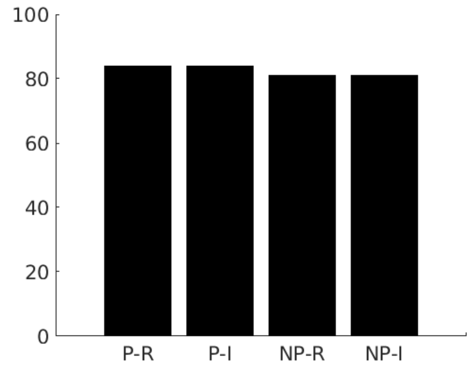



*Figure 8.5: Accuracy per music condition. P = Preferred. NP = Non-preferred. R = Regular. I = Irregular.*

was scheduled in order to collect behavioral data per trial while avoiding motor-related ERP responses during EEG recording. Accuracy scores per condition verified that participants were able to perform the syntax task during EEG (Figure 8.5). Log-transformed scores were entered into a 2 x 2 ANOVA with factors Syntax (preferred, non-preferred) and Metricality (regular, irregular), yielding no main effects or interactions (Syntax: $F(1,22) = 0.89$, p = .36, partial eta squared = .04; Metricality: $F(1,22) = 0.00$, $p = .96$, partial eta squared = .00; Syntax x Metricality: $F(1,22) = 0.01$, $p = .94$, partial eta squared = .00 ).

The post-hoc behavioral sessions demonstrated that participants correctly executed the instructions, but accuracy scores did not replicate any effects from the ERP. The absence of any effects may be attributed to participants having been exposed to all stimuli items already in a previous session; therefore the relative difficulty was reduced (or less expected items

*Table 8.5: Music P600 difference waves (non-preferred – preferred syntax ) per metricality condition per ROI, correlated with cognitive factor scores. VWM = Verbal Working Memory. TPD = Time and Pitch Discrimination. Pearson's R, two-tailed correlations, p < .05*, p < .01**. Correlations were omitted in ROIs where no P600 effect was significant.*

| | Regular | | Irregular | | Regular -– Irregular | |
|---|---|---|---|---|---|---|
| ROI | VWM | TPD | VWM | TPD | VWM | TPD |
| anterior midline | | | | | | |
| posterior midline | -.15 | .18 | **.48**** | **.46*** | **.42*** | .21 |
| left anterior | | | | | | |
| right anterior | | | | | | |
| left middle | -.23 | .14 | **.54**** | **.42*** | **.53**** | .20 |
| right middle | -.15 | .18 | **.43*** | **.43*** | .37 | .15 |
| left posterior | .02 | .37 | **.48*** | **.53*** | .37 | .15 |
| right posterior | -.02 | .32 | .39 | **.51*** | .31 | .17 |

were expected, since they were heard before). On the other hand, the behavioral measure may not have captured the online processes revealed by ERP findings, which are discussed below in Section 8.2.4. The behavioral results will not be further discussed.

### Individual differences

As the musicians in this study were a subgroup of the large participant group in Study 1, each participant had factor scores Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD). In order to test whether the cognitive processes associated with these two factors individually modulated the amplitude of the P600, factor scores were correlated with the P600 amplitude, separately for regular and irregular metricality (Table 8.5).

Both cognitive factors played a role in syntactic integration, as hypothesized. Factor correlations showed that in the irregular metricality condition, P600 amplitudes were larger for people with higher VWM and TPD scores: the irregular P600 amplitude correlated positively with TPD in all ROIs where the effect was significant, and similarly with VWM in all (except the right posterior) ROIs. The P600 elicited in the regular-metricality condition did not correlate with either score. Furthermore, VWM correlated positively with the magnitude of difference between the regular P600 effect and the irregular P600 effect, albeit only in two ROIs. That is, the higher the VWM score, the more drastic the amplitude reduction in P600 effect from irregular to regular metricality condition, indicating that these higher VWM participants benefited more from the regular metricality than did their lower VWM counterparts.

### 8.2.4 Discussion: Music ERPs

#### Results summary

Experiment 3a aimed to assess whether syntactic resources that underlie the processing of non-preferred (but not incorrect) harmonic structure were facilitated by a regular as opposed to irregular metrical context, and whether individual differences among musicians affected the ERPs. Musicians were therefore presented with melodies that had preferred or non-preferred syntactic endings, presented in regular or irregular metrical contexts. A P600 was hypothesized to reflect reanalysis of non-preferred melodic endings, be reduced in the regular compared to irregular melodies, and its amplitude was expected to correlate with cognitive factors Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD).

The current hypotheses were all supported. The non-preferred (but not incorrect) minor melody endings elicited the hypothesized P600 (Figure 8.3), thus extending the function of the P600 in music from overt violations (e.g., Patel et al., 1998), incongruity (e.g., Featherstone et al., 2013), and within-key melodic deviations (e.g., Besson & Faïta, 1995) to structural preference, as also reported in language (e.g., Osterhout & Holcomb, 1992). The P600 was reduced in amplitude in the regular metricality condition (Figure 8.4), as hypothesized, indicating that syntactic integration of the non-preferred structure was easier in the regular condition. This result extends previous findings in language, which indicated that syntactic comprehension was influenced by the metrical composition of sentences (Roncaglia-Denissen et al., 2013; Schmidt-Kassow & Kotz, 2009a, 2009b). Individual differences played a role in the irregular but not regular metricality condition, discussed below. Participants with higher TPD and VWM scores had a larger P600 effect in the irregular metrical context, and participants with higher VWM had greater reduction in P600 amplitude when comparing regular to irregular metrical context.

A non-hypothesized negativity preceding the positivity was also elicited, which was most likely an N500 (Figure 8.3). The N500 was found first in the context of harmonic violations in music, deemed to represent structural integration of an unexpected musical chord (Koelsch et al., 2000). Later the N500 was found to interact with the N400, a component elicited by language semantic violations, and therefore deemed also to be linked to 'meaning' in music (Steinbeis & Koelsch, 2008). Both interpretations will be discussed below. The current N500 was also impacted by meter, being reduced in amplitude and distribution during melodies with irregular metricality.

### Integrating syntactically non-preferred structure in music

**P600** The positivity elicited by non-preferred melodic endings was interpreted as a P600. The posterior distribution was consistent with previous reports of the P600 (Featherstone et al., 2013; Patel et al., 1998), and the later latency can be attributed to the syntactic integration processes being first engaged at the last note, and not the antepenultimate note as hypothesized; the near 1000 ms latency places the effect precisely 600 ms after the final note, which occurred approximately 400 ms after the original trigger. This interpretation is justified when considering that the final note in the melody carries much harmonic weight; the awareness of the final tonal key among these amateur musicians may easily have occurred only at the final note, although effect latency did not correlate with measures of musicianship.

The current study is the first to show a P600 in response to this type of non-preferred harmonic structure — major as opposed to minor mode endings — that avoided overt errors. Structural preference therefore joins a body of literature that reports late positive effects for harmonic violations (Patel et al., 1998), harmonic incongruities (Featherstone et al., 2013), melodic unexpectedness Pearce et al. (2010), and in- and out-of-key harmonic deviations (Besson & Faïta, 1995). This is in line with the preferences and well-formedness concepts stipulated by the GTTM (Lerdahl & Jackendoff, 1983, see Section 3.1.1), and shows that the resources responsible for integrating musical syntactic constituents operate with a set of preexisting preferences that both guide expectancies and supersede grammatical correctness.

**N500** The negativity elicited by the non-preferred or minor mode melody endings was not hypothesized, but demonstrated analogous characteristics to a negativity labeled 'N500' found in previous music EEG studies (e.g., Koelsch et al., 2000; Steinbeis & Koelsch, 2008). Using the same argument as above, that processing of the melodies occurred at the final note, the current latency is within the typical N500 range. While the central posterior distribution of the current negativity does not match initial reports of the N500's anterior distribution, a subsequent report of an N500 with melodies (as opposed to isochronous chord sequences) included broad distributions (Featherstone et al., 2013), validating the current interpretation as N500.

The exact nature of the N500 varies among reports in the literature. Three possible interpretations of the N500 will be discussed here, as pertaining to: (syntactic) musical expectancy, (semantic) harmonic resolution, and holistic perception of harmonic resolution (without subscription to syntax or semantics). First, the initial discovery of the N500 was related to musical expectancy, when chords that did not fit their preceding harmonic context elicited a larger negativity than chords that were fully expected. The N500 was in this case deemed to represent a "higher degree of integration needed to process unexpected chords," (Koelsch et al., 2000) and attributed to syntactic processing. Second, the N500 was elicited in a study that investigated potential overlapping semantic resources in language and music. Dissonant chords were found to elicit an N500, which interacted with an N400 elicited by language semantic violations; thus the N500 was thought to provide a "route to meaning in music" via tension-resolution of harmonic dissonance (Steinbeis & Koelsch, 2008) and attributed to semantic processing. Third, Featherstone et al. (2013) recently pointed out that most ERP studies reporting an N500 were using non-musician participants and musical stimuli which ended in a permanent (unexpected) key change, leaving participants without a sense of harmonic resolution. The authors themselves also found an N500 in harmonically unresolved harmonic sequences, in non-musicians, attributed to a 'holistic' type of listening employed by non-musicians (their musicians had no N500, but rather a P600).

Applied to the current study, the first syntactic account would imply that the N500 was related to the assessment that the minor end did not fit its preceding harmonic context, and, in this sense, reflected syntactic integration of an unexpected harmonic event. As to the second interpretation, the current N500 may have reflected some overall response to the qualitative aesthetic difference between major and minor melodic endings, as the minor key is often associated with sadness or nostalgia compared to happy or upbeat associations to the major mode (Nieminen et al., 2012). Regarding the third possibility, the current N500 elicited by non-preferred melodic endings could be in line with (simply) a perceived lack of harmonic resolution, but the fact that it was elicited in musicians speaks against the N500 resulting from a general holistic listening strategy.

In the current study, it is not possible to distinguish between the syntactic and semantic accounts for the N500—future ERP studies with the current melodies could systematically explore syntactic versus semantic processes, for example, by altering the instructions as an experimental variable (to focus on meaning vs. structure) to see whether the N500 is modulated, and by giving detailed semantic vs. syntactic ratings to the stimuli. The syntactic and semantic processing of these stimuli may also have been collapsed, in that the syntactically non-preferred structure provided musical meaning to the participants (Koelsch, 2011a). To further explore this latter possibility, future studies would have to investigate non-preferred structure and meaning separately from major vs. minor modes, to eliminate the cultural associations exogenous to the music (like happy and sad). Perspective N500 paradigms are suggested in Section 10.2.

#### The impact of metrical context on syntactic integration (and related music cognition)

This ERP study showed that hierarchical metrical organization impacts the cognitive processing of music. The Metricality x Syntax interaction is in line with the current neurocognitive entrainment approach, which suggests that syntactic integration will be facilitated in a regular compared to irregular metrical context due to both the theoretical matching of metrical and syntactic constituent boundaries and to the neural consequences of entrainment (see Chapters 2 & 3).

Accordingly, when accepting the premise that entrainment impacted cognition, this experiment showed that syntax-related processes are more contingent upon higher levels of metrical organization than the tactus. If lower-level temporal regularity were the most decisive influence on cognition, then there should not have been a Metricality x Syntax interaction, since entrainment to the individual level (Experiment 2b) was equal in both metricality conditions. This result therefore demonstrates the importance of acknowledging hierarchical metrical organization in studies of music cognition.

**P600** The reduced P600 in the regular meter condition supported the hypothesis, which was that a reduced processing cost owing to the regular metrical context would result in a reduced P600 amplitude. The fact that these musician participants entrained more to the higher metrical levels in regular as opposed to irregular melodies, in Experiment 2b, further supports this interpretation. Moreover, the evidence aligns with other musicians' results with these stimuli in Experiment 1b, where reaction time and accuracy of note detection were enhanced in current regular compared to irregular melodies, consistent with entrained cognition in the regular meter context.

The increased entrainment in the regular melodies (Experiment 2b) could also be independent of the reduced P600 in regular melodies, i.e. that there is no impact of entrainment but rather some other property is causing the interaction. For example, the irregular metrical structure could have been simply distracting to the musicians, and therefore increased the processing cost of syntactic integration in the irregular metrical condition. However, the convergence of results from Experiment 1b, Experiment 2b and Experiment 3a, along with the presented theoretical framework, speaks strongly for the ERP interaction to be linked to greater entrainment in regular melodies.

**N500** Regarding the reduced N500 in the irregular condition, it may be that the syntactic or semantic processes reflected by the N500 were effectively blocked by the cognitive processing necessary for the different-than-usual metrical context. Indeed different neural mechanisms have been proposed to underlie different attending modes of the DAT (Henry & Hermann, 2014), attending modes that are unique to regular ("future-oriented attending") as opposed to irregular ("analytic attending") metrical contexts (see Jones & Boltz, 1989). Regular meter is the default in Western Tonal music, thus in the current participants, future-oriented attending likely most often accompanies the music cognition underlying the

N500. Perhaps traditional elicitation of the N500 is entrenched in the neural mechanisms that underpin future-oriented attending, and the cognitive system wrote off the N500 in the recalibrated attending mode.

The reduced N500 in the irregular condition could also be evidence in support of the N500 reflecting exogenous, cultural 'meaning' associated with major and minor modes, that this meaning is also tied to some kind of automatic association with Western Tonal metrical structure (e.g., music from other traditions may be cued by their different rhythmic structure, in which case assumptions of what modes 'mean' might fall away). Again, full interpretation of the N500 interaction must await future studies (see Section 10.2)—for example, would the N500s reported in previous studies disappear if the meter were re-written to be irregular? An informative starting point would be to replicate a paradigm known to produce an N500 (e.g., Koelsch et al., 2000), and then to introduce metrical irregularity to see how the N500 is modulated.

#### Individual differences in syntactic integration

This is the first ERP study to show that individual differences in the cognitive ability among musicians were linked to syntactic integration of harmonic constituents, and moreover that individual differences were impacted by metrical context. Correlations with factor scores were significant for the P600 effect elicited by irregular melodies, but not by regular melodies.

The absence of correlations in the regular metricality condition may be due to the regular metrical context 'equalizing' cognitive differences with respect to syntax processing; the temporally predictable, regular metrical context served as a processing scaffold or tool, which lessened differences that are more evident in the irregular metrical context. A concrete analogy is useful for this explanation: A group of architects is asked to draw a perfect circle, free-hand. In this group, some will have better drawing skills than others—there will be natural variability in the quality of free-hand circles, which may perhaps correlate positively with some measure of fine motor skills (the better the measurement, the better the circle). Next, all architects are given a protractor and asked to draw another perfect circle. Now the variability among the group is reduced, the perfection of the circle no longer related to their artistic ability or hand-eye coordination—all of them are able to draw a perfect circle with their trade-tool. Relating back to the current study, the individual ability would be in perceiving acoustic information and subsequent neural synchrony to that information, and the regular meter serves to provide everyone with equal opportunity to synchronize to the highly temporally predictable information. The cognitive factors TPD and VWM measured the individual variability, and showed that when the regular meter 'trade tool' was missing, successful syntactic integration relied on the TPD and VWM ability.

The direction of the correlation between irregular P600-effect amplitude and cognitive factors was positive, and moreover, participants with higher VWM benefited more from the regular metricality (in two of six ROIs where the P600 was found), represented by the correlation with VWM and the magnitude of the P600 reduction in the irregular metrical context. These correlations showed that the musicians with higher cognitive scores had the higher P600 amplitude, and that this amplitude was reduced when processing cost was reduced. This direction is counterintuitive at first glance: If a higher amplitude reflects higher processing costs (as previously stated), why do not higher-scoring participants have a lower amplitude to indicate that they integrate the syntactic anomaly more efficiently?

The higher processing cost among participants with higher cognitive scores is likely due to their better detection of the major-minor difference and subsequent greater salience of the syntactic structure (i.e., “there was more to integrate”). Musicians with superior working memory and auditory discrimination abilities may be more explicitly aware of the syntactic structural differences between melodic phrases; previous studies have indicated that music syntax ERP effects are limited to participants who can perceive nuanced syntactic structure in the first place (Besson & Faïta, 1995; Fitzroy & Sanders, 2013; Koelsch et al., 2002).

These musician ERP responses to melodies contribute much to the literature when standing alone, yet address only half of the current investigation. The following experiments round out the story that is sought after by Study 3—to see whether these ERP phenomena manifest comparably in music and language. A behavioral experiment (3b) first addressed subject vs. possessor relative clause attachment preference among native German speakers (in a new participant group), in order to most accurately code preferred and non-preferred conditions in the language ERP experiment (3c). Experiment 3c then addressed syntactic reanalysis of non-preferred linguistic structure in the musicians from Experiment 3a, whether their reanalysis was facilitated by entrainment to a regular metrical context, and what role individual differences played in the process.

## 8.3 Experiment 3b: Attachment preference in auditory sentences

### 8.3.1 Introduction

The purpose of Experiment 3b was to determine attachment preference in German native speakers for the stimuli used in the language ERP experiment. This experiment was motivated by mixed results in the literature when it comes to preference for German relative clause attachment: Reading studies have shown a clear preference for high-attachment (e.g., Hemforth et al., 2000), whereas studies investigating heard sentences have shown varied preference, depending on the type of attachment, and in particular a preference for low-attachment in the case of Genitive possessives, the type of attachment in the current sentence stimuli (Augurzky, 2006).

A full theoretical account of relative clause attachment preference in the current German sentences was outside the current scope of the study[2], whose purpose was instead to see if reanalysis of non-preferred attachment can be facilitated by a regular metrical context, regardless of what that preference is. Thus to accurately code attachment preference in the subsequent language ERP experiment, a sentence-completion experiment was devised in an attempt to categorize and understand preferences for the attachment of relative clauses in the current stimuli. The design allowed the participants to spontaneously attach the relative clause of each sentence either 'high' to the subject or 'low' to the subject's genitive possessor, according to their own initiation, while taking a neutral approach to theoretical causes of attachment preference.

Just as attachment preference evidence differs across the literature, so do accounts of individual differences. For example, previous evidence from English showed high working memory capacity to be linked to attachment preference in opposite directions (e.g., Traxler 2007 to preferred high attachment; Swets, Desmet, Hambrick, and Ferreira 2007 to preferred low attachment) or not at all (Hocking, 2003). These studies were in the visual domain, assessing read sentences and a single reading span measure; a current consensus in the literature is that future investigations need to be more specific about which individual resources are involved in ambiguity resolution and what exactly they reflect (e.g., Caplan & Waters, 2013; Van Dyke, Johns, & Kukona, 2014). The cognitive factors Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD) determined by Study 1 therefore offered a unique opportunity to explore how (or if) richly nuanced individual measures from the auditory domain affect attachment preference in the auditory domain.

Thus the current experiment was designed to aid attachment preference coding for sentence stimuli in a subsequent ERP experiment, and carried out in a group of monolingual native German speakers (separate from the musicians tested in ERP Experiments 3a and 3c). Individual differences in cognitive factors Verbal Working Memory and Time and Pitch Discrimination were assessed in an exploratory fashion to see if cognitive functions associated with these factors play a role in relative clause attachment preference in the current sentences.

[2]Relative clause attachment preference is different cross-linguistically, and has many influential factors ranging from prosodic cues (which were eliminated by digital cross-splicing of current materials) to type of determiner between competing nouns that the relative clause can modify (Augurzky 2006 outlines these factors in German). A recent state-of-the-art in ambiguity resolution literature may be found in Sanz et al. (2013).

### 8.3.2 Methods

#### Participants

Thirty monolingual native German speakers (14 male), aged between 21 and 34 years (*M* = 25.2 years, *SD* = 3.12), were recruited from the participant database of the Max Planck Institute for Human Cognitive and Brain Sciences and were paid 7 € per hour. Participants were not part of the EEG study and had a wide range of musicianship (e.g., from self-reported non-musicians to students of the local music conservatory; see Appendix C).

#### Materials

Selected sentence stimuli (used in Study 2, described in Section 5.3.2 and below in Section 8.4.2) were edited with Praat 5.2 to stop before the final verb (Example 8.3.1). Sentences included in Experiment 3b were selected such that in the complete original versions, half were attached high (subject relative clause) and half were attached low (possessor relative clause), thus participants were presented with each lexical combination only once but the items were balanced to equally represent pre-critical-words in both syntactic conditions.

**Example 8.3.1.** Da hinten schuften die Diener der Ärztin, die Leeds vor kurzem-
*Over there are toiling the servants of the doctor, who Leeds recently-*

Stimuli for sentence completion. Participants were instructed to complete sentences with their preferred conjugation of the word 'besuchen'/ *to visit*.

#### Procedure

Participants were seated in a quiet room and listened to sentences via headphones at a comfortable volume. Participants were instructed to speak their desired conjugation of the word 'besuchen'/*to visit* into a microphone, and completed a short training to test their understanding of the instructions to assure that their enunciation was clear. Testing lasted approximately 15 minutes. The experimenter logged answers and interviewed the participants post-hoc for an explanation of their relative clause attachment preference. Answers and interviews were recorded by a microphone, and used post-hoc to verify the answer log. After the experiment, participants completed the seven diagnostic tests (Wechsler's forward and backward digit span, modified listening span, non-word repetition, anisochrony detection and Musical Ear Test rhythm and melody subtests; see Experiment 1a).

### 8.3.3 Analysis and results

#### Attachment preference

Since the goal of this experiment was to accurately categorize relative clause attachment preference in an ERP experiment, sentence-completion results were initially assessed with respect to the actual individual trials completed by each participant as well as the overarching self-reported preference of each participant (in an offline, post-hoc judgment). The two measures may well reflect different cognitive processes (Traxler, 2007): self-reported measures are an offline judgment, while the spontaneous completion of sentences in each trial is closer to an online measure (though not as sensitive as EEG or eye-tracking, for example), thus agreement among the measures increases validity of any discovered phenomena. These two approaches also allow later coding either based on majority of individual trial answers (e.g., if an overwhelming majority of sentences were completed with the same attachment preference over the group, this blanket preference would be assigned in the subsequent ERP experiment coding) or on individual self-reported preference (e.g., if across the group, participants reported a post-hoc preference that was consistent with the way they answered the majority of their trials, a post-hoc questionnaire could reliably be used to individually code attachment preference in the ERP experiment).

Attachment preference per trial was determined from the verb conjugation spoken by participants. Participants answered spontaneously in both simple past tense (providing the critical morpheme as the final syllable in the verb 'besuchen'/ *to visit*) or present or past perfect tense (providing the past participle 'besucht' along with a conjugated form of the auxiliary verb 'haben'/ *to have*). Plural conjugation indicated a subject relative clause (high attachment) while singular conjugation indicated a possessor relative clause (low attachment).

Across participants, the post-hoc questionnaire showed that preference was split for high- and low-attachment of the relative clause, the majority favoring low-attachment (64% low-attachment preference, or 16 of 25 participants who reported a stable attachment preference). Participants' post-hoc reported preference matched their online experimental sentence-completion answers: in a regression analysis using the enter method, reported preference predicted the percentage of high-attachment answers ($F(1,28) = 15.23$, $p = .001$, $R^2 = .35$, $R^2{}_{adjusted} = .33$), thus validating an approach for the different participant group of the EEG experiment to code a *preferred* condition based on individual self-reported preference. Both the experimental sentence-completion answers and the post-hoc self-reported attachment preference were assessed with respect to the individual cognitive factor scores, described below in Section 8.3.3.

### Influence of lexical semantics

With the forethought that native German speakers may individually vary in preference, participants were interviewed after the experiment in an attempt to capture potential causes of differences in preference. Participants were asked for their city of birth origin, that of their parents, whether they were raised speaking formal German or a colloquial dialect, and they were asked to explain in detail what motivated their sentence completion answers. No pattern emerged with respect to participant (or their parents') city of origin, or usage of formal German as opposed to dialect, and neither is further reported here. The detailed explanations as to sentence completion, however, yielded a very interesting an unexpected result: many participants began to talk about the semantic meaning of the words in the sentence, asserting that this influenced their relative clause attachment more than their general syntactic preference for high vs. low attachment.

A categorical difference emerged among participants: approximately a third had strong syntactic preferences to which they consistently adhered in the sentence completion; approximately a third had a general syntactic preference which they consistently overrode, citing that 'semantic content' of lexical items in sentences made them sometimes answer individual trials against their general syntactic preference; and approximately a third had a weak syntactic preference, and who based their sentence-completion almost exclusively on the semantic content of lexical items. For example, in the item "servants of the doctor, who went to Leeds" (Example 8.3.1), these latter participants said servants were less likely to visit Leeds than a doctor, thus conjugated the final verb to assign attachment to the doctor, even if they globally thought attachment to the first-occurring noun sounded more natural.

Out of 25 participants who reported a stable attachment preference, nine reported a large influence of semantics, eight reported a partial influence of semantics, and eight reported that semantics played no role on their sentence-completion answers. These participants were respectively categorized into three groups (1: large semantic influence; 2: partial semantic influence; 3: no semantic influence). These post-hoc groups were used to assess consistency of answers with the self-reported degree of semantic influence (similar to the above assessment of consistency between self-reported attachment preference and actual attachment answers), and for individual difference testing with cognitive factors, described below in Section 8.3.3.

To assess the impact of semantics in the individual sentence-completion trials, actual answers were recoded to represent a 'semantic' score. Whereas in Section 8.3.3, where participants were given a score to represent % trials answered with high-attachment, the semantic score was calculated (per participant) by finding the % trials answered according to the post-hoc reported relative clause attachment preference. (For example, if two participants each answered 70% high attachment, but reported opposite attachment preference, then the participant with self-reported high-attachment preference would be given a semantic score of 70%, and the participant with self-reported low-attachment preference would be given a semantic score of 30%). An analysis of these scores showed that participants reliably classified their own degree of 'susceptibility' to semantics, hence the categorical grouping of the self-reported role of semantics on relative clause attachment (large, partial, or no role) was validated to use in further statistical testing: in a regression analysis, the self-reported degree of semantic influence accurately predicted this 'semantic score' ($F(1,23)$ = 72.38, $p < .001$, $R^2 = .76$, $R^2{}_{adjusted} = .75$).

In an effort to see whether this discovered role of semantics was linked to attachment preference, a $\chi^2$ test was performed with the two self-report measures (role of semantics and syntactic attachment preference). The result indicated that the two measures were independent ($\chi^2 = .043$, $p = .979$), thus it is unlikely that the two form a causal relationship in either direction, i.e., attachment preference was not dictated by nor did it influence the likelihood of participants to be influenced by semantic content in the sentences.

### Individual differences

The current paradigm offered an opportunity to assess individual differences in cognitive resources related to relative clause attachment in the auditory domain, in the face of various inconsistent evidence from the visual domain (Hocking, 2003; Swets et al., 2007; Traxler, 2007). The cognitive factors were used to assess individual differences both in relative clause attachment preference and in the newly discovered influence of semantics. Both the self-reported categories (high vs. low preference; large, partial, or no influence of semantics) and the individual trial answers (% high attachment, % syntactic preference) were used in correlations with VWM and TPD scores, which were assigned to participants in Study 1 based on a factor analysis of seven diagnostic tests. One further participant was removed from the sentence completion experiment on grounds of being an extreme outlier in the diagnostic tests, thus the following correlations were carried out with 24 participants.

Regarding syntactic preference for high vs. low relative clause attachment, the self-reported attachment preference and the individual trial answers were each correlated separately with the factor scores, which showed only marginal effects, shown in Table 8.6. Neither correlation held when followed up with a larger participant group[3], thus the correlations were likely spurious.

[3]The two marginal correlations were followed up with a larger participant group, considering that all 84 participants from Study 1 (including the current group) had given a self-report of their preferred relative clause attachment. With 74 participants who had either high or low attachment preference (10 excluded who had none or a waffling preference), a regression analysis was conducted to see if TPD and VWM could predict attachment preference. The regression showed that neither TPD nor VWM predicted attachment preference ($F(2,72) = 0.28$, $p = .76$, $R^2 = .09$, $R^2{}_{adjusted} = -.02$). Thus the two marginal findings in the current experiment

*Table 8.6: Correlations of syntax and semantics scores and cognitive factor scores. All trial-by-trial correlations are Spearman's ρ (due to non-normal distribution), and all post-hoc report correlations are Pearson's R. Significant correlations in bold.*

| | Correlation statistic (*p*) | |
|---|---|---|
| | VWM | TPD |
| Syntax | | |
| Trial-by-trial | -.352 (.092) | .064 (.765) |
| Post-hoc report | -.13 (.537) | .391 (.059) |
| Semantics | | |
| Trial-by-trial | .336 (.108) | .122 (.569) |
| Post-hoc report | .214 (.314) | **.628 (.001)** |

Regarding the role of semantic influence, the self-reported influence of semantics (large, partial, or none) and the individual trial answers (% answers deviating from listed syntactic preference) were again each correlated separately with the factor scores (Table 8.6). Results indicated that VWM had no relationship to the role of semantic influence whereas TPD did, in that participants with higher scores reported less influence of semantics on their sentence completion. The fact that the two measures (self-report and individual trial answers) behaved differently in correlations could be attributed to distinct cognitive processes involved in the offline self-report or the online sentence completion (Traxler, 2007), and the offline measure is likely the more accurate one, considering that all sentences did not influence a semantic interpretation to the same degree.

### 8.3.4 Discussion: Behavioral sentence completion

The goal of Experiment 3b was to find a reliable way to code 'preferred' sentence conditions for a subsequent ERP experiment, and the goal was achieved. Considering that participants varied in preference, individual coding would be necessary for ERP participants. Participants' post hoc report of their own relative clause attachment preference accurately reflected the way that they answered individual sentence completion trials, thus a valid approach for Experiment 3c was to code the EEG preferred condition based on the post-hoc self-reported preference of participants.

The majority of participants (slightly less than 2/3) preferred to complete sentences such that low attachment occurred, assigning the relative clause to the second of two nouns (e.g. in "the servants of the doctor," the relative clause modified "doctor"). This result supports previous findings for low-attachment preference in German speakers in the auditory domain

were likely spurious.

(Augurzky, 2006). Since participants who dissented (e.g., they preferred high attachment, or to modify the first noun "servants") were in part very strong in their preference, ERP would nevertheless be most accurate in reflecting reanalysis of non-preferred structure if preference were determined on a per-participant basis.

An unexpected finding was that independent of syntactic preference, participants were variably influenced by the semantic content of the sentences. For example, participants who syntactically preferred high attachment (to the first noun or subject), but were influenced by semantics, reported completing the sentence "the servants of the doctor, who went to Leeds," such that the doctor (the second noun or possessor) went to Leeds because it was less plausible that servants could afford such a trip. This result emerged from participant interviews, such that approximately a third of participants reported being largely influenced by semantics, a third partially influenced by semantics, and a third not influenced by semantics at all. Among the participants who were influenced by semantic content in the first place, the frequency of lexical items probably generally biased toward low attachment of the relative clause (in addition to some sentences which were more plausible due to the financial circumstances of who could take a trip): The constrained nature of the sentence meter resulted in a tendency for more general, higher frequency words in the subject position (friends, customers) and lower-frequency, perhaps topically more exotic words in the possessor position (Latvian woman, heiress). This means that for participants who have no preference for either attachment (who were excluded in the current analysis), the current stimuli set would likely bias their preferred attachment on the whole to the possessor, or low attachment, and that such participants in the EEG should be assigned a low-attachment "preferred" condition. That the role of semantic influence was linked to Time and Pitch Discrimination scores is discussed below.

When cognitive factor scores Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD) were correlated with results of syntactic attachment preference, there appeared to be no systematic way to account for participants' relative clause attachment preference, although marginal tendencies were that participants with higher TPD scores tended to prefer high attachment and participants with higher VWM completed sentences more often with low relative clause attachment. Considering that a follow-up regression analysis with a larger participant group found neither of these correlations to be significant (see Footnote 3), the evidence most strongly suggests that neither set of resources associated with the cognitive factors is involved with whatever processes determine syntactic preference with respect to relative clause attachment in the auditory domain.

While neither cognitive factor was convincingly linked to syntactic attachment preference itself, TPD was strongly linked to the degree to which semantic content of sentences motivated participants to attach relative clauses contrary to their syntactic preference. Participants with higher TPD scores were less influenced by semantic content, and answered more consistently with their own syntactic preference (be it high- or low-attachment). In a practical sense, this result can additionally aid the EEG experimental coding. If participants who have no attachment preference also have lower TPD scores (compared to others in the group), an assigned low-attachment preference would be strengthened by the likelihood that they depend on semantic content to determine a preferred attachment, which as stated above is biased to low in the current sentences.

This intriguing finding, that lower TPD accompanied a predisposition to set aside syntactic preference in favor of semantic content, may be interpreted in a way consistent with arguments presented in the dissertation thus far: participants with a higher TPD can 'afford' to stick to their preference, whereas adhering to syntactic preference is too costly for participants with lower TPD, who in turn look more to context cues to complete the auditory task. Higher TPD means higher auditory perceptual advantage; in Study 1, when the factor scores were established (Experiment 1a), higher TPD scores were associated with musicianship (and musicians who practiced more). In the subsequent behavioral experiment (Experiment 1b), higher TPD meant better discrimination in melodies with regular metricality and sentences with irregular metricality, which was suspected to be a reflection of musicians' entrainment to these conditions (but not successfully proven in Study 2). Moreover, in ERP Experiment 3a, higher TPD scores were linked with better syntactic integration.

Returning to the current experiment, participants who attended better to acoustic information (as indicated by their higher TPD scores), who could 'afford' to stick to their preference, completed sentences according to their natural preference more often. Participants who were less efficient at processing auditory information (as indicated by their lower TPD scores) completed sentences less often according to their natural preference, instead weighing semantic context more heavily. This could be an indication that participants with better auditory perception have never needed to adopt a listening strategy that checked lexical context against their syntactic preference, since they could always accurately rely on their comprehension in auditory conversational contexts to keep them 'in the loop.' People with less auditory perceptive ability may be used to having to rely on other strategies in everyday comprehension, setting aside their own syntactic preferences in favor of more frequent context updating (such as semantic cues) in their comprehension. Thus the current result, that semantics influenced syntactic preference more in people who have less auditory perceptive ability, could be a reflection of listening and conversation strategies developed based on the degree of challenge those participants usually face in auditory communication.

In sum, attachment preference varied unsystematically among native German speakers, and simply asking participants which noun they preferred to modify with a relative clause was a reliable method for these sentences, and can be used to code a baseline preferred-syntax condition in the EEG Experiment 3c. Therefore in the subsequent ERP experiment, relative clause attachment preference will be assigned individually based on a post-hoc questionnaire.

The global intentions of Study 3 were to see whether metrical context could facilitate the integration of non-preferred syntactic structure (regardless of which of two alternate structures is preferred), and to take into account individual differences in Verbal Working Memory and Time and Pitch Discrimination. Now that the language syntax ERP coding has been determined, the next section recapitulates some relevant findings of the P600 and individual differences before presenting the ERP Experiment 3c.

## 8.4 Experiment 3c: ERP responses to sentences with varying metricality and syntax

### 8.4.1 Introduction

As introduced at the beginning of this study (and see Section 3.1.2), the integration of syntactically ambiguous or non-preferred structure in language reportedly elicits a P600 effect (Kaan & Swaab, 2003; Osterhout & Holcomb, 1992). Thus here it was predicted that non-preferred relative clause attachment would elicit a P600.

The syntactic P600 is generally influenced by the temporal and metrical composition of words (e.g., Schmidt-Kassow & Kotz, 2008, 2009b). One ERP study (with non-musicians) compared sentences with regular and irregular metrical contexts, and found the P600 (elicited by a high-cost, object relative clause attachment) to be reduced when occurring in the regular meter context (Roncaglia-Denissen et al., 2013). This was interpreted such that the regular meter facilitated syntactic integration, consistent with the current neurocognitive entrainment approach that predicts entrained syntactic processes to be facilitated (see Section 3.2). Considering that these EEG participants entrained more to the current regular-metricality sentences (Experiment 2b), it was predicted that a reduced P600 elicited by the non-preferred relative clause attachment would be a marker of facilitated syntactic integration in the regular compared to irregular metricality condition.

Finally, the amplitude of a P600 elicited by similar, high-host relative clause attachment was correlated with individual working memory measures previously in non-musicians (Bornkessel et al., 2004). Thus, here it was expected that individual differences measured by cognitive factors Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD) would correlate with P600 amplitude.

#### Hypotheses

1. A P600 should be elicited by the non-preferred syntax condition (according to participants' self-report of attachment preference).

2. Syntax should interact with metricality (regular-meter sentence context should facilitate processing, resulting in a reduced P600 in the regular-meter condition).

3. Individual differences should play a role for the P600 amplitude (cognitive factor scores should correlate with P600 amplitude).

### 8.4.2 Methods

#### Participants

Data from Experiment 2b Sentence sessions were analyzed in this experiment. Thus, participants were the same twenty-eight right-handed, monolingual native German speakers (see Section 7.2.2) whose seven diagnostic scores were previously collected in Study 1. Methods relevant to this ERP analysis are listed below.

#### Materials

Sixty sentences were used in this experiment (used in Experiment 1b, Experiment 3b and Study 2) with four versions each (for a total 240 sentences): Each syntactic condition contained an identical first phrase with an established subject and genitive possessor, followed by a second phrase beginning with an ambiguous relative pronoun which was resolved in the final verb as being either attached high (to the subject) or attached low (to the possessor), thus creating either a subject or possessor relative clause. The two initial syntax conditions were thus high or low, subsequently coded as *preferred* and *non-preferred* based on a post-hoc questionnaire (see Experiment 3b). The two metricality conditions were *regular* (predictably weighted syllables in groups of three beats) and *irregular* (less predictably weighted syllables in groups of two, three, or four beats; see Experiment 2a for a detailed description of the metricality conditions). Example sentences are shown in Figure 8.6. See Section 5.3.2 for a full stimuli description and Section 3.1.2 for a review of relative clauses.

Fillers were presented in 50% of the trials. As mentioned in Experiment 2b, the metricality conditions of fillers were similar to stimuli items. With respect to the syntax conditions, filler items consisted of a similar two-phrase structure that satisfied task demands, also with a syntactic ambiguity (the same ambiguous relative pronoun as in stimuli resolving to either subject- or object relative clause attachment) to ensure that participants did not distinguish stimuli from fillers and develop some sort of strategy. See Figure 5.10 in Chapter 5 for an example. All filler items are listed in Appendix B.2.

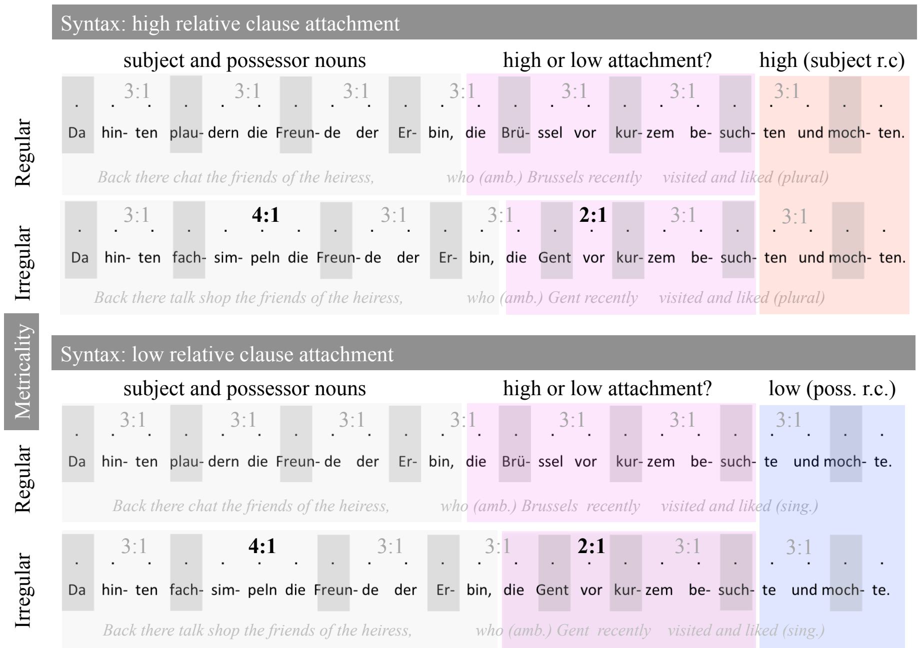



*Figure 8.6: Language stimuli design. r.c. = relative clause.* ***Syntax****: Sentence stimuli consisted of two clauses, the main clause containing a subject (e.g. friends) and a genitive possessor (e.g. [of the] heiress). The relative clause — with an ambiguous relative pronoun — was either attached to the subject (high attachment) or to the possessor (low attachment), resolved by the final plural or singular verb conjugation.* ***Metricality****: Regular sentences have even groups of three syllables, and irregular sentences have alternating groups of two, three, or four syllables. The words distinguishing the two metricality conditions were semantic equivalents with different syllable counts. Strong syllables are shaded.*

#### Procedure

Participants first completed a short computerized training (8 trials), with feedback, where they explicitly had to answer the question “Does the second phrase contain the same subject as the first phrase?” Participants had to successfully complete the training before continuing with the experiment, correctly answering the question after all condition types (for stimuli and fillers). To avoid artifactual brain responses related to motor preparation and execution, instructions during the EEG experiment were to mentally answer the question after each sentence, and to answer with a button press when the question was presented visually after 10% of the trials. Behavioral data for each trial was collected in a separate post-hoc session (no recorded EEG) wherein they repeated the entire experiment with the instruction to answer the question via button press after each trial.

In both EEG and behavioral sessions, blocks were presented with stimuli and fillers of only one metricality condition separated by the question and a self-paced break, the order of blocks counterbalanced across participants. Items were pseudo-randomized such that no more than four stimuli or fillers appeared successively, nor did any syntactic condition appear more than four times successively.

### 8.4.3 Analysis and results

#### Alternate ending preference

In the sentence-completion experiment (Experiment 3b) it was shown that native German speakers vary in attachment preference in the auditory domain, and that participants accurately self-report their own attachment preference. Based on this evidence, the preferred syntax condition was determined on a per-participant basis in the current experiment. If participants stated post-hoc a clear preference for one alternate structure over the other, their listed preference was assigned as ‘preferred.’ If participants stated no preference, low-attachment relative clauses were designated as ‘preferred.’ This latter group exhibited lower TPD scores than the other two groups, supporting results from Experiment 3b (Section 8.3.3). The final participant assignment was 12 high-attachment and 11 low-attachment (six reported preference, five no preference).

#### ERPs

In preprocessing, one electrode within a ROI was interpolated for one participant. Otherwise preprocessing was identical to the sentence data in Experiment 2b. Participants with fewer than 40 trials remaining after trial rejection were excluded from further analysis (final group size was 23). The trigger or zero-point in Experiment 3c was the onset of the critical

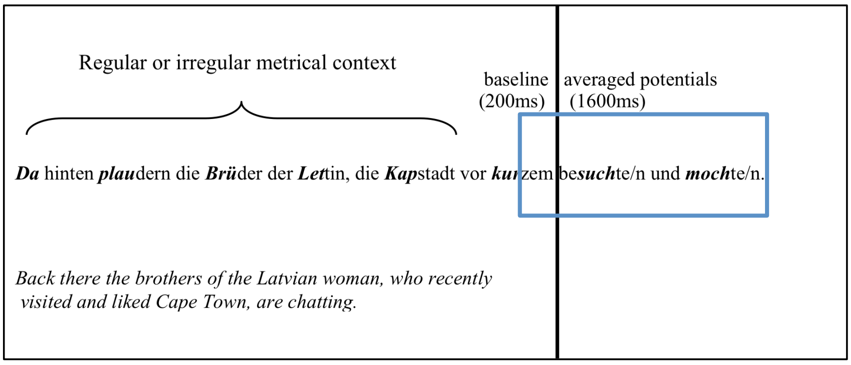



*Figure 8.7: Language ERP epoch. The epoch is marked by the blue box. The vertical black line indicates the trigger.*

verb 'besuchten' or 'besuchte'/ *to visit (plural or singular)*. As with music epochs, the preceding 200 ms of data were normalized to create an absolute baseline. Epochs extended for 1600 ms, which included the remainder of the sentence and the following ISI.

Cluster-based permutations (Oostenveld et al., 2011) were performed with grand averages of preferred- and non-preferred- syntax conditions (non-preferred minus preferred relative clause attachment, collapsed across meter type). The same ROIs as in Experiment 3a were used to examine any found syntax effect. Averaged-over-ROI, averaged-over-time-windows amplitude values were used in (1) a midline ANOVA with factors Syntax (preferred, non-preferred), Metricality (regular, irregular), and AntPost (anterior midline, posterior midline) and (2) a lateral ANOVA with factors Syntax (preferred, non-preferred), Metricality (regular, irregular), Hemisphere (left, right) and AntPost (anterior, middle, posterior).

A positivity was found between 1032 ms and 1206 ms. The late positivity was elicited in response to the non-preferred final verb (determined on a per-participant basis, as described above). Despite a somewhat later latency, this effect was labeled a P600 (e.g., Osterhout & Holcomb, 1992; see Section 8.4.4 for a discussion of the label).

As the amplitude values of the elicited effect were not normally distributed for all ROIs, data were log-transformed, and ANOVAs performed on the transformed dataset. All other reported values (mean amplitude, etc.) represent untransformed data. The results of Syntax effects in midline and lateral ANOVAs may be found respectively in Tables 8.7 and 8.8. See Experiment 2b for examination of the Metricality effect in a separate frequency analysis.

**P600** In the P600 time window, a main effect of Syntax was found both in the midline-ROI ANOVA ($F(1,22) = 15.20$, $p = .001$, partial $\eta^2 = .409$) and lateral-ROI ANOVA ($F(1,22) = 7.68$, $p = .011$, partial $\eta^2 = .259$), supporting the findings of the cluster-based permutation test. The ERP waveform is shown in Figure 8.3 and the amplitudes of the P600 per ROI are

*Table 8.7: ANOVA of language P600 ERP amplitudes, midline ROIs. S = Syntax. M = Metricality. A = AntPost. ANOVA performed on log-transformed values. Significant effects in bold.*

| Effect | $F(1,22)$ | partial $\eta^2$ | $p$ |
|---|---|---|---|
| S | 18.80 | .46 | **<.001** |
| S x M | 5.71 | .21 | **.026** |
| S x A | 0.44 | .02 | .52 |
| S x M x A | 4.35 | .17 | **.049** |

*Table 8.8: ANOVA of language P600 ERP amplitudes, lateral ROIs. S = Syntax. M = Metricality. A = AntPost. H = Hemisphere. ANOVA performed on log-transformed values. Significant effects in bold.*

| Effect | $F(1,22)$ | partial $\eta^2$ | $p$ |
|---|---|---|---|
| S | 13.71 | .38 | **<.001** |
| S x M | 4.08 | .156 | **.056** |
| S x H | 0.30 | .01 | .59 |
| S x M x H | 0.07 | .003 | .80 |
| S x A | 1.73 | .14 | .20 |
| S x M x A | 2.15 | .17 | .14 |
| S x H x A | 0.17 | .016 | .847 |
| S x M x H x A | 1.31 | .111 | .292 |

*Table 8.9: Mean P600 amplitudes ($\mu$V; ± SD), non-preferred minus preferred syntactic structure, per ROI.*

| ROI | P600 | |
|---|---|---|
| anterior midline | 1.08 | (± 1.25) |
| posterior midline | 1.01 | (± 1.40) |
| left anterior | 0.53 | (± 1.02) |
| right anterior | 0.54 | (± 0.99) |
| left middle | 0.65 | (± 0.81) |
| right middle | 0.70 | (± 0.99) |
| left posterior | 0.68 | (± 1.29) |
| right posterior | 0.72 | (± 1.25) |

listed in Table 8.9.

**The impact of Metricality on the P600 amplitude** The P600 effect was reduced in amplitude and more focally posterior in the irregular compared to regular condition, partially supporting the hypothesis that Metricality would interact with Syntax. The interaction Syntax x Metricality occurred in both midline and lateral ANOVAs (midline: $F(1,22) = 5.71$, p =.026, partial $\eta^2 = .206$; lateral: $F(1,22) = 4.08$, $p = .056$, partial $\eta^2 = .156$). Specifically, when the metrical context was irregular, the P600 was reduced in amplitude, which was revealed when the Syntax effect was evaluated separately in regular and irregular metricality conditions in midline and lateral ANOVAs (regular: midline, $F(1,22) = 17.39$, $p < .001$, partial $\eta^2 = .442$; regular, lateral: $F(1,22) = 14.94$, $p = .001$, partial $\eta^2 = .404$; irregular: midline, $F(1,22) = 6.58$, $p = .018$, partial $\eta^2 = .230$; irregular: lateral, $F(1,22) = 3.24$, $p = .086$, partial $\eta^2 = .128$). The amplitudes of the P600 per metricality condition per ROI are listed in Table 8.10.

**The impact of Metricality on the P600 amplitude and distribution** In addition to a smaller P600 amplitude in the irregular condition, the P600 was more focally distributed in posterior ROIs in the irregular compared to regular condition (Figure 8.9), indicated by the midline Syntax x Metricality x AntPost interaction ($F(1,22) = 4.35$, $p = .049$, partial $\eta^2 = .165$). This was deduced by separating the ANOVA into anterior and posterior ANOVAs: in the anterior region the Syntax x Metricality interaction was significant ($F(1,22) = 8.01$, $p = .010$, partial $\eta^2 = .267$), whereas it was not significant in the posterior region ($F(1,22) = 1.22$, $p = .282$, partial $\eta^2 = .053$). That is, once the meter was regular, the large area of significance over the central scalp area subsided.

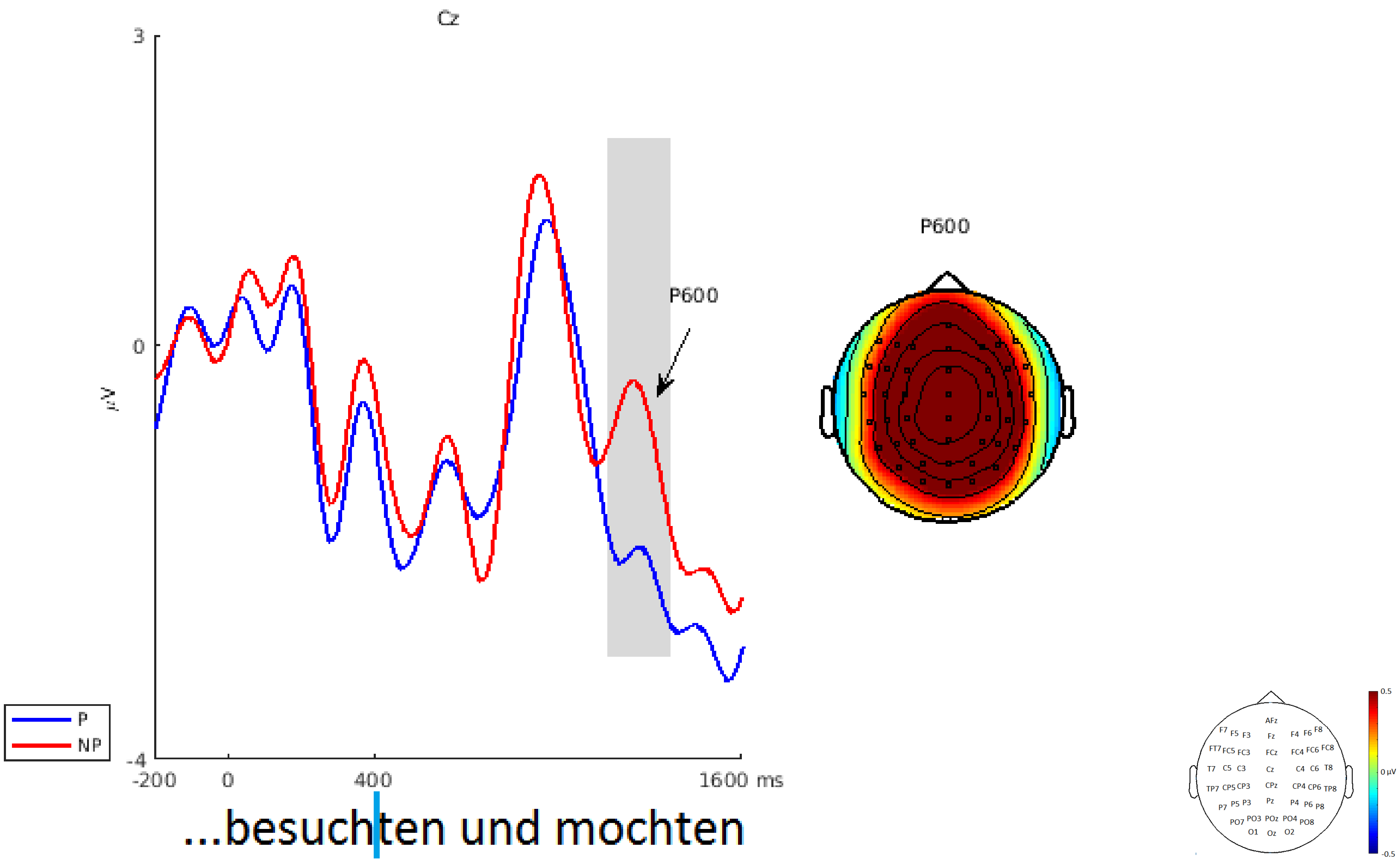



*Figure 8.8: The language syntax P600. P = Preferred. NP = Non-preferred. While ERPs were time-locked (0 seconds) to the beginning of the critical verb, which consistently indicated relative clause attachment, effect latency of the P600 was consistent with processes that were triggered by the critical morpheme (vertical blue line, ca. 400 ms). The effect time window is shaded in gray. Scalp topography shows the shaded region.*

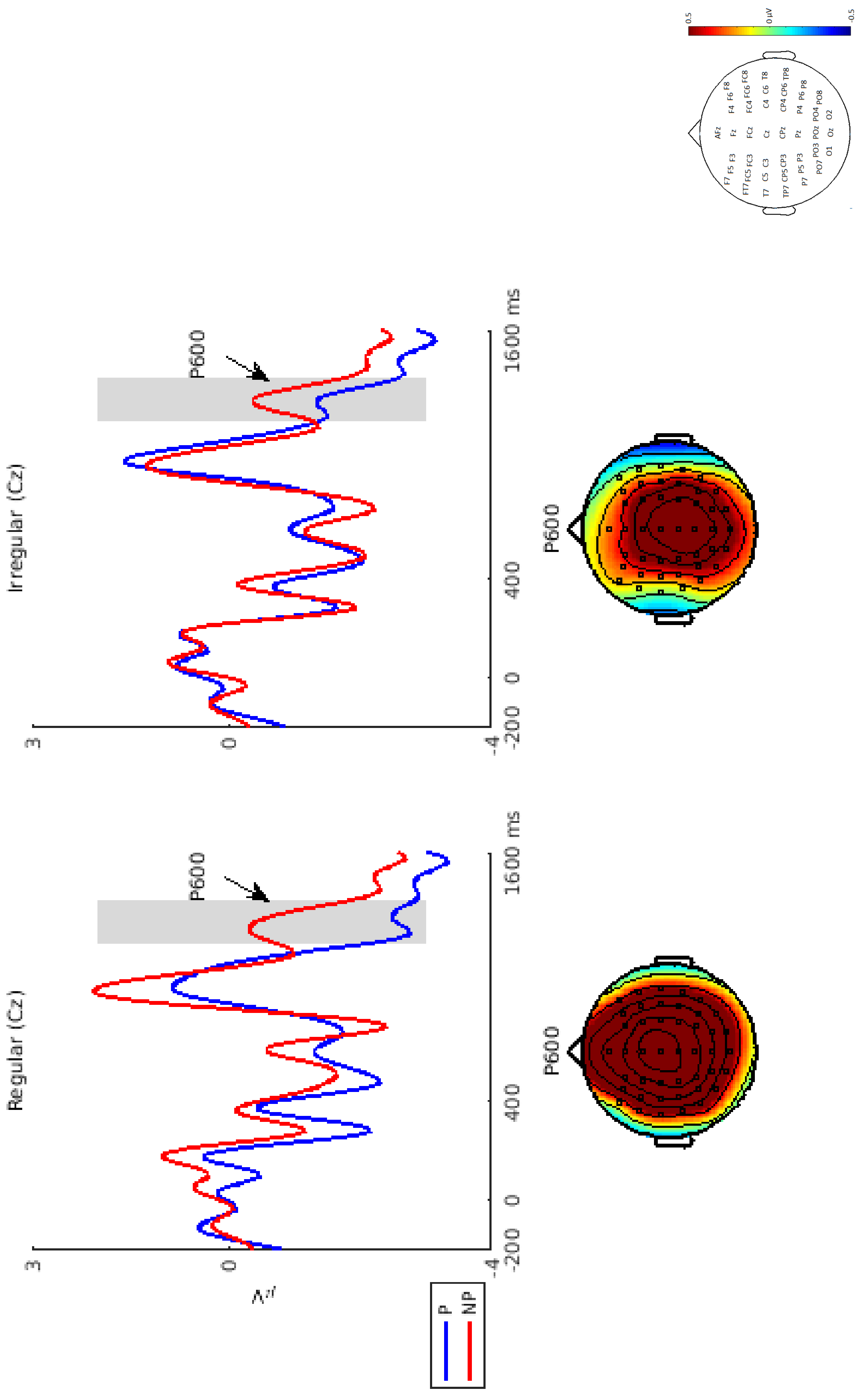



*Figure 8.9: Syntax x Metricality interaction in the language P600. P = Preferred. NP = Non-preferred. The language P600 was reduced in size and distribution in the irregular-metricality condition. The effect time window is shaded in gray. Scalp topography shows the shaded region.*

*Table 8.10: Amplitude of language P600 difference-waves (non-preferred minus preferred syntactic structure), averaged per metricality condition per ROI (μV; ± SD)*

| ROI | Regular | | Irregular | |
|---|---|---|---|---|
| anterior midline | 1.56 | (± 1.69) | 0.56 | (± 1.42) |
| posterior midline | 1.20 | (± 1.83) | 0.82 | (± 1.47) |
| left anterior | 0.78 | (± 1.14) | 0.30 | (± 1.47) |
| right anterior | 0.82 | (± 1.28) | 0.27 | (± 1.09) |
| left middle | 0.98 | (± 1.21) | 0.33 | (± 0.97) |
| right middle | 0.97 | (± 1.19) | 0.43 | (± 1.25) |
| left posterior | 0.83 | (± 1.62) | 0.52 | (± 1.34) |
| right posterior | 0.89 | (± 1.50) | 0.55 | (± 1.45) |

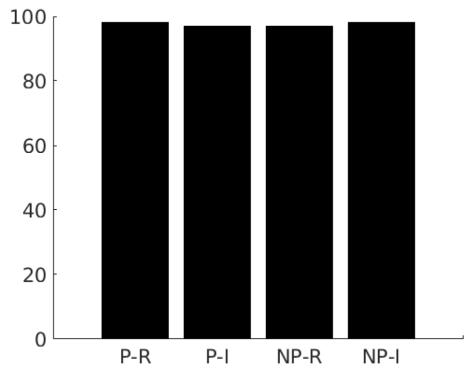



*Figure 8.10: Accuracy per language condition. P = Preferred. NP = Non-preferred. R = Regular. I = Irregular.*

### Behavior

In a separate behavioral session, at least two days after the music EEG experiment, participants repreated the experiment with the same task (to answer whether the subject of the second phrase was the same as the subject in the first phrase), with the exception that participants answered with a button press after each item as opposed to after 10% items. Accuracy scores per condition verified that participants were able to perform the syntax task during EEG (Figure 8.10). Near-100% accuracy suggested a ceiling effect. A 2 x 2 ANOVA performed on log-transformed scores with factors Syntax (preferred, non-preferred) and Metricality (regular, irregular) yielded no main effects or interactions (Syntax: $F$(1,22) = 0.00, p = .96, partial $\eta^2$ = .00; Metricality: $F$(1,22) = 0.00, $p$ = .95, partial $\eta^2$ = .00; Syntax x Metricality: $F$(1,22) = 2.24, $p$ = .15, partial $\eta^2$ = .09 ). These results indicated that the online processes captured by the ERPs were not captured by the behavioral data.

*Table 8.11: Correlation of language P600 difference waves per ROI with cognitive factor scores. Pearson's R, two-tailed correlations. No significant correlations. Correlations were omitted in ROIs where no P600 effect was significant. VWM = Verbal Working Memory, TPD = Time and Pitch discrimination.*

| | Regular metricality | | Irregular metricality | |
|---|---|---|---|---|
| ROI | VWM | TPD | VWM | TPD |
| anterior midline | .06 | .28 | .02 | -.15 |
| posterior midline | - | - | - | - |
| left anterior | .23 | .31 | .27 | 0 |
| right anterior | .09 | .22 | -.07 | -.27 |
| left middle | -.003 | .16 | .14 | -.15 |
| right middle | -.09 | .27 | -.13 | -.33 |
| left posterior | | - | - | - |
| right posterior | | - | - | - |

The post-hoc behavioral sessions demonstrated that participants correctly followed the instructions, but accuracy scores did not replicate any effects from the ERP. As in music, absence of any effects may have been due to the fact that participants had already been exposed to all stimuli items, therefore the relative difficulty was reduced (or less expected items were in essence expected, since they were heard before). On the other hand, the behavioral measure simply may not have captured the online processes revealed by ERP findings, which are discussed below. Either way, the null result in the behavior does not detract from the ERP findings, and the behavioral results will not be discussed further.

**Individual differences**

As with music, factor scores from Study 1 were correlated with the P600 amplitude, separately for regular and irregular metrical context (Table 8.11) in order to test whether the cognitive processes associated with these two factors individually influenced the amplitude of the P600. There were no significant factor score correlations with amplitude of the effect, rejecting the hypothesis that individual differences play a role in the amplitude of this effect.

### 8.4.4 Discussion: Language ERPs

Experiment 3c presented musicians with sentences that had preferred or non-preferred syntactic endings presented in regular or irregular metrical contexts. A P600 was hypothesized to reflect integration of non-preferred relative clause attachment, be reduced in the regular compared to irregular sentences, and its amplitude was expected to correlate with cognitive factors Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD).

#### Results summary

The current hypotheses were partially supported. The non-preferred (but not incorrect) relative clause attachment, determined individually based on participant feedback (per Experiment 3b), elicited the hypothesized P600 (Figure 8.8). Thus previous results with respect to syntactic preference were replicated, incidentally supporting reports that complex but correct structure involves processes whose P600 elicitation is distributed in antero-central (e.g., Friederici et al., 2002; Osterhout & Holcomb, 1992) as opposed to posterior distribution (see Section 3.1.2). Findings also established the P600 as reflecting auditory high vs. low attachment preference in German subject vs. possessor relative clauses.

The P600 was reduced in amplitude in the irregular metricality condition, which was opposite to the hypothesized direction. This result speaks for a mediating role of metrical and temporal predictability in language syntactic integration (e.g., Schmidt-Kassow & Kotz, 2008, 2009b), however; unlike previous results showing regular meter to facilitate processing in language (Cason & Schön, 2012; Roncaglia-Denissen et al., 2013; Rothermich et al., 2012), here it appears that syntactic integration of the non-preferred structure was easier in the irregular condition. Neither cognitive factor score correlated with P600 amplitude, regardless of metrical context, indicating that individual differences did not impact musicians' language syntactic integration with these stimuli. These results are discussed below, where the deviations from the hypotheses, where they occurred, are largely attributed to the musical expertise.

#### Integrating syntactically non-preferred structure in language

Although the P600 usually is earlier, the late positivity elicited by current sentence stimuli was labeled a P600. The critical verb had three syllables, the third carrying the actual identifying information as to the conjugation. Taking into account the length of the first two syllables, the latency of the effect is appropriate to identification as the P600 in response to syntactic incongruities reported elsewhere (e.g., Hagoort, Wassenaar, & Brown, 2003). Moreover, the distribution of the current positivity is in line with previous reports of the P600 (e.g., Osterhout & Holcomb, 1992). Thus the initial hypothesis was supported: a P600 effect was elicited, reflecting the integration of a non-preferred complex syntactic structure.

The P600 elicited here is in line with countless psycholinguistic paradigms that tested integration of syntactic structure (see Gouvea et al., 2010), specifically, syntactic integration of ambiguous relative clause attachment and of non-preferred syntactic structures (e.g., Gunter, Friederici, & Hahne, 1999; Kaan & Swaab, 2003; Osterhout & Holcomb, 1992). Such studies attribute the P600 to the integration of syntactic constituents with their preceding context. The current paradigm expanded the P600 repertoire to represent abstract structural preference in the auditory domain, independent of specific linguistic identity of constituents (since opposite attachment conditions were preferred across the group, and the presence of a P600 validated the approach to individually assign preferred conditions).

#### The impact of metrical context on syntactic integration

While the metrical context of the sentences unquestionably impacted the resolution of the local ambiguity, the direction of the interaction was opposite to the hypothesis. Previous literature reported entrainment to facilitate linguistic processing (Cason & Schön, 2012; Roncaglia-Denissen et al., 2013; Rothermich et al., 2012), marked by reduced ERP effects. Thus here, the reduced P600 in irregular sentences implied that the regular meter hindered rather than facilitated syntactic integration. Considering that these musician participants entrained more to regular than to irregular sentences (Experiment 2b), it would seem that, in conflict with previous reports, the greater entrainment to the regular sentences was detrimental to syntactic integration.

While the above mentioned ERP studies report facilitation by regular meter in non-musicians (Cason & Schön, 2012; Roncaglia-Denissen et al., 2013; Rothermich et al., 2012), one recent behavioral study (Menninghaus et al., 2015) suggested that cognition cannot be considered monolithic when it comes to the impact of meter. Menninghaus et al. (2015) presented participants with proverbs that were manipulated according to meter, with the task to rate the aesthetic qualities vs. comprehensibility of each proverb. While regular meter facilitated certain aspects of cognition (e.g., perceived beauty, succinctness, and persuasiveness), comprehension of the linguistic content was actually detrimented by the regular meter. Accordingly in the current experiment, the regular meter, while facilitating general perception of and neural entrainment to the sentences, may have ultimately hindered access to the linguistic content of the sentences.

In line with the interpretation that the regular condition was ‘too regular’ and distracted musicians when integrating the syntactic structure, Chapter 4 presented evidence from the literature that musicians have enhanced auditory processing of language compared to non-musicians, including that they were more sensitive to metrical structure of words (Marie et al., 2011). Thus it is not inconceivable that a metrical context in language that is akin to musical meter (a strong beat every three beats) would be distracting to musicians, and draw some attentional resources away from syntactic integration. If this is the case, than this represents an extension of previously hindered semantic processing in metered stimuli (Menninghaus et al., 2015) to hindered syntactic processing.

A second and compatible possibility is that the irregular meter entrained participants as well, to a lesser extent than the regular meter, and that this lesser 'dosage' of regularity facilitated syntactic integration (Roncaglia-Denissen et al., 2013). The irregular metrical context in current sentence stimuli was still more regular than ordinary conversation, resembling nursery rhyme structure (a 3-4-3, 3-2-3 beat groups), a pattern likely predictable enough for musicians to entrain. Accordingly, the EEG frequency analysis (Experiment 2b) did not exclude the possibility of entrainment to the irregular sentences, but rather showed that entrainment was comparatively greater in regular as opposed to irregular sentences. The musicians also did not show any difference in RT or accuracy when detecting syllables in regular and irregular conditions (Experiment 1b), consistent with entrainment to both conditions. The null result of individual differences, discussed presently, compliments this interpretation.

In sum, the regular sentences seem to have been 'too regular'; while they were entrained more than regular sentences (Study 2), the increased entrainment to regular metricality proved to encumber syntactic integration as opposed to facilitating it. A secondary possibility is that irregular sentences were entrained more than ordinary prose would have been, and that this lesser entrainment in turn facilitated syntactic integration.

#### Individual differences in syntactic integration

When P600 amplitude was correlated with factor scores, there were no significant correlations in either metrical context, rejecting the hypothesis that individual differences play a role in the amplitude of this effect. This is despite previous reports of individual differences impacting P600 amplitude in relative clause attachment in non-musicians (Bornkessel et al., 2004). Although it is possible that neither Verbal Working Memory (VWM) nor Time and Pitch Discrimination (TPD) is involved in this type of ambiguity resolution, an explanation consistent with arguments thus far is that the musicians' language syntactic integration is too homogeneous to capture individual differences, due potentially to globally enhanced linguistic processing (e.g., Strait et al., 2014), entrainment to both conditions (as previously suggested) or a combination thereof. Above it was predicted that conversational sentence meter (with analogous syntactic manipulation) would elicit a larger P600 in musicians, compared to both current metricality conditions; accordingly, such a future paradigm might also show individual differences in VWM and TPD in the P600 amplitude elicited in a conversational meter context.

So far, Study 3 has presented and discussed the results of Experiment 3a and Experiment 3c separately. The following sections present the music and language results side-by-side and interpret them in a broader cross-domain perspective.

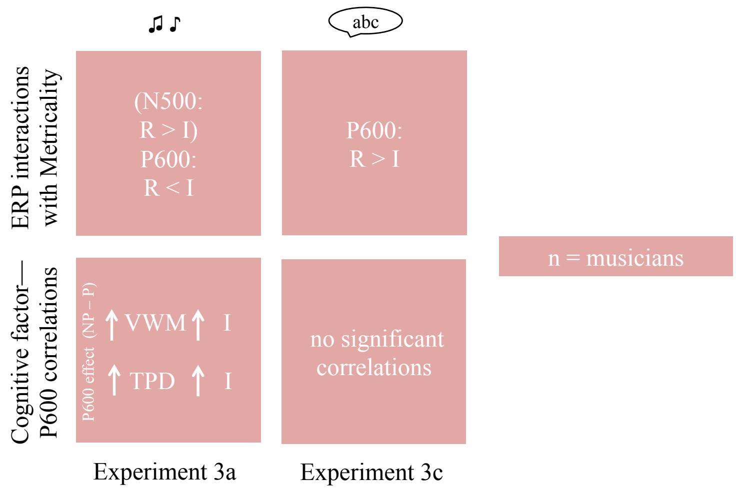



*Figure 8.11: Study 3 ERP results summary. R = regular, I = irregular, NP = non-preferred syntax, P = preferred syntax. In lower left panel, two upwards arrows represent a positive correlation between listed variables.* ***Experiment 3a, Music****: the P600 effect was reduced in the regular compared to irregular metricality condition.* ***Experiment 3c, Language****: the P600 was reduced in the irregular compared to regular metricality condition, opposite from the music condition.* ***Individual differences****: only in music, cognitive factors correlated positively with P600 effect waves (non-preferred minus preferred) in the irregular condition.*

## 8.5 Discussion: ERPs across domains

### 8.5.1 Experiment 3a and Experiment 3c results summary

Experiment 3a and 3c results are summarized in Figure 8.11. The two ERP experiments compared the syntactic integration of non-preferred (error-free) syntactic structures in regular as opposed to irregular metrical contexts, in melodies and sentences. Individual differences in auditory abilities were assessed in relation to ERP amplitude. It was found that the hypothesized ERP, the P600, indeed represented the integration of a non-preferred (but not incorrect) structure in both domains. The P600 was modulated by metrical context in both domains, however; while the direction of the syntax-metricality interaction was as hypothesized in music (that regular context facilitated processing, indexed by a reduced amplitude), the direction was opposite in language (a reduced P600 occurred in the irregular metrical context). Cognitive factors Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD) correlated with music P600 amplitude in irregular metrical context only, but did not correlate with the language P600 in either metrical context. In music, an additional negativity was elicited by the non-preferred endings, interpreted as an effect called the N500. The following discussion will focus on specifically hypothesized-for effects (see above Section 8.2.4 for discussion of the N500).

### 8.5.2 Shared resources for processing syntactic preference in music and language: SSIRH

Experiments 3a and 3c showed the syntactic integration of non-preferred structure in both melodies and sentences to be represented by a late positivity, interpreted as a P600. This result generally compliments previous suggestions that the P600 reflects syntactic integration in both domains (e.g., Patel et al., 1998), with the added paradigmatic advantage that error-free syntax conditions indeed compared syntactic integration within-participant, and not violation-detection mechanisms.

To this extent the SSIRH (Patel, 2003, and see Section 3.1.3 ) is supported, such that integration of non-preferred (but not incorrect) structure is represented by a similar evoked potential in music and language domains. Since no violation was present, current results suggest the resources responsible for integrating musical and linguistic syntactic constituents to operate with a set of preexisting preferences that both guide expectancies and supersede grammatical correctness. In other words, music–language comparison extends to ecologically valid, subtle syntactic manipulations that could affect everyday comprehension, such as how the natural preference of a listener guides him/her to expect certain harmonic and linguistic structures in the face of multiple interpretations.

### 8.5.3 The impact of metrical context on cross-domain syntax processing: SSIRH meets OPERA hypothesis

It was hypothesized that regular metrical context (weighted note/syllable every three beats) would facilitate syntactic integration of melodies or sentences more so than irregular metrical context, which had a lesser degree of predictability (weighted note/syllable every two, three or four beats). This phenomenon is explained by the current neurocognitive entrainment hypothesis (Section 2.3) such that acoustic information with more temporal predictability would beget entrainment (e.g. Large & Jones, 1999), and entrained neural responses in turn have a greater degree of neural synchrony, including hierarchical phase coupling among delta, theta and gamma bands (Giraud & Poeppel, 2012; Schroeder & Lakatos, 2009), which generally results in more efficient cognition (Singer, 2013; Varela et al., 2001). Such facilitated cognition in contexts of metrically predictable auditory information has been seen in language (Roncaglia-Denissen et al., 2013; Rothermich et al., 2012) and argued to underlie the benefits of music listening (Thaut et al., 2005), but before the current experiment has never been tested in music and language in a within-participants design.

While the concept of shared syntactic resources was reinforced by the P600 response to non-preferred syntactic structure in both domains, these syntactic resources did not respond analogously to metrical regularity in both domains. Instead, what was found was an interaction with metrical context in the two domains, in opposite directions. This is surprising in light of the current neurocognitive entrainment hypothesis, which proposed entrainment to facilitate syntactic processing regardless of domain, based on multiple converging theories from both domains (e.g., Giraud & Poeppel, 2012; Large & Jones, 1999; Pitt & Samuel, 1990; Schroeder & Lakatos, 2009). However, an explanation for the discrepant results may lie in a recent theory (Patel, 2011) that proposed musicians, due to a combination of training and experience, to have asymmetrical perception of music and language.

The OPERA hypothesis (Patel, 2011) brought together a large body of empirical evidence (including many of the studies cited in Section 4.1) and postulated that musicians' improved perception and performance in the language domain is due to five key conditions: (1) an overlap (O) of brain networks serving the aspect (acoustic feature) of language that is improved and the acoustic feature of music that is trained by the musician; (2) demands on the nervous system that perception of this acoustic feature be more precise (P) in music than in usual language encounters; (3) emotional (E) reward from and (4) repetition (R) of active listening and performance which feed off of each other to promote long-term plasticity; and (5) the attention (A) to musical features (as opposed to passive listening) in the lifestyle of a musician that affords the opportunity to improve the perception of musical features in the first place, a starting-point for the plasticity which leads to better overall perception. Thus in sum, the demands on musicians to constantly and accurately perceive acoustic aspects of music (for 'normal' musical communication) translate to a heightened perception of similar linguistic acoustic features (above and beyond the perception required for 'normal' linguistic communication). This hypothesis was not tested in the current paradigm, e.g., no data was collected to verify the presence of all five OPERA elements in participants' musical history. But, if one assumes the presence of those elements (which is reasonable), the current data certainly align with the phenomena described by the hypothesis, particularly if the precisely trained, overlapping acoustic feature is amplitude envelope.

The amplitude envelope was the feature according to which meter was defined and analyzed in this dissertation (see Experiment 2a). Since perception of the amplitude envelope is a skill commonly required for musicians, to perceive not only aspects of meter but also to distinguish timbre quality in various instruments (see Patel, 2011), it follows that an increased sensitivity in language compared to music would cause the current language meter to acoustically pop out, while the same meter in music is perceived as 'normal'.

While meter serves as an organizational principle in both music and language (see Chapter 2), (Western Tonal) musical meter is highly regular as a baseline, whereas (Germanic) language meter is highly structured in circumstances that use the meter as a special tool of delivery as opposed to baseline communication. For example meter in language may be highly regular for rhetoric or poetry, in order to deliver a memorable, persuasive message (such as in artistic, religious, or political settings; Menninghaus et al., 2015; Obermeier et al., 2016). Thus, regular meter in music often simply organizes the harmony, whereas regular meter in language often accompanies some added dimension of communication beyond the words that comprise it. In musicians who already have a boosted perception of linguistic meter, the highly regular structure in the current stimuli could trigger cognitive processes that compete with or impede syntactic integration. In other words, along with increased entrainment to regular linguistic meter, musicians might be trying to interpret an over-salient meter along an additional dimension beyond the surface meaning of the words.

The OPERA hypothesis, which would predict that musicians have 'normal' processing of musical meter and above-normal processing of linguistic meter (providing that certain practice conditions were met), is supported by the combined music- and language-experiment ERP results. In line with these points, behavioral data from Study 1 suggested that non-musicians did benefit from the regular meter in language; non-musicians would not have (over-) trained amplitude envelope perception and therefore would benefit from metrical regularity by the hypothesized degrees (more entrainment to more regularity), much as musicians did with music. A subsequent hypothesis for future research would be that non-musicians would have facilitated syntactic integration in the regular-meter sentences of the current stimuli.

### 8.5.4 Individual differences in the music domain

The correlations between cognitive factors VWM and TPD and music and language P600 amplitudes tell a similar story to the above-described one: Hypotheses based on non-musicians in language studies, e.g. that individual differences in working memory correlate with P600 amplitude, are only supported in musicians when applied to the music domain. As predicted, individual differences surfaced in syntactic integration of non-preferred harmonic structure when cognition was not facilitated by regular metrical context. Contrary to predictions based on non-musician syntax ERP evidence (Bornkessel et al., 2004, see Chapter 4), no individual differences surfaced at all with respect to syntactic integration of non-preferred linguistic structures in either metrical context. Thus it would seem that in musicians, cognitive resources associated with perception of auditory features are gaged for music cognition much the way non-musicians' resources are gaged for language cognition (as per the OPERA hypothesis; Patel, 2011). Further research examining individual differences among musicians in other music- vs. language cognition tasks is necessary to see whether the current result (is simply a null language result, or) reflects a larger pattern in musicians' cross-domain cognition.

## 8.6 Conclusion

This ERP study addressed whether metrical context influences the integration of non-preferred syntactic structure in melodies and sentences, and how individual differences factor into this reanalysis. The fact that a P600 was elicited by the structural complexity in each domain lent credence to the SSIRH (Patel, 2003), though diverging results regarding the influence of meter suggest that (musicians') syntactic integration mechanisms do not have straightforward comparability across domains: Intuitively, greater entrainment to regular meter in melodies (shown in Exp. 2b) facilitated the syntactic integration of non-preferred harmonic structure (as reflected by a reduced P600 amplitude; Exp. 3a), linking entrainment to facilitated cognition for musicians in the music domain. Counter-intuitively, greater entrainment to regular meter in sentences (shown in Exp. 2b) resulted in hindered syntactic integration, suggesting a more complex relationship between neural entrainment and cognitive facilitation for musicians in the language domain. Individual differences results suggested that, among the four experimental conditions, musicians most rely on Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD) abilities when doubly confronted with an unaccustomed, irregular meter and a non-preferred syntactic structure.

### 8.6.1 Cross-domain syntactic integration and metrical context

It may be concluded from this study's results that syntactic integration of non-preferred structure is analogous in language and music, but that musicianship seems to be too influential a factor when it comes to answering whether syntactic integration in both domains is similarly influenced (e.g. facilitated) by metrical regularity. Despite a paradigm that established very similar regular and irregular metrical contexts in music and language, musicians' perception of the meter was not the same in both domains. While previous evidence showed non-musicians' language syntactic integration to be facilitated by regular meter (Roncaglia-Denissen et al., 2013), current results suggested that musicians' improved auditory linguistic perception (Patel, 2011) confounded the entrainment to the metrical context in sentences, such that the current (semi-regular) irregular context provided ample temporal predictability for entrainment and facilitated processing, while the 'too regular' regular context was over salient and distracted them.

### 8.6.2 Individual differences

Cognitive factors indexing abilities in auditory perception and working memory only played a prominent role in syntactic reanalysis in the condition which was least ecologically valid, e.g. in melodies with an irregular metricality condition. Although the null language result does not necessarily mean that there were no individual differences, it conveniently supports the idea that musicians were able to supplement their cognition with entrainment to both regular and irregular metrical contexts (despite the fact that regular was distractingly regular), which lessened the cost of processes relying on those cognitive factors in both sentence conditions. The next chapter discusses the findings of this dissertation's three studies in a broader literary context.

# Chapter 9

# General Discussion

This dissertation investigated neurocognitive entrainment and syntactic comprehension in music and language. Accordingly, three core research questions were whether meter is perceived comparably across domains, whether meter comparably impacts syntactic integration in both domains, and whether individual differences impact the entrainment or the syntactic integration. To address these questions, auditory melody and sentence stimuli were created with a 2 x 2 design, whose regular and irregular metricality conditions served the cross-domain investigation of entrainment and whose alternate endings provided non-preferred syntactic structures to be comprehended. All participants underwent a diagnostic test battery that focused on aspects of expertise and auditory cognitive ability as part of the individual-differences approach. This chapter discusses the behavioral and EEG findings from seven experiments across three studies.

## 9.1 Summary of empirical findings

**Study 1** investigated individual differences and cognitive entrainment, using statistical methods to create individual-difference metrics based on a series of diagnostic tests (Experiment 1a), and behavioral methods to examine individual differences and the influence of expertise on cognitive entrainment to syntactically and metrically manipulated musical and linguistic stimuli (Experiment 1b). A factor analysis in Experiment 1a showed the diagnostic test scores to load onto two distinct factors, and each participant received factor scores, which enabled an individual-differences approach for further experiments. The cognitive functions associated with tasks in each factor gave rise to factor names Verbal Working Memory (VWM) and Time and Pitch Discrimination (TPD). Musical expertise was associated with improved TPD, and within musicians, the number of practice hours correlated positively with TPD. Improved TPD in musicians (and musicians who practice more) is in line with studies pointing to better perception of (musical and linguistic) acoustic features resulting from musical training (Bidelman & Krishnan, 2010; Koelsch et al., 1999; Kraus et al., 2009, and see Section 4.1). Results thus indicated that in the current participant group, musical expertise impacted perception of features related to complex auditory streams but did not necessarily involve verbal working memory in the online computation of this information.

Accuracy and reaction time data in Experiment 1b was used to evaluate whether metrical context impacted note- and syllable discrimination in musicians and non-musicians. In music, regular metricality offered a perceptual advantage over irregular metricality for both musicians and non-musicians, while in language, this same advantage for regular metricality was shown for non-musicians only. With respect to individual differences, musicians had generally better accuracy in melodies and faster reaction times in sentences compared to non-musicians. Expertise was associated with improved TPD (echoing results of Experiment 1a) and improved note discrimination in regular melodies. Participants with better TPD (irrespective of musical expertise) had, in turn, better note discrimination in regular melodies and better syllable discrimination irregular sentences. Considering the link between TPD and musical expertise, this latter finding may indicate musicians' improved performance in regular melodies and irregular sentences compared to non-musicians.

**Study 2** investigated the presence of meter-representative frequencies in melody- and sentence-stimuli audio files (Experiment 2a) and sought those same frequencies in the EEG of participants who listened to the stimuli (Experiment 2b), to see whether they neurally entrained to the meter. A frequency analysis was conducted in the amplitude envelope of auditory stimuli in Experiment 2a. Results showed clear peaks at frequencies representing the two hypothesized levels in the metrical hierarchy of melodies and sentences: the group level (weighted notes and syllables) and the individual level (individual eighth-notes and syllables). All of the peaks except that of the (marginal) individual peak in regular sentences were significant, validating the metrical levels intended from stimulus composition and validating further empirical investigation as to the perception of meter. Comparing regular and irregular metricality conditions within domain (at the group level, where metricality was defined according to placement of weighted beats), the playing strategy of the pianist likely contributed to the melodies' group peak power being less in regular compared to irregular conditions, while in the sentences group peak power was greater in regular than irregular conditions. This latter finding was attributed to a more temporally consistent weighted beat in the regular condition, which, in turn, should yield a higher peak power in the averaged spectra compared to the irregular condition (whose dispersed weighted beat should bleed power across more frequencies in averaged spectra). Comparing the metrical levels across domains, group peaks were lower than individual peaks in the melodies whereas group peaks were higher than individual peaks in the sentences. This indicates that perhaps objectively, the individual level would be better acoustically perceived than group in the melodies, and the group level would be better acoustically perceived than individual level in the sentences.

An EEG frequency analysis was conducted in Experiment 2b to see whether musicians entrained to the group (upper) and/or the individual (lower) metrical levels while they listened to the stimuli. In a frontal-central auditory ROI defined by N1 significance, results showed spectral peaks representing the 2:1 harmonic of the group level in regular music and regular language conditions, and the stimulated individual level frequencies in regular and irregular music conditions. These results were interpreted to mean that consistently across domains, participants entrained more to the group level in regular compared to irregular conditions. However, inconsistently across domains, music was entrained more than language at the individual level. An absence of any peaks representing group-level frequencies in either domain indicated that the <2 Hz frequencies fell below the threshold of data analysis due to 1/f noise in the EEG. Cognitive factors did not play a role in the height of significant music individual peaks. Since especially the cognitive resources associated with TPD (discrimination of temporal and pitch features) are surely involved in auditory entrainment to musical meter, this null result is attributed to the fact that there may not have been enough spread among the musicians in TPD scores and/or the ability to entrain to a musical tactus.

**Study 3** used ERP methods to investigate whether entrainment to regular meter facilitated syntactic integration of melodies (Experiment 3a) and sentences (Experiment 3c), necessarily beforehand determining the syntactic preference for current sentence stimuli with behavioral methods (Experiment 3b). In Experiment 3a (melodies), ERP results showed that a P600 reflects syntactic integration of non-preferred harmonic structures, not just violations or incongruities (which was a limitation in previous paradigms, e.g., Featherstone et al., 2013; Patel et al., 1998). Results further indicated that a regular metrical context facilitated syntactic processing compared to an irregular one: a P600 elicited by the non-preferred harmonic ending was reduced in the regular compared to irregular metricality condition. Both VWM and TPD cognitive factors correlated with amplitude of the irregular-metricality P600 effect, supporting the notion that while syntax processing was facilitated by entrainment to a regular metrical context, syntactic processing was unaided by the irregular metrical context and thus subject to individual differences in cognitive ability. A non-hypothesized negativity preceding the positivity was also elicited, reminiscent of the N500 reported in previous literature (e.g., Koelsch et al., 2000).

In order to accurately interpret language ERPs stemming from 'preferred' and 'non-preferred' relative clause attachment, Experiment 3b was conducted with an additional participant group comprised of German native speakers (with a range of musicality) to determine auditory subject- vs. possessor relative clause attachment preference. Results showed that participants generally preferred low-attachment (a possessor relative clause), supporting previous findings from the auditory domain (Augurzky, 2006). A subset of individuals preferred high-attachment (a subject relative clause), and offline self-report was an accurate means to verify individual attachment preference. No systematic explanation was found to determine individual syntactic attachment preference, however, participants with lower TPD consistently set aside their syntactic preference in favor of what they deemed to be a more agreeable semantic context. This was interpreted as potentially reflecting pragmatic listening strategies by participants with less aptitude in auditory perception, whose persistent challenge in every conversations may have caused them to favor lexical ‘context updating’ to glean meaning in ambiguous situations over adhering to syntactic preference.

In ERP Experiment 3c (sentences), the same musicians from Experiment 3a (melodies) were tested. Results initially contributed that a P600 reflects syntactic integration of German non-preferred subject- or possessor relative clause attachment in the auditory domain. Metrical context was found to interact with syntactic processing, but in the opposite direction than hypothesized: a P600 elicited by the non-preferred relative clause attachment was reduced in sentences with the irregular metrical context. Since the regular sentences were entrained more than the irregular sentences, it seems that entrainment in this case distracted from the syntactic integration rather than enhancing it. This interpretation is in line with recent results indicating that regular meter can enhance certain perceptual fluency (Obermeier et al., 2016) while simultaneously inhibiting conceptual, or semantic fluency (Menninghaus et al., 2015). The current results potentially extend these previous results to include syntactic comprehension to processes inhibited by entrainment (but see Section 10.1), however; all research that proposes facilitation of speech perception after regular meter is based on work with non-musicians, thus the current findings may be constrained to musician participants. The regular meter may have been over-salient to the current musicians, as musicians generally are reported have superior language perception skills (see Section 4.1; Patel, 2011). There were no individual differences when VWM and TPD were correlated with P600 difference waves.

In sum, each of the three studies evaluated how the metrical context of stimuli (more or less regular) impacted a perceptual process, with respect to either cognitive entrainment (Study 1), neural entrainment (Study 2), or syntactic integration (Study 3). Study 1, conducted with musicians and non-musicians, showed that cognitive entrainment increased when music stimuli was more regular, in both participant groups. Cognitive entrainment remained the same among musicians with more- and less regular language stimuli, but non-musicians entrained more when language stimuli was more regular. Study 2 showed that musicians’ neural entrainment increased with metrical regularity in both domains. Study

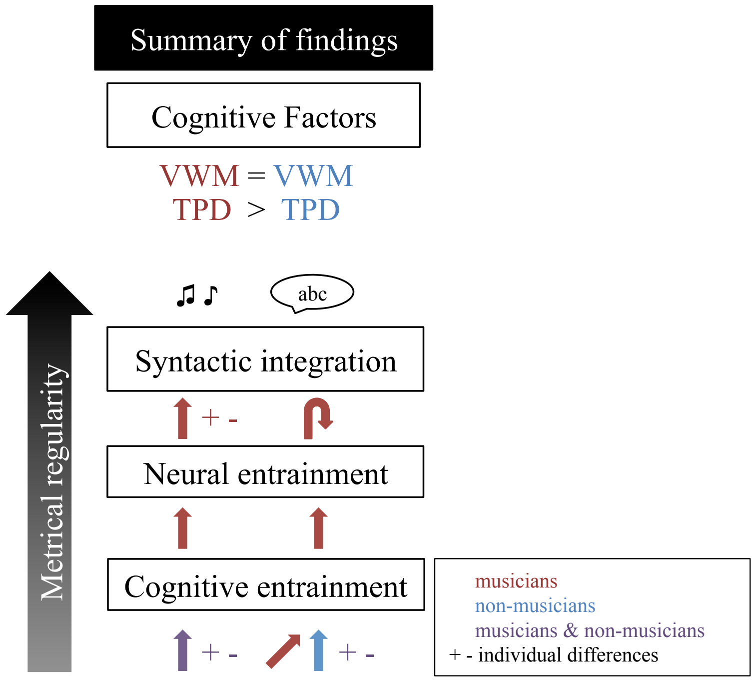



*Figure 9.1: Summary of dissertation results. VWM = Verbal Working Memory. TPD = Time and Pitch Discrimination. The large black arrow creates a context for the results of each type of process studied, on a continuum from stimuli being less- to more metrically regular. Vertical colored arrows indicate that as metrical regularity increased, experimental performance increased. The diagonal arrow indicates no impact of metrical regularity on performance. The looped arrow indicates that after stimuli became 'too regular,' performance decreased.*

3 showed that musicians' music syntactic integration was facilitated as metrical regularity increased, but that their language syntactic integration was hindered as metrical regularity increased. The findings (summarized in Figure 9.1) are discussed in the remainder of this chapter.

## 9.2 Neurocognitive entrainment

As it was used here, the concept of neurocognitive entrainment merged cognitive (attention) and neural oscillatory models of entrainment. Entrainment referred to the general phase coupling of two oscillators, in this context referring to an "internal attentional oscillator" proposed by the Dynamic Attending Theory (DAT; e.g., Large & Jones, 1999) or neurons in the auditory pathway that couple to an external auditory pulse (e.g., Large & Snyder, 2009; Schroeder & Lakatos, 2009). It was argued that these models can not only describe perception of musical meter (e.g. Large & Palmer, 2002; 'neural resonance theory,' Large, 2008), but also can describe and be combined with reports or models of perception modulated by temporal structure in speech (Giraud & Poeppel, 2012; Pitt & Samuel, 1990).

### 9.2.1 Was meter comparable across domains?

The current auditory analysis (Experiment 2a) showed that melodies and sentences with intentional, analogous metrical composition and execution indeed created a comparable frequency profile, i.e. frequencies representing upper and lower levels in a metrical hierarchy appeared across domains in comparable frequency ranges. This shows that meter is empirically malleable. Metrical structure is arbitrary and can be imposed in any manner, its physical manifestation residing in the composition and temporal execution of metrical constituents (with voice or instrument). Thus in future empirical studies, the conversation in the field of music–language research can shift from intrinsic properties in music and language (e.g. Arvaniti, 2009; Patel, Iversen, Wassenaar, & Hagoort, 2008) to an awareness of timing and stress parameters that are typically unaccounted for. That said, actual acoustic content is distinct from how that content is perceived, which is addressed in the next sections.

### 9.2.2 Did participants neurocognitively entrain to meter, similarly across domains?

#### Neural entrainment

Neural entrainment was effectively shown in an EEG frequency experiment, with musician participants (Experiment 2b). In accordance with the hypothesis, it was shown that regular meter conditions were entrained more than irregular meter conditions, consistently in both domains. In music, this finding indicated that the brain imposed top-down processes on the neural entrainment, since the stimulus analysis showed irregular melodies to have larger amplitude than regular melodies at the higher metrical level. This finding of top-down regulation of neural entrainment is consistent with previous EEG music-domain literature (Nozaradan et al., 2011, 2012). In language, the greater entrainment to regular meter was a novel finding, as hierarchical metrical perception in language has never been approached in previous EEG literature. Here it was shown that sentences with a highly structured meter, akin to oratory or poetry, can entrain the musician brain. At least at higher metrical levels, this entrainment was comparable across domains. Thus theories asserting that predictable temporal structure entrains the brain, such as neural resonance (e.g., Large & Snyder, 2009) or the neural oscillatory hypothesis (Schroeder & Lakatos, 2009) are shown to extend beyond the music domain. If language is given a hierarchically structured meter, the upper-level meter perception (group level) is similar to musical upper-level meter perception.

The same cannot be said of the lower-level meter perception: the individual level was entrained in music but not in language. In music, the entrainment pattern matched the acoustic prominence of the stimuli (large individual peaks in regular and irregular melody spectra), and is in line with previous literature (Nozaradan et al., 2011, 2012). In language, the brain down-regulated entrainment to individual beats (representing the 'beat' of individual syllables) compared to what peaks were found in the stimuli spectra. The language finding would seem to be part of a larger linguistic segmenting phenomenon hinted at by previous literature (Buiatti et al., 2009; Ding et al., 2016) that metrical organization can create boundaries in the sentence structure that are more salient than syllable boundaries. In other words, the brain down-regulates metrical information coming from the syllable once reliable upper-level information is present.

When it comes to neural entrainment, EEG frequency tagging showed that structural regularity was more salient in upper metrical levels regardless of the domain, while lower metrical beats were less salient in language compared to music. Thus nested beat structure was neurally entrained differently in music and language in these musician participants. Top-down processes seemed to be intervening, up- and down-regulating entrainment to metrical structure that was dependent on the metrical level and the domain. This may be evidence of discrete neural resources in music and language that underlie temporal processing or metrical grouping (Abrams et al., 2011; Schmidt-Kassow et al., 2011; see Section 2.5). However, before specificity of neural resources per domain can be inferred, future investigation with non-musician participants must first clarify whether this cross-domain difference is general or limited to musical experts (see Section 10.1).

### Cognitive entrainment

Cognitive entrainment was investigated in a behavioral experiment with musician and non-musician participants (Experiment 1b), where critical notes or syllables were detected in a reaction time paradigm. In music, performance was improved in regular compared to irregular melodies for both musicians and non-musicians. In language, performance was improved in regular compared to irregular melodies in non-musicians only; musicians showed no effect of metricality in language.

Results were consistent with cognitive entrainment occurring in the music domain as would be predicted by the DAT (e.g., Jones & Boltz, 1989)—participants, regardless of musical training, had improved performance in regular as opposed to irregular metrical contexts. This is in line with previous behavioral studies utilizing auditory detection tasks (e.g., Jones et al., 2002) that indicated cognitive entrainment to meter.

Results were only partially consistent with the cognitive entrainment that should have occurred in more metrically predictable sentences. Consistent with non-musician data, the DAT, though seldom tested in the language domain, would predict results that match the music domain results; moreover, the attentional bounce hypothesis would also predict improved phoneme perception in more temporally and metrically predictable contexts (Pitt & Samuel, 1990). Thus while non-musicians seemed to cognitively entrain more to regular than irregular metrical contexts regardless of domain, the musicians only exhibited the predicted behavior in music. In language, musicians either cognitively entrained to both language conditions, or to neither, or they had separate strategies in each metricality condition that were equally successful in terms of behavioral performance. The literature, discussed in the next paragraph, points to musicians entraining to both metricality conditions in language.

The OPERA hypothesis (Patel, 2011), introduced in Section 8.5, predicts that musicians have an asymmetrical processing in music compared to language. Specifically, due to a combination of training and experience in their musical education, musicians hone perception of certain acoustic features that allow ‘normal’ music perception but imply an amplification of said acoustic features in the language domain. This is corroborated by ample individual-differences evidence that shows musicians to have enhanced syllable perception compared to non-musicians (e.g., Marie et al., 2011; Parbery-Clark, Tierney, Strait, & Kraus, 2012; see Section 4.1). Applied to the current results, enhanced perception of amplitude envelope could give them expected results in music but enhanced results in language—such as cognitive entrainment to the irregular sentences, whose meter was indeed more structured than everyday speech. This argument is also invoked to explain the meter-syntax interaction in language, below in Section 9.4.1.

#### Neurocognitive entrainment

In musicians, neural and cognitive accounts seem to align in music perception and form a coherent neurocognitive account of entrainment, consistent with the account provided here: in separate experiments, musicians neurally and cognitively entrained more to regular compared to irregular melodies. Thus attention was entrained proportionally to metrical regularity according to the mechanism proposed by the DAT (e.g., Jones & Boltz, 1989; Large & Palmer, 2002), and this was reflected in greater neural resonance (e.g., Large & Snyder, 2009) in the more entrained condition.

Musicians’ neural and the cognitive entrainment in language, however, seem to form a more complex relationship than in music. It would seem that while neural resonance indicated greater entrainment in regular metricality conditions, the advantage did not translate to cognition, where there seemed to be no difference in entrainment dependent on metrical context. Thus the DAT (e.g., Large & Jones, 1999) and attentional bounce (Pitt & Samuel, 1990) theories were not supported by musician language data, nor was the current hypothesis that these theories overlap with neural resonance accounts.

However, non-musicians did perform better when language stimuli had the regular meter context. This indicated cognitive entrainment to regular language meter, indeed in support of DAT (e.g., Large & Jones, 1999) and attentional bounce (Pitt & Samuel, 1990) theories. Thus, only with additional neural-entrainment data from non-musicians can a generalized account of neurocognitive entrainment in language be made. Non-musician neural entrainment data must be compared with the current musicians' results, and then additional inferences can be drawn in combination with the current non-musician cognitive entrainment behavioral data. Otherwise, effects of musicianship can be confounded with general perceptual mechanisms when forming a neurocognitive entrainment account in language (see the outlook in Section 10.1).

These sections have discussed neurocognitive entrainment to metrical context, but this was only part of the investigation. The full neurocognitive entrainment hypothesis that served as a framework for the dissertation includes how neurocognitive entrainment facilitates syntactic integration in both domains. The next sections add syntax to the discussion.

## 9.3 Syntax

A theoretical basis for syntax existing respectively in music and language was provided in Section 3.1. The domains do not have the same syntax, but rather the similarity lies in the existence of a system of rules governing the constituents (e.g., Chomsky, 1957; Lerdahl & Jackendoff, 1983) and in the cognitive procedures (e.g., prediction, detection, integration) that draw on these rules. Considering that syntactic integration has been argued to be a cross-domain phenomenon, i.e. that a shared neural resource integrates new syntactic information into the separately stored, discrete systems of music and language syntactic rules (SSIRH, Patel, 2003), the current paradigm addressed the integration of non-preferred syntactic structure across domains, comparing melodies with alternate relative key endings to sentences with alternate relative clause endings.[1]

[1] In line with the distinction between actual syntax and the existence of syntactic principles, no claim is made here that the syntactic forms of relative key and relative clause are similar, but rather that the integration of the non-preferred (but not incorrect) structure is analogous and capable of bypassing error-detection mechanisms. In other words, while current stimuli's cross-domain syntactic structures were similar in that both alternatives were plausible (and in the temporal unfolding of the constituents), the syntactic manipulations themselves were qualitatively different in music and language, i.e., relative keys and relative clauses do not share any objective syntactic features, aside from plausibility in their alternate forms.

### 9.3.1 Is syntactic integration comparable in music and language?

In the context of the SSIRH,[2] the P600 represents a syntactic resource shared across music and language domains (Patel, 2003). Specifically, the P600 is suggested to index integration of syntactic constituents with their preceding context (e.g. Patel et al., 1998). Thus the presence of a P600 effect in both Experiments 3a and 3c was interpreted as reflecting syntactic integration of non-preferred syntax across domains. The P600 had a comparable latency but different distribution across domains, however; both distributions were within the scope of previous P600 reports from literature (e.g., Osterhout & Holcomb, 1992; Patel et al., 1998).

The integration of non-preferred, as opposed to violated syntax, was shown for the first time here in the music domain, and replicated previous language domain findings (e.g., Kaan & Swaab, 2003). Musicians were necessarily tested, especially as the task required identification and judgment of a subtle syntactic manipulation in music. It is not known whether the current melodies would elicit a P600 in non-musicians, who do not have as theoretically developed music syntax representations as musicians (see Besson & Faïta, 1995; Koelsch et al., 1999).

Current music and language stimuli contained non-preferred syntactic structures because a limitation in previous ERP paradigms was that syntactic integration was indistinguishable from syntactic error detection (e.g., Patel et al., 1998; see Section 3.1.3). The alternate structures (major vs. minor relative keys, high- vs. low relative clause attachment) were error-free while still requiring integration of syntactic constituents, and the non-preferred structure in turn evoked clear P600 ERP effects in both domains. Thus the current results extended music-language syntax comparisons firmly beyond the scope of error detection.

Whether or not the similar P600s indicated shared syntactic integration resources, as opposed to shared processes beyond syntactic integration, remains to be investigated. For example, Slevc and Okada (2015) recently attributed experimental cross-domain syntax results to domain-general cognitive control. To be accurate, the results in this dissertation are consistent with SSIRH claims *in addition to* being consistent with claims of competing theories, such as Slevc and Okada (2015). (Coming closer to) proving the same integration resource must be attempted by future studies, using for example an interaction paradigm and more extensive functional/anatomical study.

The N500 elicited in melodies has in previous studies been attributed to syntactic pro-

[2]The note and syllable discrimination task in the first study (Experiment 1b) is not included in the cross-domain syntax comparison addressed in the research question. Discrimination of features such as pitch is not related per se to integration (in fact, it was specifically mentioned in the original SSIRH publication to be outside the scope of syntactic integration; Patel 2003), and moreover the discrimination in Study 1 was only syntactic insofar as the critical notes and syllables also determined the final syntactic structure of melodies and sentences. Thus results of Experiment 1b are generalizable to music and language cognition (and the respective effects of entrainment, as addressed above), but not applicable to the SSIRH.

cessing (Koelsch et al., 2000), semantic processing (Steinbeis & Koelsch, 2008), or general holistic listening strategies where non-musicians had the impression of an 'unfinished' harmonic manipulation (Featherstone et al., 2013). Thus it is possible that the N500 in the current set-up was either part of the syntactic process that gave rise to the P600, or was its own discrete musical comprehension process. However, the behavior of the N500 in the metricality–syntax interaction suggests that it is neither strictly syntactic nor semantic as comparable to the linguistic sense, but rather an assessment of the harmony enabled by a regular-meter context and inhibited by irregular-meter context. The N500 is further discussed below (Section 9.4.1) and an outlook is provided in the next chapter (Section 10.2).

Overall, syntactic integration was shown to indeed exist comparably across domains. This was indexed by a P600, in a circumstance that did not require detection of syntactic error. Next, ERPs will be further discussed in the context of the metricality manipulation.

## 9.4 Neurocognitive entrainment and syntax

Finding a comparable cross-domain meter-syntax interaction, or more specifically, establishing whether neurocognitive entrainment to regular metricality can facilitate syntactic integration across domains, was a primary goal of the current work. The chosen paradigm therefore was carefully designed to induce, across domains, more entrainment in the regular compared to irregular metricality conditions, and successfully showed this in Study 2. Since meter and syntax constituent boundaries overlap in both music and language (Lerdahl & Jackendoff, 1983; Selkirk, 2011), and since entrainment to syntax constituents has been shown in previous EEG studies (Ding et al., 2016; Schmidt-Kassow & Kotz, 2008) it was hoped that consequences of this entrainment would be evident in the ERPs elicited by the syntax manipulation in Study 3 (Experiments 3a & 3c). The expected consequence of entrainment to regular meter was facilitated syntactic integration (Roncaglia-Denissen et al., 2013), in accordance with increased attention (Large & Jones, 1999; Pitt & Samuel, 1990), and more efficient neural processing downstream of more synchronous neural communication (Thaut et al., 2005; Varela et al., 2001).

### 9.4.1 Does entrainment to metrical context facilitate syntactic integration comparably across domains?

More entrainment to regular metricality conditions did not facilitate syntactic integration similarly across domains. The direction of the interaction was as hypothesized in the music domain (Experiment 3a), but ERP data showed the opposite metricality-syntax interaction in the language domain (Experiment 3c). The meaning of 'facilitation,' and current findings' discrepancy across domains, are discussed in the next sections.

**Music**

Entrainment to regular meter seemed to facilitate syntactic integration in music: The music P600 was reduced in the regular metricality condition, and individual differences suggested that processing was more similar among participants when they had the opportunity to entrain to melodies with more (temporally) predictable weighted notes. This is exactly as predicted, and supports the theoretical framework that led to the following hypotheses: The DAT predicts more entrainment to regular meter. Syntactic integration should in turn be facilitated by entrainment as the attentional pulse is sharper and more accurate in the entrained conditions, and neural synchrony is increased (Large & Jones, 1999; Thaut et al., 2005, see Chapter 2). The fact that previous non-musician, language ERP data has shown similar reduction of ERPs in regular-meter conditions (Cason & Schön, 2012; Roncaglia-Denissen et al., 2013; Rothermich et al., 2012) seemed initially to extend the concept of cross-domain shared resources to include a shared facilitative effect by regular meter (but was negated by current language results).

Alternatively, the impact of regular metricality in music is perhaps not exactly facilitative, but rather it induces a qualitatively different perception of the harmony. The main support for this view is that the non-hypothesized N500 was increased in amplitude in the regular metricality condition; whatever process underlies the N500 would have to be facilitated by the irregular metricality if one follows the logic that a reduced amplitude indicates facilitated processing (Cason & Schön, 2012; Roncaglia-Denissen et al., 2013; Rothermich et al., 2012). This does not seem likely, considering that irregular metricality in Western Tonal music is unusual and should increase processing difficulty. Moreover, a previous study that reviewed the common musical circumstances eliciting N500 and P600 pointed out that ‘unfinished-sounding’ harmony, much like the current minor-ending melodies, typically elicited an N500 in music ERP research (Featherstone et al., 2013): all of these cited studies used melodies with a regular metrical context. Thus the neural process eliciting the current N500 may only occur when melodies have a regular meter (and perhaps the P600 in response to the current harmonic manipulation may only occur in irregular melodies).

These interpretations are not mutually exclusive. It is possible that, especially considering the validity of all arguments leading to the facilitated syntactic integration by regular metrical context, the regular metricality condition enabled an additional value or meaning assessment of the melodies that surfaced as an N500. Regular meter is most common in Western Tonal music, thus the familiar cultural value assigned to the major and minor modes (happy vs. sad, respectively; Hevner, 1935) could have been triggered as part of the familiar attending mode used for Western Tonal music (e.g. Jones & Boltz, 1989). Then, in the absence of the familiar metrical context, this value-assessment was dispensed with. Participants were engaged in their syntactic integration task, increased in difficulty in the irregular metrical context. This possible, but speculative, reconciliation of the ERP results can only be elucidated by future empirical work. For example, the paradigms that previously elicited N500 effects could be reproduced, with the added variable of metrical regularity, to see how the N500 reacts to changes in meter (see Section 10.2).

**Language**

Irregular metricality seemed to facilitate syntactic integration compared to the regular metricality in language: the P600 was reduced in the *irregular* metricality condition. This interaction was opposite from the hypothesis, which stated that the *regular* metricality should facilitate syntactic integration, as it did in the current music ERP experiment and in a previous language ERP experiment with non-musicians (Roncaglia-Denissen et al., 2013). Since musician participants entrained more to the regular sentences (Experiment 2b), the most logical interpretation of the data is that more entrainment hindered rather than helped syntactic integration. A similarly complex entrainment narrative has recently surfaced in the field of empirical aesthetics: regular meter in language facilitated perception (Obermeier et al., 2016), while hindering the comprehension of meaning (Menninghaus et al., 2015). Thus what was found here might be part of a larger phenomenon that entrainment produces disparate effects along different dimensions of language cognition. Whereas this was found previously in relation to semantics, here it was the case with syntax.

Specific perceptual phenomena in musicians could also account for the reversed direction of the metricality x syntax interaction. The theoretical framework of the SSIRH can be combined with proposals from the OPERA hypothesis (Patel, 2011, see Section 8.5) to suggest that musicians' qualitatively different metricality–syntax interaction across domains is the result of a perceptual hypersensitivity to acoustic features of language. A cross-domain syntactic-integration neural resource could interact with metrical information in a manner that is respectively skewed across music and language. The same resource is operating with two qualitatively different sets of relevant information and the outcome is therefore different across domains.

To follow up these interpretations, non-musicians need to be tested (see Section 10.1). If non-musicians show the same interaction direction as current musicians, results could be seen to trace a larger phenomenon that entrainment can facilitate some while inhibiting other aspects of language cognition. If, however, regular language meter facilitates syntactic integration in non-musicians, then the current pattern of results could be seen as specific to musical experts, supporting more the SSIRH/OPERA framework.

Regardless of which degree of regularity facilitates what process across domains, the current findings prove meter to be linked to processing of non-preferred syntax, in both music and language. This encourages future work in the area of syntax processing to acknowledge and explore entrainment to meter. Moreover, discussing the impact of metrical context on syntax processing across domains also requires first acknowledging the role of (cognitive factors trained by) musical expertise. The next section entails a full discussion of the individual-differences aspect of the current paradigm.

## 9.5 Individual differences

This dissertation focused on musical expertise, temporal processing, and working memory as three measurable attributes to monitor among participants. All three aspects were shown in the introduction to impact the initial perception and/or integration of meter and syntax (Chapter 4): Musical expertise is associated with improved pre-attentive auditory processing and more robust mental representations of meter and syntax (Fitzroy & Sanders, 2013; Geiser et al., 2010; Kraus & Chandrasekaran, 2010), temporal perception both influences auditory event perception at the sub-second order (e.g. metrical and syntactic events) and is influenced by large-scale temporal organization of events (Jones et al., 2002; Magne et al., 2007), and working memory has strong links to language syntax processing (Daneman & Carpenter, 1980). The current paradigm therefore condensed diagnostic measures of musical expertise, temporal perception, and working memory into two factor scores in order to monitor individual differences among all participants across tasks.

### 9.5.1 Cognitive factors

The names used for the two factors were based on the cognitive functions associated with their respective auditory tasks (described in Section 5.2). Factor Verbal Working Memory (VWM) represents combined performance scores from forward and backward digit span, non-word repetition, and modified listening span tasks (Daneman & Carpenter, 1980; Mottier, 1951; Tewes, 1994), and factor Time and Pitch Discrimination (TPD) represents combined performance scores from anisochrony detection, melody discrimination, and rhythm discriminations tasks (Dalla Bella et al., 2016; Wallentin et al., 2010). All of the tasks in VWM were delivered via a spoken medium, that is, the materials were all linguistic sounds delivered by a speaker (hence the *Verbal* Working Memory nomenclature, as these materials should be involved with the phonological loop; Baddeley & Hitch, 1974). All of the tasks in TPD were without spoken delivery of materials, in other words, all participants heard were tones or clicks throughout tasks that targeted temporal and pitch discrimination.

It is important to acknowledge that the name TPD, while indeed reflecting the aims of underlying tasks, may not comprehensively include all resources involved in these tasks. Specifically, non-verbal, auditory working memory may be a necessary component to complete the tasks that required participants to store standard melodic phrases, rhythmic phrases, or temporal intervals, and compare them to probe items. If (non-verbal) auditory working memory is considered a component of factor TPD, a vein of previous research falls into place with the current results: many studies have linked auditory working memory to both musical expertise and improved perception of linguistic events at a sub-second time scale (Kraus et al., 2012; see Chapter 4). Thus if auditory working memory is considered a component of TPD, the fact that TPD was improved in musicians and correlated with note and syllable discrimination measures supports previous findings. Further testing would be necessary to assess this claim, for example, a task clearly delineated to target non-verbal auditory working memory may load onto TPD in a factor analysis together with the current seven diagnostic tests.

The use of factor score correlations across these studies helped to address a limitation in several previous musician–non-musician paradigms, namely that grouping according to musical expertise potentially masks individual difference in nuanced cognitive ability. In particular, the fact that musicians are typically associated with better temporal processing and working memory (Ho et al., 2003; Kraus et al., 2009) means that improved perception of meter and syntax among musicians (e.g., Fitzroy & Sanders, 2013; Geiser et al., 2010) is indistinguishably traced back to musicianship and respectively temporal perception or working memory. This limitation was effectively addressed here: evidence collected across the three studies suggested that individual differences in perceptual processing were attributable to cognitive resources distinctly from expertise, but that, in an important distinction, musical training was so influential on certain resources (like TPD) that group performance differences emerged. Thus, especially where perception of auditory signals is concerned (for tasks not requiring express knowledge such as musical theory), nuanced individual-difference data collection is more informative than musician/non-musician grouping. In the paradigms mentioned above, then, group differences between musicians and non-musicians are potentially due to cognitive resources associated with the TPD factor but not necessarily differences associated with VWM.

### 9.5.2 How do individual differences affect entrainment to meter and syntax processing?

Regarding the specific role that the cognitive factors played in the current investigation, the story that emerges from the findings across studies is that TPD is involved in both the detection and the integration of syntactic constituents (pitches and phonemes), while VWM is primarily involved in syntactic integration. TPD and VWM are moreover both likely involved in interactions between meter and syntax.

In the two ‘detection’ places in this series of studies, the note and syllable discrimination (Experiment 1b, which measured immediate application of detection ability in accuracy and

RT), and the adherence to syntactic preference in relative clause attachment (Experiment 3b, seen as an offline, long-term consequence of event detection ability), TPD surfaced as a correlate to behavioral measures. The potential role that TPD played in detection traces back to the attentional oscillator described by the DAT: the detection of events is sharper when the attentional pulse is peaking, and peaks in the attentional pulse are more closely coupled with salient information when the error-correction mechanism of the oscillator is most agile (in the case of TPD, captured by the anisochrony detection scores).

When syntactic integration was tested in Study 3 (Experiment 3a, melodies with preferred and non-preferred alternate syntactic structures), both cognitive factors played a role: VWM and TPD correlated with the P600 amplitude elicited by non-preferred melodies, signaling their involvement in musicians' musical syntactic integration. (Verbal) working memory is intuitively involved in this later-stage integration process, as the storage and retrieval of information in the working memory system presumably has to take place after the initial perception of events. Thus Study 3 provided a direct link between verbal working memory and music syntax processing, adding to previous research in the language domain (e.g., Bornkessel et al., 2004). The role of Time and Pitch Discrimination in music syntactic integration replicates previous findings that timing is crucial to harmonic judgment (Bigand et al., 1999; Boltz, 1993), and is attributable to large-scale incorporation of pitch events and their temporal placement at the phrasal level, i.e. the evaluation of event-timing, and which temporal deviances are salient to the meter of the melody as opposed to being natural temporal variation in the pianist's production (e.g. Large & Palmer, 2002).

Verbal Working Memory and TPD were also both likely involved in the metricality x syntax interaction in Study 3 (Experiment 3a). The correlations with the music P600 amplitude occurred only in the condition where the musicians were presumably not entraining (or entraining less) to higher metrical levels, the irregular metricality condition. This result was interpreted as an equalization of individual differences in TPD and VWM when participants entrained, i.e. the entrainment provided a scaffold for cognitive resources, helping the attentional oscillator to better time the pulse to salient events, which lessened the group spread in cognitive performance. Upon further speculation, the cognitive factors may be actively involved in this entrainment process. Participants with higher VWM were able to take better cognitive advantage of the entrained attention in the regular metricality condition (the higher VWM score, the greater reduction in P600 amplitude in regular compared to irregular metricality), which could mean that participants with better VWM are able to better strategize and consolidate cognitive resources, given the opportunity. This is plausible especially when considering evidence in the literature that (verbal) working memory is tied to general intelligence (Conway, Kane, & Engle, 2003). Accepting the arguments presented above for TPD representing part of the error-correction mechanism in the DAT's internal oscillator, this would place TPD as (circularly) both agent and receiver in the entrainment process: the ability to entrain is constrained by the initial detection and error correction of temporally predictable events, so the utilization of regular metricality to entrain first requires TPD. Then, once entrained, the aspect of this resource that compiles larger scale temporal (and pitch) patterns can more efficiently assist later-stage cognitive processes.

Experiment 3b, looking at sentence completion, revealed an unforeseen role of TPD. Participants with higher TPD scores completed an ambiguous sentence according to their syntactic preference, while participants with lower TPD scores relied on the semantic content of the sentences to resolve the ambiguity, often contrary to their syntactic preference. Findings were interpreted to mean that participants with higher TPD could 'afford' to adhere to their syntactic preference, and it was speculated that participants with lower TPD may have developed a strategy in everyday communication to boost their comprehension with multiple semantic cues.

To make one final point, the roles of VWM and TPD in the current tasks suggest a broadened scope of individual measures for future investigations of individual differences in auditory language comprehension: much (language) comprehension research focuses on verbal working-memory related aspects, which here were shown to be less involved than temporal and pitch discrimination (and any related non-verbal working memory). Thus a contribution of the current research is to suggest for future auditory individual-differences paradigms a shifted focus from single verbal working memory measures to multiple auditory perception measures, both verbal and non-verbal (temporal, pitch, and working memory).

In sum, the individual-differences approach showed musical training to be beneficial to information processing in complex auditory signals, and the resource that appears to be most directly trained by music practice is that underlying the cognitive factor TPD (temporal, pitch and rhythmic discriminations tasks). The cognitive factors helped to refine this paradigm, and combined evidence across experiments offered a coherent account of their respective roles in meter and syntax processing. Suggestions for future individual difference investigations are suggested in the next chapter (Section 10.3)

## 9.6 Closing remarks

The first chapter of this dissertation outlined a hypothesized explanation for related phenomena found in the literature, all pertaining to meter and syntax in music and language. The current approach constructed a single underlying framework for the previous findings: metrical regularity promotes neurocognitive entrainment, and increased neurocognitive entrainment in turn facilitates syntactic integration, similarly across domains and with individual differences. Multiple cognitive, neural and formal theories comprised the framework.

Musicians followed predictions in the music domain only. Across three studies, measures of cognitive entrainment, neural entrainment, and syntactic integration all consistently improved when the metricality was regular as opposed to irregular, in the music domain. Musician performance did not follow predictions in the language domain. Only regarding neural entrainment (Study 2) did musicians follow the predicted enhanced entrainment to regular meter; otherwise, no difference in meter was found in cognitive entrainment (Study 1), and the reverse of the hypothesized effect was found with syntactic integration (Study 3).

The fact that the language results were different does not diminish the significance of the music domain findings in this dissertation. Never before has neural entrainment been linked to cognitive entrainment with the same stimuli set, and never before has neural entrainment been linked with such advanced cognitive processes as syntactic integration. The music domain evidence shown here supports the theoretical framework laid out in the introduction. The cognitive attention described by the DAT (e.g., Jones, 1976) was reflected in Study 1, and the accompanying entrained delta-theta range neural oscillations (Henry & Hermann, 2014; Schroeder & Lakatos, 2009) and neural embodiment of meter perception (Large, 2008) was reflected in Study 2. Moreover, the suggestion that entrainment to meter could facilitate syntax processing was realized in Study 3 – the introduction attributed this to overlapping meter and syntax constituent boundaries (Lerdahl & Jackendoff, 1983), as well as general increased neural synchronization in entrained conditions (Singer, 2013; Thaut et al., 2005). Thus in music, the framework suggesting regularity in metrical structure to promote neurocognitive entrainment, which in turn could facilitate syntactic integration, proved consistent with the findings. The mutual overlap of the concatenated theories across the three studies is a groundbreaking step in cognitive science: cognitive attending and neural entrainment and syntactic integration can all interlace in music, with musician participants.

The current findings suggest that something about musicians as a group impacted the cognitive aspect of their entrainment in the language domain, or at least the cognitive consequences of neural entrainment. In line with a previous behavioral study (Menninghaus et al., 2015), different dimensions of language cognition were impacted disparately when encountering regular meter. In Menninghaus et al. (2015), the perception of stimuli was improved when meter was regular, including the overall aesthetic pleasantness, but comprehension of semantic meaning was hindered by regular meter. Non-musicians must be tested to see whether, for example, neural entrainment to language meter detriments comprehension-related processes overall, or whether this is limited to musicians in the language domain. Previous language-domain EEG data, all with non-musician participants, suggest that regular meter does facilitate syntactic (Roncaglia-Denissen et al., 2013), semantic (Rothermich et al., 2012), and phonological processes (Cason & Schön, 2012). Indeed the cognitive entrainment in Study 1 did seem to follow predictions among non-musicians in language; non-musician neural data must be collected with the current paradigm to get the full story. Thus there is the potential that among non-musicians, the overlapping language-domain theories proposed in the introduction—the attentional bounce hypothesis (Pitt & Samuel, 1990), the Giraud and Poeppel (2012) speech perception model—would manifest as the current approach predicted.

## 9.7 Conclusion

This dissertation aimed to investigate meter perception, syntactic integration, and the impact of meter on syntactic integration, in both music and language domains. Behavioral and EEG methods were used in an individual-differences approach. Among musicians, top-down regulation of neural entrainment was suggested when regular meter was neurally entrained more than irregular meter across domains. Metrical regularity interacted with syntactic integration in musicians, but musicians' neural entrainment to meter had a qualitatively different impact on syntactic integration across domains. This implied either asymmetrical music and language cognition among musicians, or a complex impact of entrainment on language cognition. Non-musician results suggested a cognitive benefit from metrical regularity in both domains, warranting further investigation as to the differences between musicians and non-musicians in their affinity to entrain to language, and how this impacts syntactic comprehension. The individual-differences approach revealed the emerging importance of temporal and pitch discrimination in meter perception and syntactic comprehension across domains, in that it seemed to influence perception more globally than verbal working memory. These findings encourage future work in cross-domain meter perception and syntactic comprehension and espouse the consideration of individual differences in auditory perception paradigms.

# Chapter 10

# Outlook

Based on behavioral and EEG data, this dissertation was able to draw conclusions about musicians' neurocognitive entrainment to musical meter, and the subsequent impact on music syntactic comprehension. Moreover, the individual difference metrics proved informative, showing that Time and Pitch Discrimination played a role in cognitive entrainment and syntactic comprehension. Several issues, however, remained unclear. The role that neurocognitive entrainment played in language syntactic comprehension was not straightforward. The consequence of neurocognitive entrainment on additional aspects of music cognition was also unclear. The current chapter suggests future lines of research to address these latter two issues, in addition to potential further applications of the individual-differences metrics.

## 10.1 Neurocognitive entrainment and syntactic comprehension in language: Non-musicians

The literature informing the current approach pointed toward regular meter facilitating linguistic processes related to comprehension (e.g., Cason & Schön, 2012; Roncaglia-Denissen et al., 2013; Rothermich et al., 2012), which did not bear out in the current EEG findings with musician participants. Crucially, all of this prior research was conducted with non-musician participants. Applying the current EEG paradigm to non-musicians is the next logical step in extending the current investigation.

Current behavioral results did point toward non-musicians benefiting from the regular meter in language: non-musicians' syllable discrimination performance was improved in regular compared to irregular metricality conditions (Experiment 1b), consistent with previous entrainment paradigms (e.g., Jones et al., 2002). This convergence of non-musician reports from the literature and the current non-musician behavioral results suggests that non-musicians may neurocognitively entrain to regular language meter, and that unlike current musicians, the entrainment may facilitate their syntactic integration in language.

## Hypotheses

Applying the current language-domain EEG paradigm to non-musicians, the accompanying hypothesis would be that syntactic integration of the relative clauses would indeed be facilitated by regular metricality, resulting in a reduced P600 in regular as compared to irregular sentences. To show accompanying neural entrainment, a frequency analysis of their EEG data should reveal peaks representing the weighted syllables (the 'group' metrical level) in the regular- more than in the irregular condition, reflected by a peak at the 2:1 harmonic of the group level (Tierney & Kraus, 2015; Experiment 2b). A large participant group may also allow to test the spectrum near 1.7 Hz, the actual group frequency of the weighted syllable, by increasing the spectral power.

## Implications

### Delta-theta oscillations

The use of non-musician participants in an EEG frequency experiment would clarify whether the cross-domain differences found in neural entrainment to nested beat levels (Experiment 2b) were due to musical expertise or whether this reflected a general cognitive phenomenon. If the individual metrical level (syllable) is not entrained whereas the group-level (weighted syllable) is entrained (replicating current musician results), the argument would be strengthened that top-down processes downregulate language-domain neural entrainment to the syllable. If on the other hand the individual-level is entrained in addition to the group-level, the downregulation of individual-level in current musician results may be attributed again to differences due to expertise, and the simultaneous entrainment to multiple nested beat levels can indeed occur in the language domain similarly to music. Thus non-musicians' results could inform models of cross-domain neural entrainment that account for top-down regulation, for example extensions of neural resonance theory (Large, 2008; Large & Snyder, 2009) to the language domain.

### ERPs

If non-musicians' language data replicate current musicians' language ERP data (entrainment to regular meter hinders syntactic integration, Experiment 3c), first of all, an investigation would need to be launched into the particular syntactic manipulation used in the current sentences, since several other linguistic stimuli sets showed non-musicians' linguistic processing to be facilitated by entrainment (e.g., Cason & Schön, 2012; Roncaglia-Denissen et al., 2013; Rothermich et al., 2012). The syntactic preference used here is based on individual coding; in current filler sentences, the same syntactic ambiguity occurred as in Roncaglia-Denissen et al. (2013) and is consistent across native German speakers. Thus the regular-meter-hindrance should first be replicated in filler sentences of the present study (Section 5.3.2 provides an example), with multiple musicians and non-musician groups, to be sure that the phenomenon is a consistent one.

If regular-meter hindrance in language does turn out to be consistently replicated, then assumptions made by the current neurocognitive-entrainment framework (see Section 2.3) need to be revisited. For example, since the critical syntactic morpheme occurs on a weak metrical beat in the current paradigm, perhaps in the language domain 'analytic attending' (see Section 2.1.1), or dynamic attending based on local relationships, is a cognitively more-efficient strategy than future-oriented attending (referred to throughout this dissertation as 'cognitive entrainment'), or dynamic attending based on expected temporal locations. Thus syntactic integration could be facilitated by the better attending strategy that is specific to the materials; accordingly, if the critical morpheme were located on a strong beat, then the future-oriented attending should be more beneficial. In this case a regular meter, which would induce the future-oriented attending, should facilitate syntactic integration.

Depending on whether non-musician participants neurally entrained to the meter (in a replication of Experiment 2b), this might also introduce a dissociation between cognitive entrainment and neural entrainment, e.g., more synchronized neural communication based on the neural entrainment could occur independently from whether attention to auditory events is more accurate. The current paradigm assumes that the two operate in conjunction, but, as shown in the results summary in the previous chapter (Figure 9.1), neural entrainment and cognitive entrainment already dissociated among (albeit different groups of) musician participants in the language domain. Linking this to another area of the literature where a similar dissociation has been discussed, for example neural entrainment could be tied more to perceptual fluency and cognitive entrainment tied more to conceptual fluency (Menninghaus et al., 2015; Obermeier et al., 2016). The fact that this dissociation is surfacing means that future research on the topic of 'entrainment' needs to be very exact in defining the processes being empirically investigated.

If non-musicians' language data negate current musicians' language data (and regular meter facilitates syntactic integration), then the current language-domain, counter-hypothesis results (Experiment 3c) can be attributed to differences between musician and non-musician data (as argued in the discussions of Study 3 and Chapter 9). This would also uphold the cross-domain, neurocognitive framework laid in the introduction, but would show the need to qualify the framework to account for differences in expertise. If this is the case, then perhaps the facilitative effect of neural entrainment operates on a sort of gradient, modulated by individual differences such as the current Time and Pitch Discrimination metric. Future prospects for individual differences are revisited in Section 10.3.

## 10.2 N500

The current music EEG results suggest that syntactic integration, indexed by the P600, was facilitated by entrainment to regular meter. Interestingly, a non-hypothesized N500 also surfaced, which showed the opposite behavior from the P600: it was larger in the regular condition and virtually nonexistent in the irregular meter condition. The N500 is typically associated with either musical meaning deriving from music-internal properties (Koelsch, 2011a) or with the integration of harmonic structure (Koelsch et al., 2000, 2013), both explanations which are mutually compatible (see Study 3, Section 8.2.4). The current N500 interaction with meter is an interesting contribution to N500 literature and can inform some of the debate (see Koelsch, 2011b) surrounding the comprehension of musical meaning. The current N500 results can therefore inform future N500 research paradigms in multiple ways, in rather an 'erpology' (Luck, 2005) fashion. Three such paradigms are suggested below.

### Exploratory paradigms

The current results used a relatively new harmonic manipulation in music-syntax EEG paradigms (major or relative minor melodic endings), and one that may likely rely on theoretical music syntactic representations too advanced for non-musicians (see Besson & Faïta, 1995). Therefore, previously executed paradigms that reported an N500 are a fruitful starting-point for systematic exploration of N500 interactions with meter.

1. For example, the five-chord sequence presented by Koelsch et al. (2000) and reused in multiple paradigms, could be adjusted dynamically and temporally to first create a sense of regular meter (e.g., 3/4), and then an alternative irregular meter (e.g., alternating 2/4, 3/4, and 4/4 time signatures within a sequence); the N500 typically evoked by the Neapolitan final cadence in these sequences may or may not interact with the meter conditions.

2. The music stimuli set created by Featherstone, Waterman, and Morrison (2011) is practically a ready-made setup; its original conception was made for harmonic and rhythmic experimental exploration. Moreover, the harmonic incongruities from this stimuli set have previously elicited an N500 from non-musician participants (Featherstone et al., 2013). In a future paradigm exploring the N500's reaction to metrical

manipulation, the stimuli set could be used to combine harmonic and rhythmic incongruities in a 2 x 2 design.

3. The Koelsch et al. (2013) paradigm found an N500 evoked by hierarchically violated musical structure, suggesting the N500 to be linked to processes of recursion in music. The musical sequences were extracted passages based on Bach chorales, written in 4/4 meter. Multiple voices were already used, thus the creation of a sense of irregular meter could easily be achieved.

The above three paradigms would touch upon different harmonic manipulations known to evoke an N500, including local and non-local syntactic dependencies. The usage of both musician and non-musician participants is necessary as well, since differences in the N500 in specific (Featherstone et al., 2013) and meter perception in general (see Section 4.1.1) have been reported across the groups.

### Implications

The above suggestions are all exploratory, thus the results and implications of those results are highly speculative. Nevertheless, three possible patterns of results carry distinct implications (all assuming that musicians and non-musicians were tested, and that an N500 main effect was present in both groups). First, if the current N500-metricality interaction is not replicated, then the source of the current interaction resides with the major-minor harmonic manipulation. If such is the case, then maybe the absence of the N500 in irregular meter melodies lies in external cultural association with the minor mode (see Hevner, 1935), a cultural association that is not made once the meter is no longer typically Western. Second, if the N500-meter interaction occurs only in musicians, then perhaps the source of the interaction is musicians' extra sensitivity to musical meter (Section 4.1.1). Perhaps the cognitive resources allocated to attending the irregular meter disrupt the cognitive processes that underlie the N500, precluding the effect. Non-musicians, in contrast, would not dedicate extra resources to the irregular meter and process the harmonic manipulation as-usual (see Featherstone et al., 2013). Third, if the current N500-meter interaction is present in musicians and non-musicians in multiple harmonic conditions, then perhaps there is a general tie between meter and harmony that creates musical meaning. In other words, the meter contributes to the meaning created by the harmony, and a changed meter strips the meaning from the baseline condition (eliminating the difference between harmonic conditions). Since meter has not been manipulated in N500 paradigms until now, the contribution of meter to meaning has not yet been considered.

These implications are subsequent to exploratory research, and are limited to paradigms set in Western tonal music. Hopefully the coming decades will bring a rich investigation into musical meaning and its ties to meter and rhythm, and broaden the scope to include musical traditions from non-Western cultures.

## 10.3 Individual differences: TPD-centered pedagogical intervention tool

The role of working memory is connected by much literature to intelligence and academic performance (e.g., Au et al., 2015; Conway et al., 2003; Nisbett et al., 2012). While games and training tools are claimed to improve auditory working memory (Anguera et al., 2013; Spencer-Smith & Klingberg, 2015), there has been reception from research community that challenges claims of effectiveness (e.g., Hulme & Melby-Lervåg, 2012), and working-memory-targeted training that improves working memory beyond specific task performance is debated (e.g., Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Redick et al., 2013; Shipstead, Redick, & Engle, 2012). However, music training is linked to improved auditory working memory (Kraus et al., 2012; see Section 4.3.1), thus development of intervention tools that tap into musical tasks may be promising.

Current Time and Pitch Discrimination (TPD) scores, developed through a factor analysis of seven diagnostic tests with over 80 participants (Experiment 1a), may be an effective starting point for the development of an intervention tool for poor language-comprehending students. Time and Pitch Discrimination scores were improved in current musicians compared to non-musicians, and among musicians, improved with weekly practice hours (Experiment 1a). Moreover, current TPD correlated with various behavioral measures of language phonological or syntactic proficiency (Experiments 1b and 3b), and electrophysiological measures of music syntactic integration (Experiment 3a). Thus a (very) long-term perspective of this research would be to develop an effective TPD-centered pedagogic tool aimed at improving language comprehension, for the classroom population on the lower end of the normal range, i.e. those not breezing through everyday situations with full comprehension (perhaps, those who cannot afford to adhere to their syntactic preference, seen in Experiment 3b).

### Research sketch

The immediate steps for future research in this direction would be to replicate the cognitive factor loadings of the diagnostic tests used here (see Study 1) to verify the consistency of TPD. Participants should amount to at least 10 per test battery for best factor analysis results (Field, 2005). Participants should ideally have a varied musical history (including musicians and non-musicians), and musician participants should additionally answer detailed questionnaires about their training history and current practicing habits. Then, a heterogeneous set of diagnostic and standardized academic cognitive tests should be also undertaken, in order to verify that TPD indeed surfaces as a correlate of academic performance beyond the current experimental settings.

Provided that the results of the above-suggested research shows that TPD is a key cognitive resource for academic performance, then a research program could begin which systematically identifies the aspects of musical training that are most associated with higher TPD, based on the musical training questionnaires for musician participants who took the test batteries to verify TPD. Aspects of their training that statistically relate to higher TPD scores could then be consolidated to create exercises to train TPD, and these exercises could then be entered into their own randomized controlled trials to see if they improve academic performance in low-performing students.

While many unforeseen factors would likely affect this proposed sketch, the potential for TPD to be trained was demonstrated by its correlation with practice hours of musicians at their time of testing, and its role in the current cognitive performance of participants was unquestionable. Thus TPD is a good candidate for the development of intervention tools, albeit first requiring a dedicated, long term research program. Considering that music is already offered in many mainstream curricula, incorporation of an intervention tool that is based off of the most effective aspects of musicians' training should be relatively easy to integrate into school curricula and therefore reach a large body of students. Optimally, this could help academic performance among low-achieving students by improving their TPD, in turn improving the degree to which students learn auditorily imparted information.

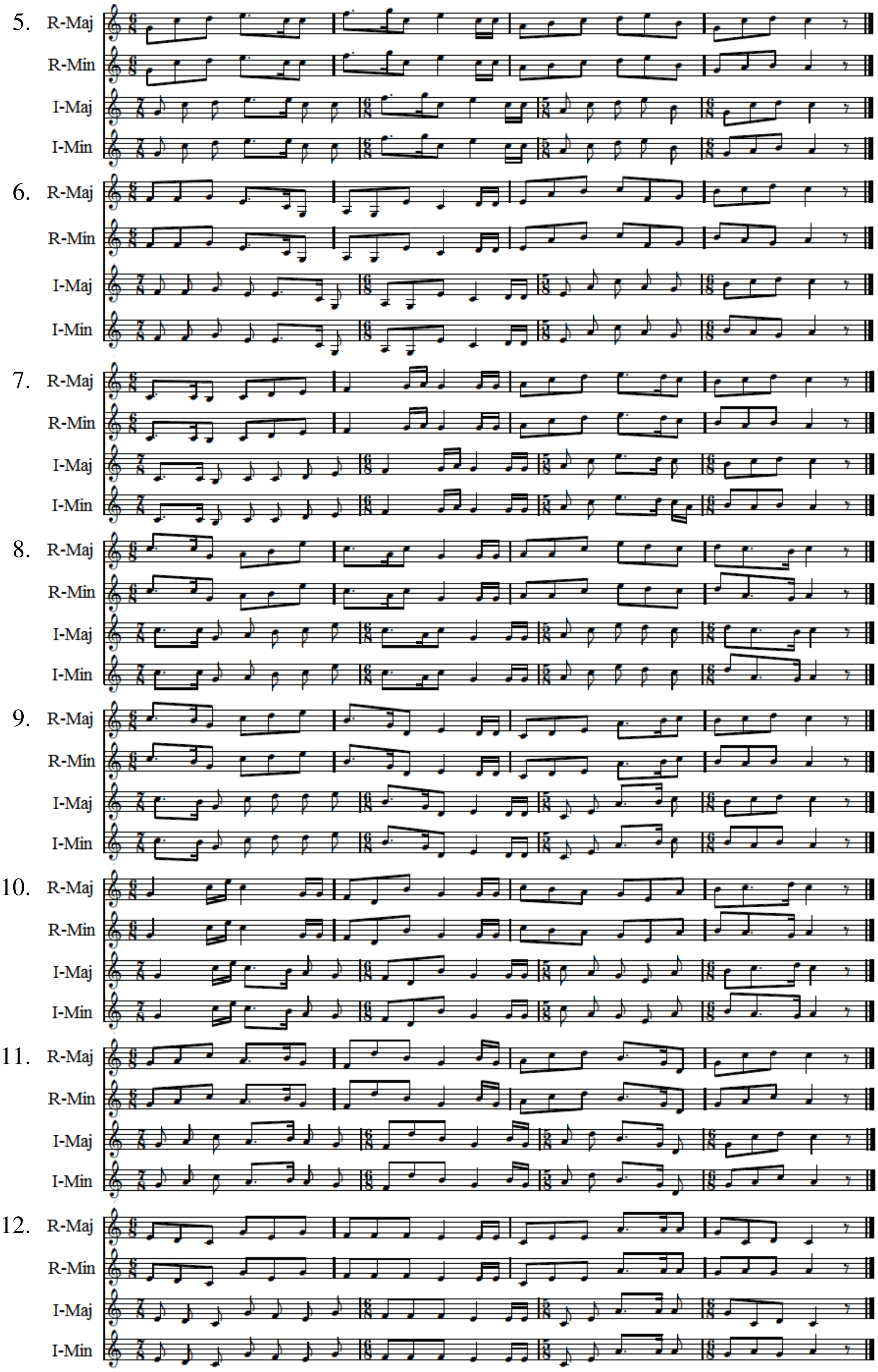
In sum, three major outlooks for future implementations of the current fundamental research include repeating the current EEG paradigm with non-musician participants, exploring previous N500 paradigms with respect to metrical context, and developing an extensive research program to create an academic intervention tool to improve Time and Pitch Discrimination. These are but some of the options for continuing the research in this dissertation, which has forged multiple novel concepts that may become tributaries to future bodies of work.

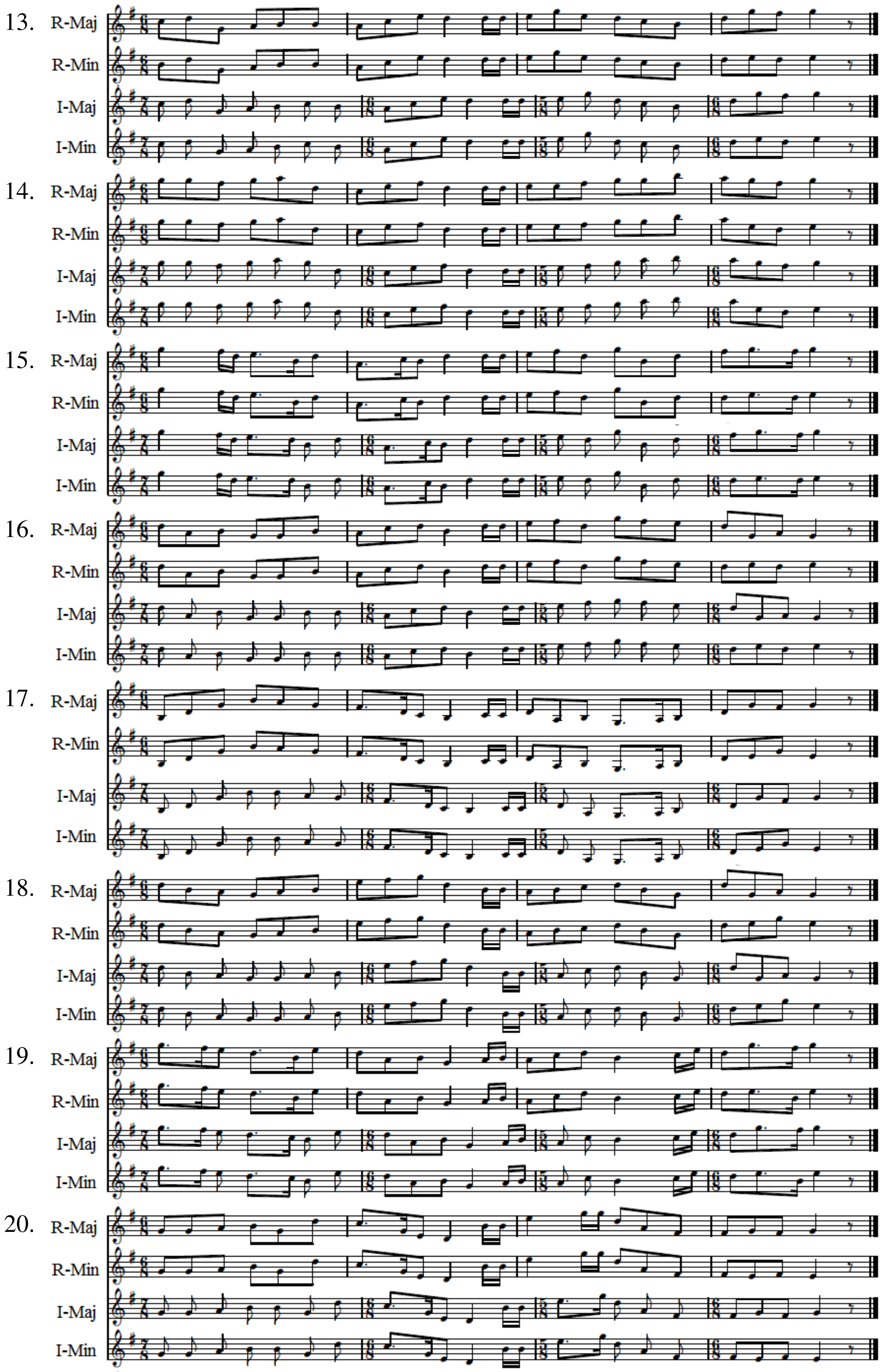
# Appendix A

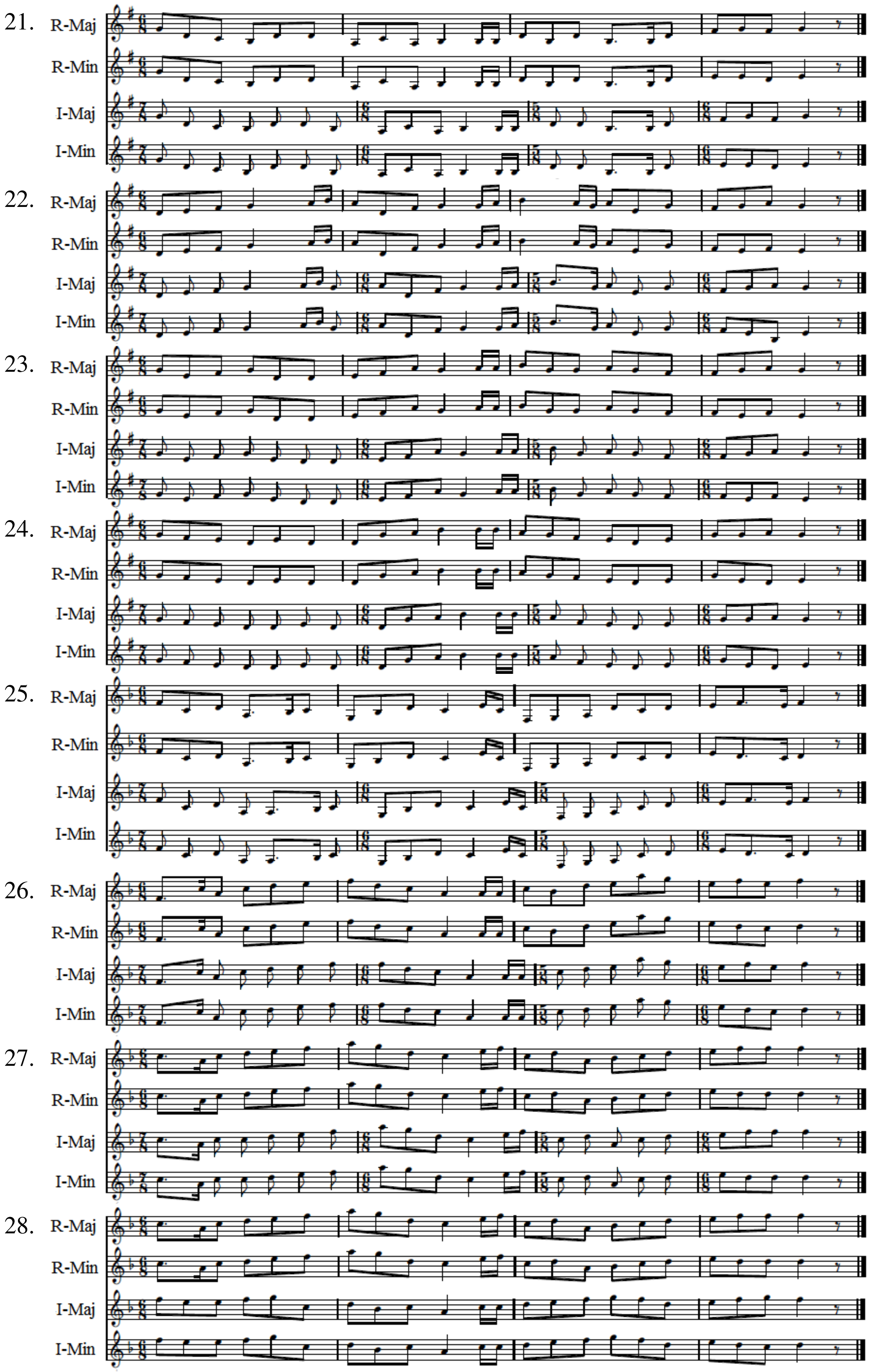
# Music materials

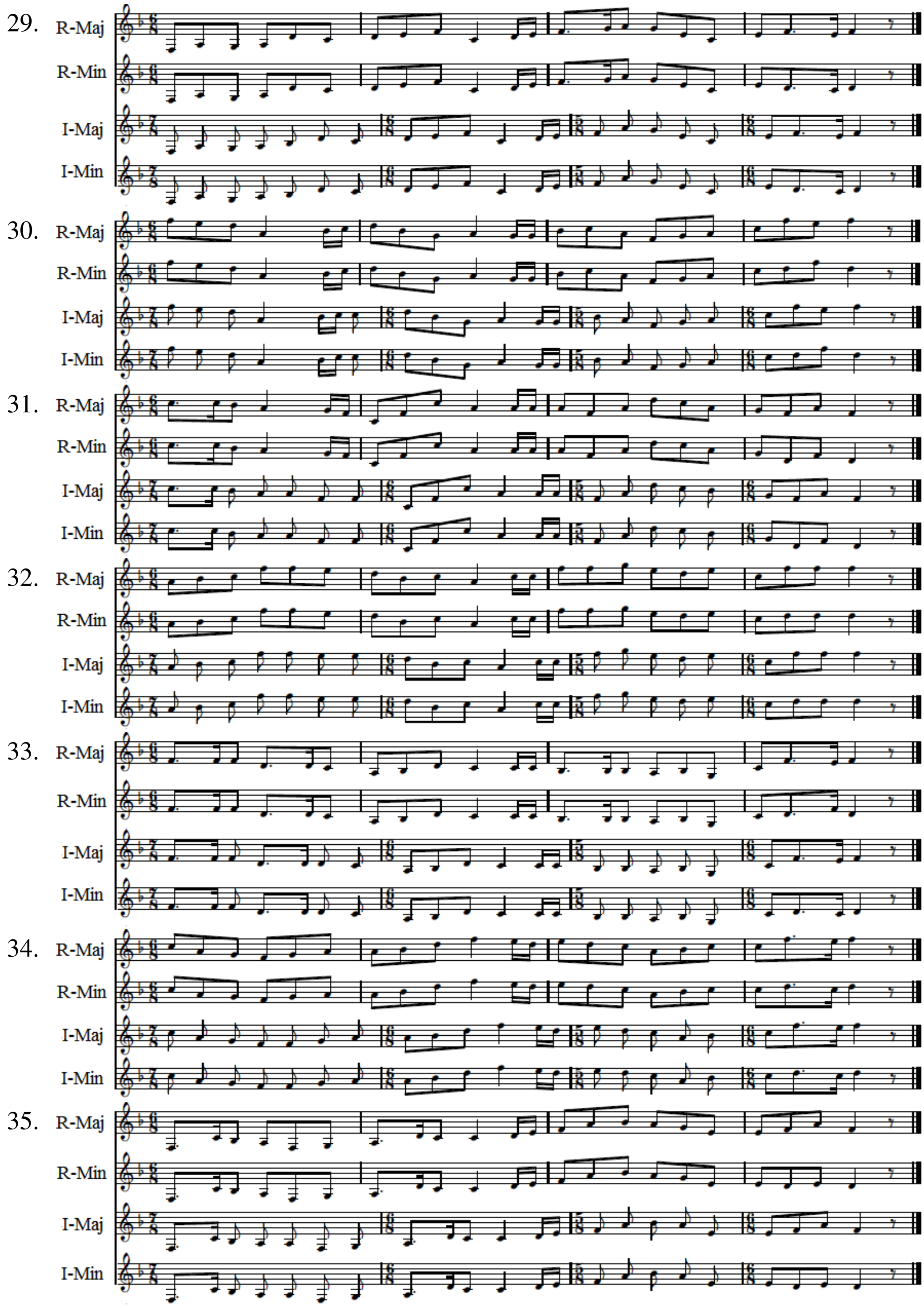
## A.1 Stimulus items

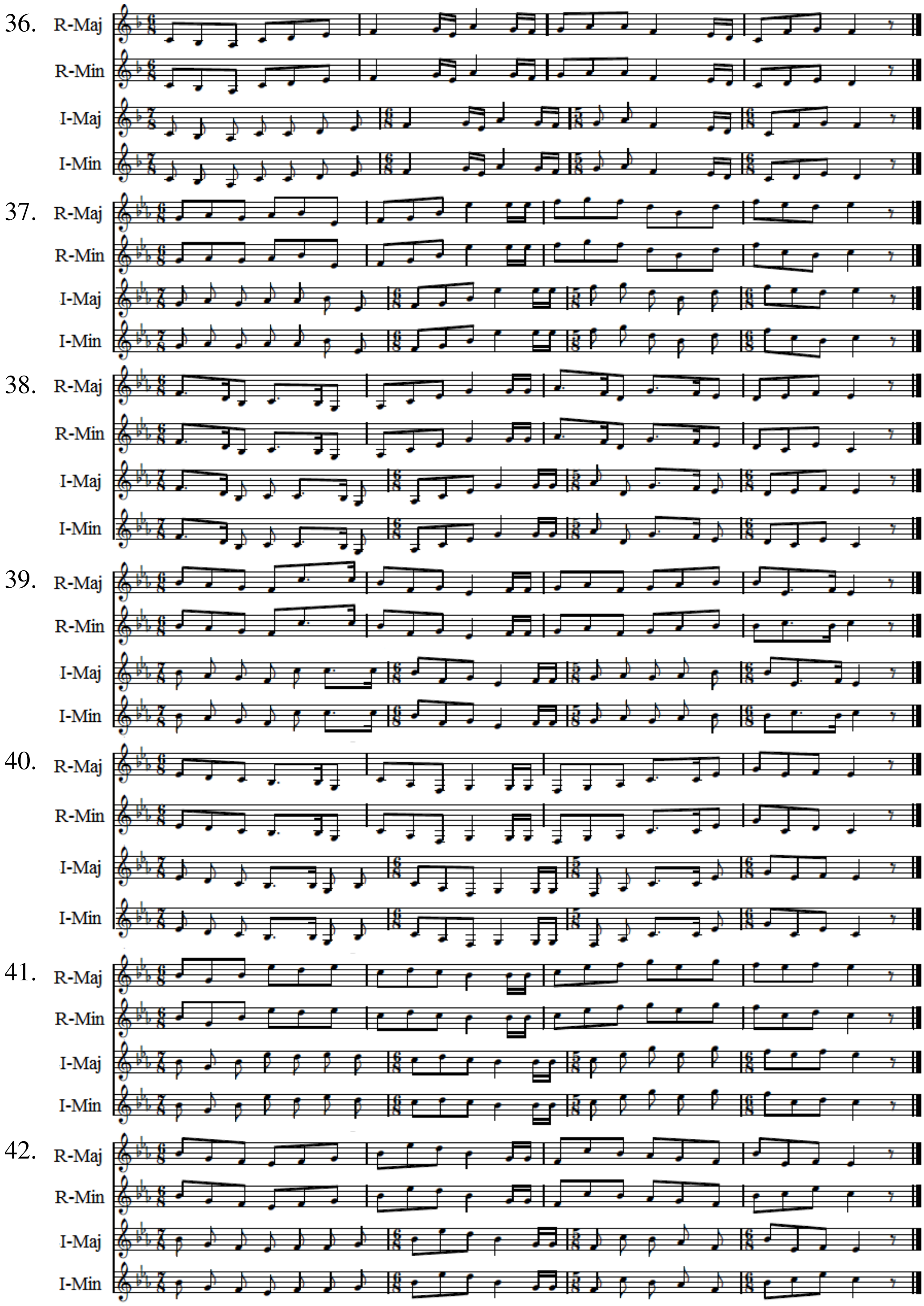
*Music stimuli items: Each of 60 melodies has four permutations. R = Regular metricality. I = Irregular metricality. Maj = Syntax remains in original major key. Min = Syntax modulates to relative minor key. Melodies span five tonal keys (C/a,G/e,F/d,$E^b$/c,D/b).*

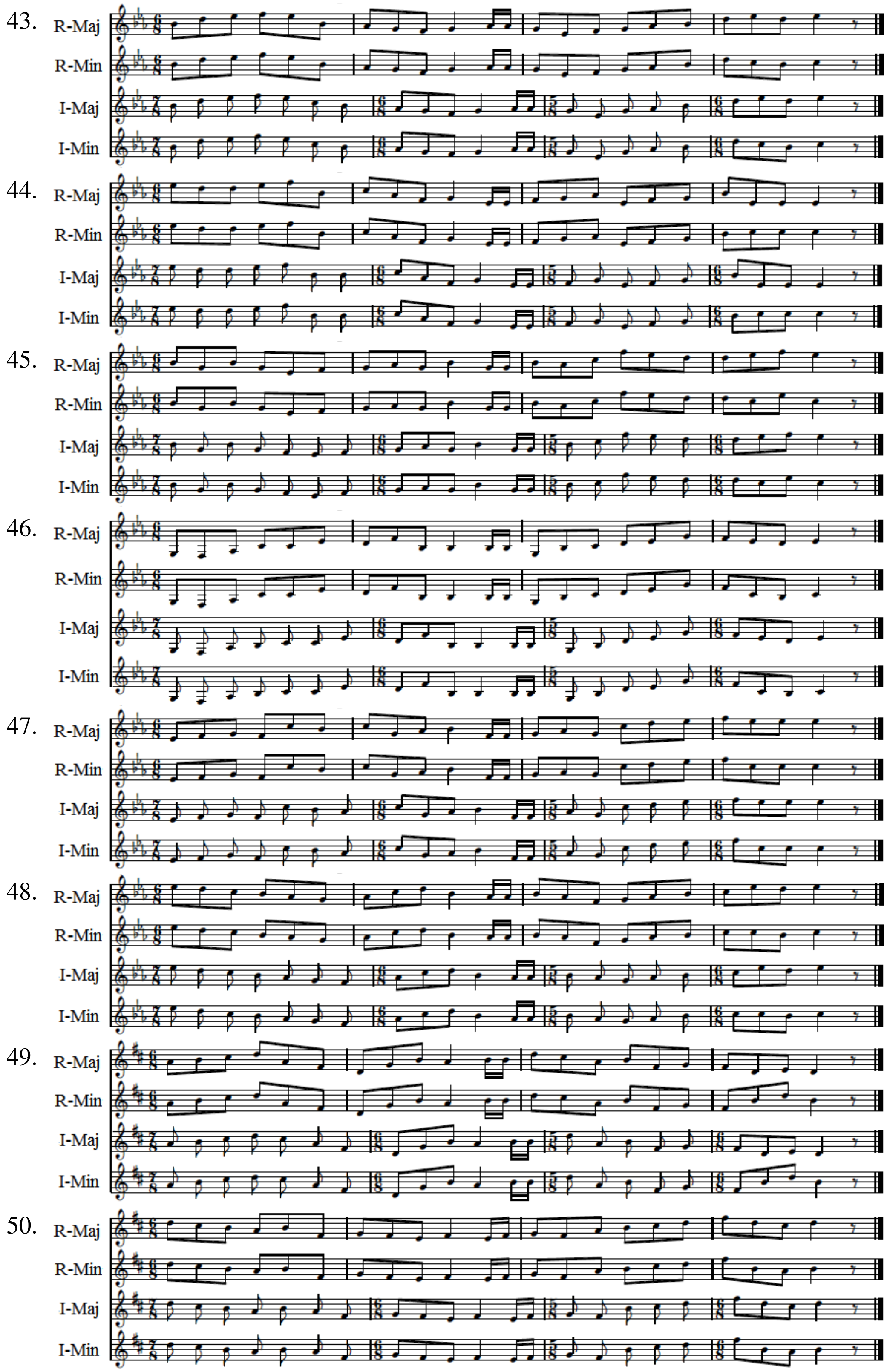

5. R-Maj
R-Min
I-Maj
I-Min
6. R-Maj
R-Min
I-Maj
I-Min
7. R-Maj
R-Min
I-Maj
I-Min
8. R-Maj
R-Min
I-Maj
I-Min
9. R-Maj
R-Min
I-Maj
I-Min
10. R-Maj
R-Min
I-Maj
I-Min
11. R-Maj
R-Min
I-Maj
I-Min
12. R-Maj
R-Min
I-Maj
I-Min

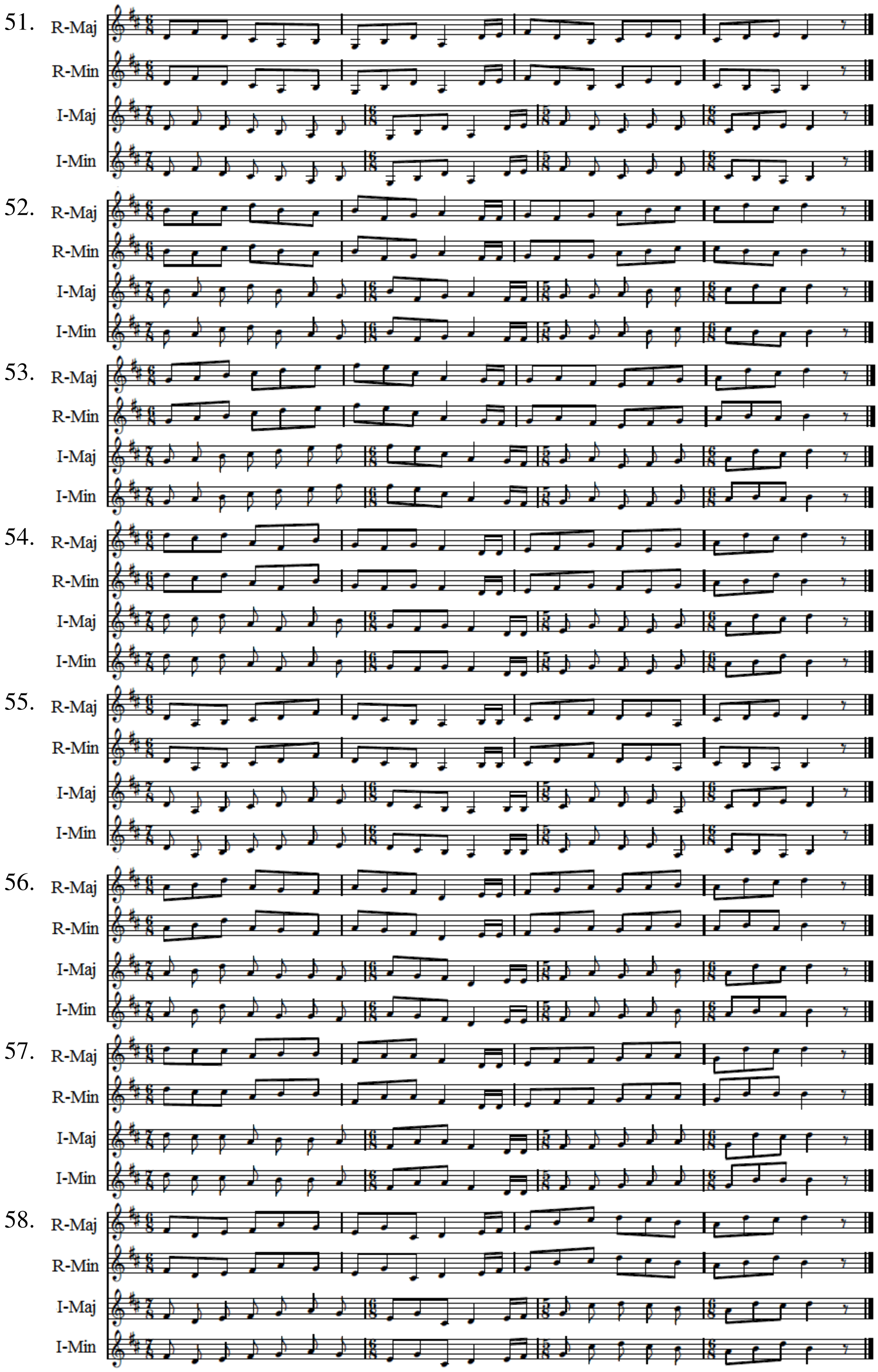

13. R-Maj
R-Min
I-Maj
I-Min
14. R-Maj
R-Min
I-Maj
I-Min
15. R-Maj
R-Min
I-Maj
I-Min
16. R-Maj
R-Min
I-Maj
I-Min
17. R-Maj
R-Min
I-Maj
I-Min
18. R-Maj
R-Min
I-Maj
I-Min
19. R-Maj
R-Min
I-Maj
I-Min
20. R-Maj
R-Min
I-Maj
I-Min


21.
R-Maj
R-Min
I-Maj
I-Min
22.
R-Maj
R-Min
I-Maj
I-Min
23.
R-Maj
R-Min
I-Maj
I-Min
24.
R-Maj
R-Min
I-Maj
I-Min
25.
R-Maj
R-Min
I-Maj
I-Min
26.
R-Maj
R-Min
I-Maj
I-Min
27.
R-Maj
R-Min
I-Maj
I-Min
28.
R-Maj
R-Min
I-Maj
I-Min


29.
R-Maj
R-Min
I-Maj
I-Min
30.
R-Maj
R-Min
I-Maj
I-Min
31.
R-Maj
R-Min
I-Maj
I-Min
32.
R-Maj
R-Min
I-Maj
I-Min
33.
R-Maj
R-Min
I-Maj
I-Min
34.
R-Maj
R-Min
I-Maj
I-Min
35.
R-Maj
R-Min
I-Maj
I-Min

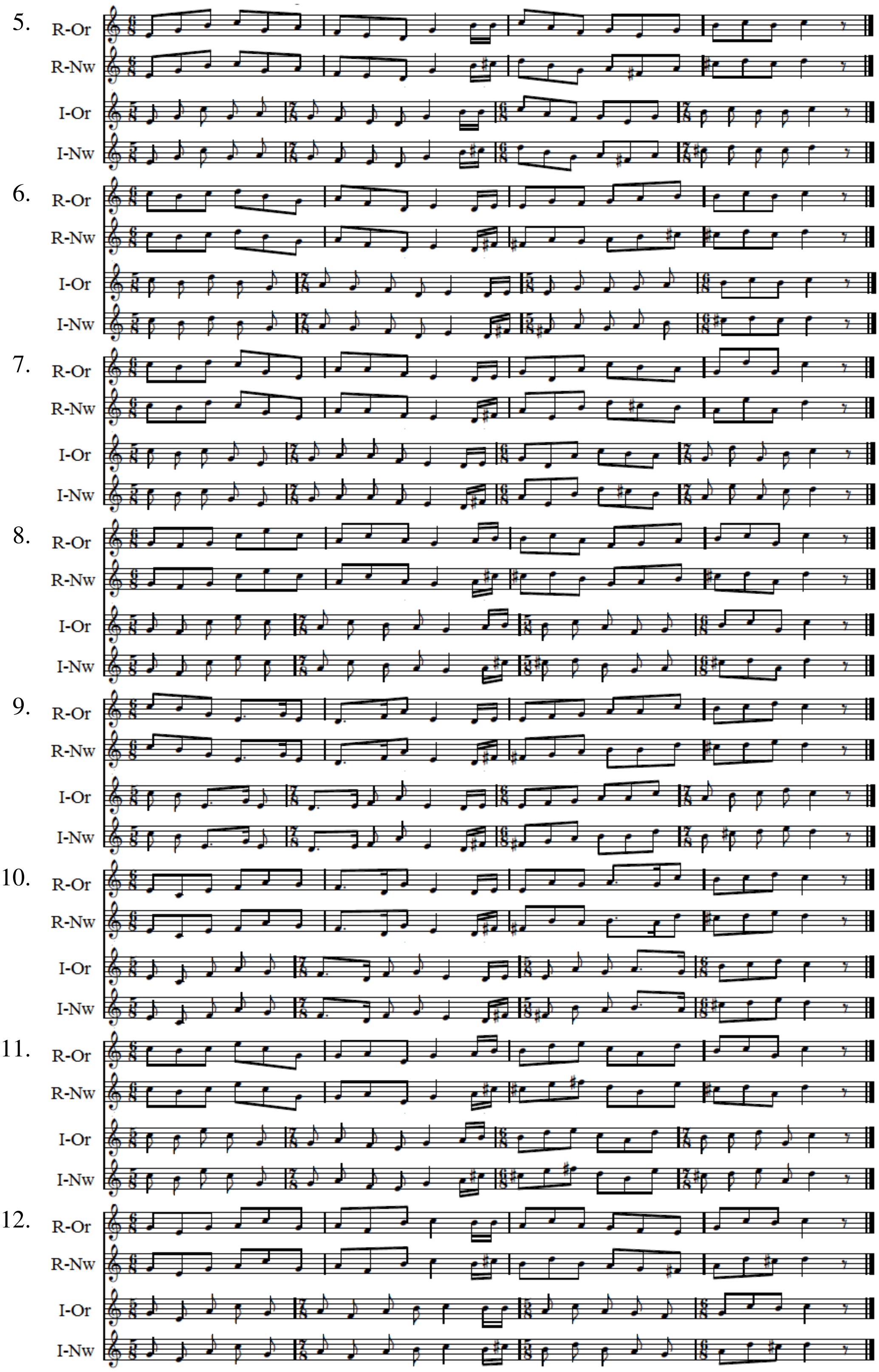

36.
R-Maj
R-Min
I-Maj
I-Min
37.
R-Maj
R-Min
I-Maj
I-Min
38.
R-Maj
R-Min
I-Maj
I-Min
39.
R-Maj
R-Min
I-Maj
I-Min
40.
R-Maj
R-Min
I-Maj
I-Min
41.
R-Maj
R-Min
I-Maj
I-Min
42.
R-Maj
R-Min
I-Maj
I-Min


43. R-Maj
R-Min
I-Maj
I-Min
44. R-Maj
R-Min
I-Maj
I-Min
45. R-Maj
R-Min
I-Maj
I-Min
46. R-Maj
R-Min
I-Maj
I-Min
47. R-Maj
R-Min
I-Maj
I-Min
48. R-Maj
R-Min
I-Maj
I-Min
49. R-Maj
R-Min
I-Maj
I-Min
50. R-Maj
R-Min
I-Maj
I-Min


51.
R-Maj
R-Min
I-Maj
I-Min
52.
R-Maj
R-Min
I-Maj
I-Min
53.
R-Maj
R-Min
I-Maj
I-Min
54.
R-Maj
R-Min
I-Maj
I-Min
55.
R-Maj
R-Min
I-Maj
I-Min
56.
R-Maj
R-Min
I-Maj
I-Min
57.
R-Maj
R-Min
I-Maj
I-Min
58.
R-Maj
R-Min
I-Maj
I-Min


## A.2 Filler items

*Music filler items: Each of 60 melodies has four permutations. R = Regular metricality. I = Irregular metricality. Or = Syntax remains in the original major key. Nw = Syntax modulates to a new major key. Melodies span five tonal keys (C/D,G/A,F/G,E$^b$/F,D/E).*


5.
R-Or
R-Nw
I-Or
I-Nw
6.
R-Or
R-Nw
I-Or
I-Nw
7.
R-Or
R-Nw
I-Or
I-Nw
8.
R-Or
R-Nw
I-Or
I-Nw
9.
R-Or
R-Nw
I-Or
I-Nw
10.
R-Or
R-Nw
I-Or
I-Nw
11.
R-Or
R-Nw
I-Or
I-Nw
12.
R-Or
R-Nw
I-Or
I-Nw

13.

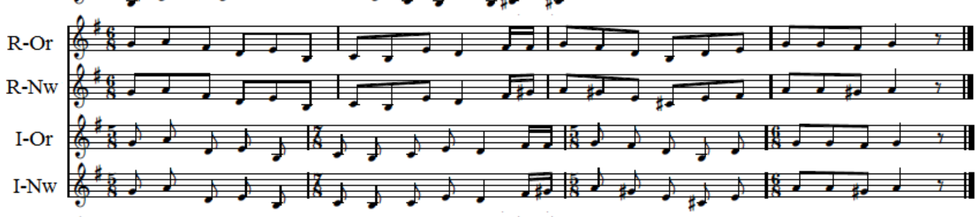

14.


15.

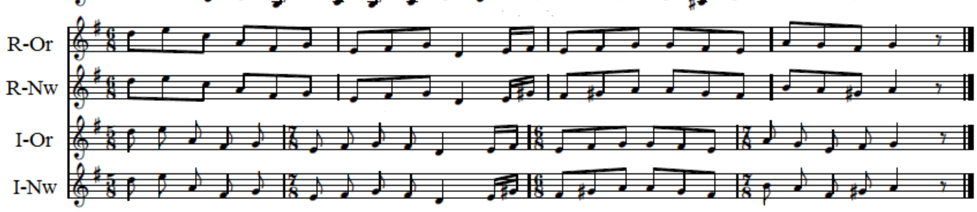

16.


17.


18.




19.




20.

21.


22.


23.


24.


25.


26.


27.


28.

29.


30.


31.


32.


33.

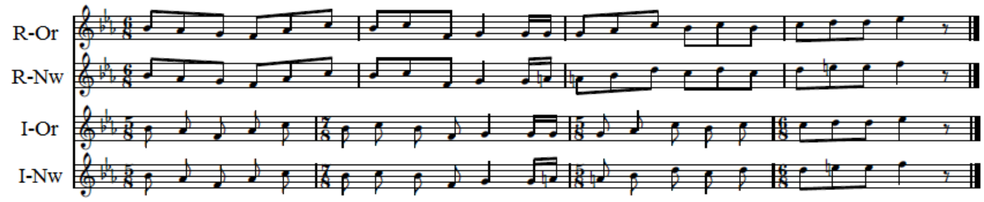

34.


35.


36.

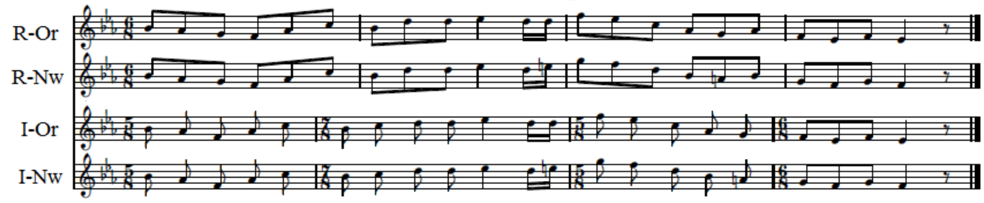
37.



38.



39.



40.



41.

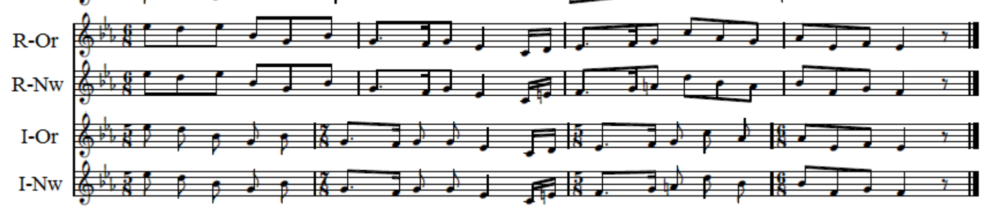


42.



43.



44.


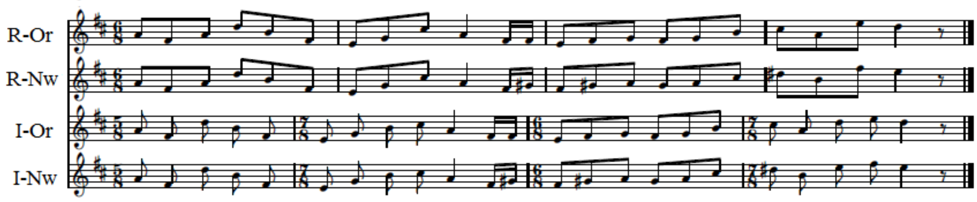
45.


46.


47.

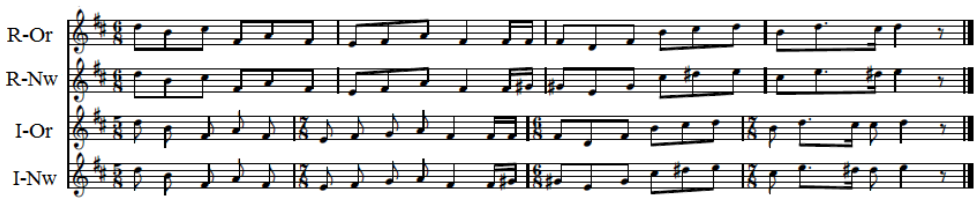

48.


49.

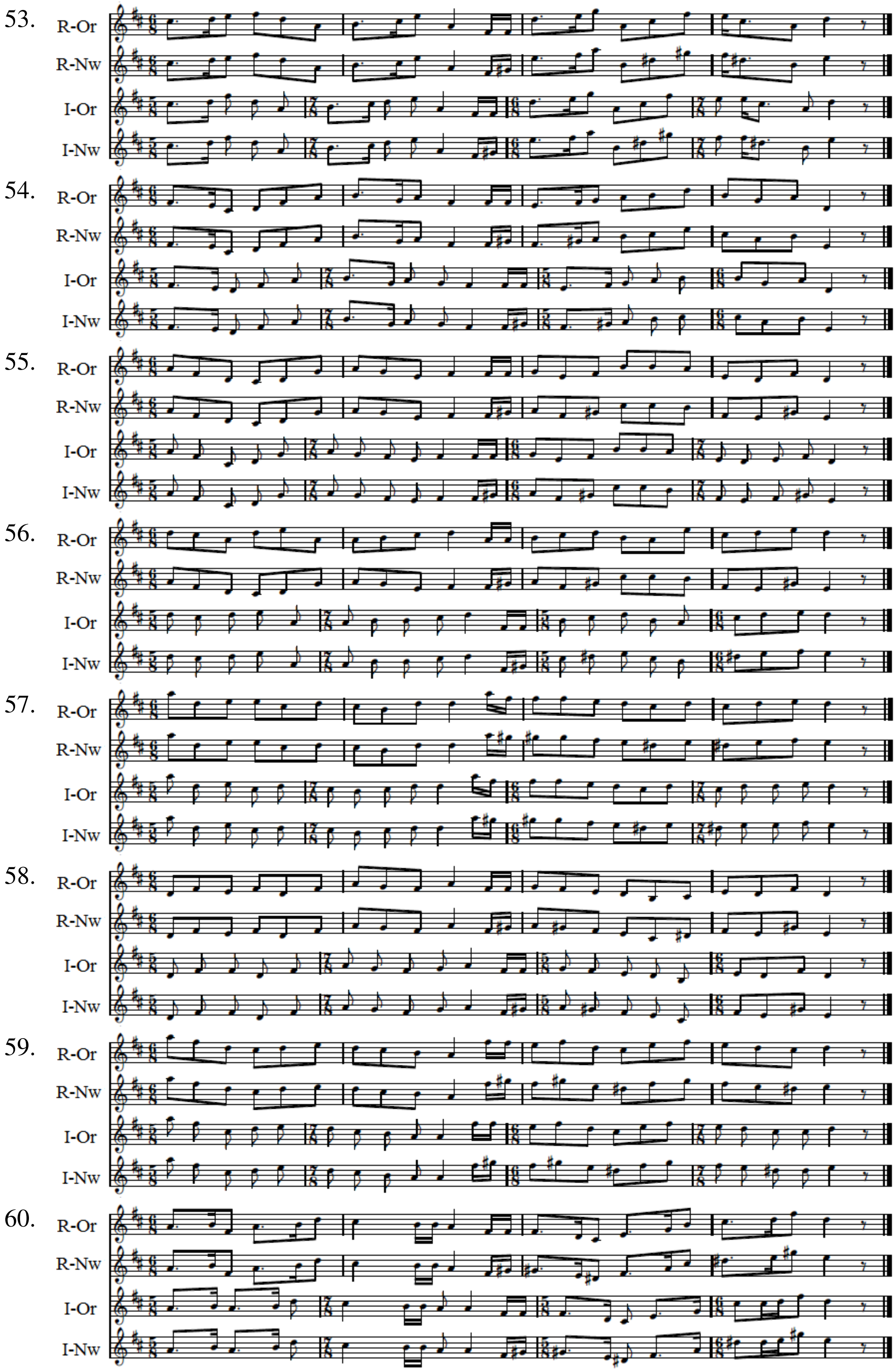



50.


51.




52.


53.
R-Or
R-Nw
I-Or
I-Nw
54.
R-Or
R-Nw
I-Or
I-Nw
55.
R-Or
R-Nw
I-Or
I-Nw
56.
R-Or
R-Nw
I-Or
I-Nw
57.
R-Or
R-Nw
I-Or
I-Nw
58.
R-Or
R-Nw
I-Or
I-Nw
59.
R-Or
R-Nw
I-Or
I-Nw
60.
R-Or
R-Nw
I-Or
I-Nw

# Appendix B

# Language materials

## B.1 Stimulus items

*Language stimulus items: Each of 60 sentences has four permutations. R = Regular metricality. I = Irregular metricality. H = High-attached relative clause. L = Low-attached relative clause.*

1. R-H Da draußen plaudern die Brüder der Ärztin, die Salzburg vor kurzem besuchten und mochten.
   R-L Da draußen plaudern die Brüder der Ärztin, die Salzburg vor kurzem besuchte und mochte.
   I-H Da draußen fachsimpeln die Brüder der Ärztin, die Wien vor kurzem besuchten und mochten.
   I-L Da draußen fachsimpeln die Brüder der Ärztin, die Wien vor kurzem besuchte und mochte.
2. R-H Dort drüben rasten die Brüder der Chefin, die Moskau vor kurzem besuchten und mochten.
   R-L Dort drüben rasten die Brüder der Chefin, die Moskau vor kurzem besuchte und mochte.
   I-H Dort drüben faulenzen die Brüder der Chefin, die Perm vor kurzem besuchten und mochten.
   I-L Dort drüben faulenzen die Brüder der Chefin, die Perm vor kurzem besuchte und mochte.
3. R-H Da draußen schuften die Brüder der Türkin, die Kapstadt vor kurzem besuchten und mochten.
   R-L Da draußen schuften die Brüder der Türkin, die Kapstadt vor kurzem besuchte und mochte.
   I-H Da draußen arbeiten die Brüder der Türkin, die Paarl vor kurzem besuchten und mochten.
   I-L Da draußen arbeiten die Brüder der Türkin, die Paarl vor kurzem besuchte und mochte.
4. R-H Da draußen speisen die Brüder der Russin, die London vor kurzem besuchten und mochten.
   R-L Da draußen speisen die Brüder der Russin, die London vor kurzem besuchte und mochte.
   I-H Da draußen frühstücken die Brüder der Russin, die Leeds vor kurzem besuchten und mochten.
   I-L Da draußen frühstücken die Brüder der Russin, die Leeds vor kurzem besuchte und mochte.
5. R-H Dort vorne plaudern die Diener der Türkin, die Lübeck vor kurzem besuchten und mochten.
   R-L Dort vorne plaudern die Diener der Türkin, die Lübeck vor kurzem besuchte und mochte.
   I-H Dort vorne fachsimpeln die Diener der Türkin, die Trier vor kurzem besuchten und mochten.
   I-L Dort vorne fachsimpeln die Diener der Türkin, die Trier vor kurzem besuchte und mochte.

6. R-H Da draußen rasten die Diener der Witwe, die Basel vor kurzem besuchten und mochten.
   R-L Da draußen rasten die Diener der Witwe, die Basel vor kurzem besuchte und mochte.
   I-H Da draußen faulenzen die Diener der Witwe, die Bern vor kurzem besuchten und mochten.
   I-L Da draußen faulenzen die Diener der Witwe, die Bern vor kurzem besuchte und mochte.
7. R-H Dort drinnen schuften die Diener der Schwedin, die Brüssel vor kurzem besuchten und mochten.
   R-L Dort drinnen schuften die Diener der Schwedin, die Brüssel vor kurzem besuchte und mochte.
   I-H Dort drinnen arbeiten die Diener der Schwedin, die Gent vor kurzem besuchten und mochten.
   I-L Dort drinnen arbeiten die Diener der Schwedin, die Gent vor kurzem besuchte und mochte.
8. R-H Da vorne speisen die Diener der Russin, die Moskau vor kurzem besuchten und mochten.
   R-L Da vorne speisen die Diener der Russin, die Moskau vor kurzem besuchte und mochte.
   I-H Da vorne frühstücken die Diener der Russin, die Perm vor kurzem besuchten und mochten.
   I-L Da vorne frühstücken die Diener der Russin, die Perm vor kurzem besuchte und mochte.
9. R-H Dort hinten plaudern die Freunde der Polin, die München vor kurzem besuchten und mochten.
   R-L Dort hinten plaudern die Freunde der Polin, die München vor kurzem besuchte und mochte.
   I-H Dort hinten fachsimpeln die Freunde der Polin, die Fürth vor kurzem besuchten und mochten.
   I-L Dort hinten fachsimpeln die Freunde der Polin, die Fürth vor kurzem besuchte und mochte.
10. R-H Dort vorne rasten die Freunde der Lettin, die Lübeck vor kurzem besuchten und mochten.
    R-L Dort vorne rasten die Freunde der Lettin, die Lübeck vor kurzem besuchte und mochte.
    I-H Dort vorne faulenzen die Freunde der Lettin, die Trier vor kurzem besuchten und mochten.
    I-L Dort vorne faulenzen die Freunde der Lettin, die Trier vor kurzem besuchte und mochte.
11. R-H Da hinten schuften die Freunde der Erbin, die Brüssel vor kurzem besuchten und mochten.
    R-L Da hinten schuften die Freunde der Erbin, die Brüssel vor kurzem besuchte und mochte.
    I-H Da hinten arbeiten die Freunde der Erbin, die Gent vor kurzem besuchten und mochten.
    I-L Da hinten arbeiten die Freunde der Erbin, die Gent vor kurzem besuchte und mochte.
12. R-H Da drüben speisen die Freunde der Kundin, die London vor kurzem besuchten und mochten.
    R-L Da drüben speisen die Freunde der Kundin, die London vor kurzem besuchte und mochte.
    I-H Da drüben frühstücken die Freunde der Kundin, die Leeds vor kurzem besuchten und mochten.
    I-L Da drüben frühstücken die Freunde der Kundin, die Leeds vor kurzem besuchte und mochte.
13. R-H Da vorne plaudern die Gäste der Lettin, die Moskau vor kurzem besuchten und mochten.
    R-L Da vorne plaudern die Gäste der Lettin, die Moskau vor kurzem besuchte und mochte.
    I-H Da vorne fachsimpeln die Gäste der Lettin, die Perm vor kurzem besuchten und mochten.
    I-L Da vorne fachsimpeln die Gäste der Lettin, die Perm vor kurzem besuchte und mochte.
14. R-H Da drüben rasten die Gäste der Dänin, die Frankfurt vor kurzem besuchten und mochten.
    R-L Da drüben rasten die Gäste der Dänin, die Frankfurt vor kurzem besuchte und mochte.
    I-H Da drüben faulenzen die Gäste der Dänin, die Mainz vor kurzem besuchten und mochten.
    I-L Da drüben faulenzen die Gäste der Dänin, die Mainz vor kurzem besuchte und mochte.
15. R-H Dort draußen schuften die Gäste der Griechin, die Stuttgart vor kurzem besuchten und mochten.

R-L Dort draußen schuften die Gäste der Griechin, die Stuttgart vor kurzem besuchte und mochte.
I-H Dort draußen arbeiten die Gäste der Griechin, die Ulm vor kurzem besuchten und mochten.
I-L Dort draußen arbeiten die Gäste der Griechin, die Ulm vor kurzem besuchte und mochte.
16. R-H Da drinnen speisen die Gäste der Polin, die Kapstadt vor kurzem besuchten und mochten.
R-L Da drinnen speisen die Gäste der Polin, die Kapstadt vor kurzem besuchte und mochte.
I-H Da drinnen frühstücken die Gäste der Polin, die Paarl vor kurzem besuchten und mochten.
I-L Da drinnen frühstücken die Gäste der Polin, die Paarl vor kurzem besuchte und mochte.
17. R-H Dort drinnen plaudern die Kumpel der Polin, die Salzburg vor kurzem besuchten und mochten.
R-L Dort drinnen plaudern die Kumpel der Polin, die Salzburg vor kurzem besuchte und mochte.
I-H Dort drinnen fachsimpeln die Kumpel der Polin, die Wien vor kurzem besuchten und mochten.
I-L Dort drinnen fachsimpeln die Kumpel der Polin, die Wien vor kurzem besuchte und mochte.
18. R-H Da vorne rasten die Kumpel der Dänin, die Bregenz vor kurzem besuchten und mochten.
R-L Da vorne rasten die Kumpel der Dänin, die Bregenz vor kurzem besuchte und mochte.
I-H Da vorne faulenzen die Kumpel der Dänin, die Linz vor kurzem besuchten und mochten.
I-L Da vorne faulenzen die Kumpel der Dänin, die Linz vor kurzem besuchte und mochte.
19. R-H Da draußen schuften die Kumpel der Griechin, die Frankfurt vor kurzem besuchten und mochten.
R-L Da draußen schuften die Kumpel der Griechin, die Frankfurt vor kurzem besuchte und mochte.
I-H Da draußen arbeiten die Kumpel der Griechin, die Mainz vor kurzem besuchten und mochten.
I-L Da draußen arbeiten die Kumpel der Griechin, die Mainz vor kurzem besuchte und mochte.
20. R-H Da hinten speisen die Kumpel der Britin, die München vor kurzem besuchten und mochten.
R-L Da hinten speisen die Kumpel der Britin, die München vor kurzem besuchte und mochte.
I-H Da hinten frühstücken die Kumpel der Britin, die Fürth vor kurzem besuchten und mochten.
I-L Da hinten frühstücken die Kumpel der Britin, die Fürth vor kurzem besuchte und mochte.
21. R-H Da draußen plaudern die Kunden der Finnin, die Warschau vor kurzem besuchten und mochten.
R-L Da draußen plaudern die Kunden der Finnin, die Warschau vor kurzem besuchte und mochte.
I-H Da draußen fachsimpeln die Kunden der Finnin, die Lodz vor kurzem besuchten und mochten.
I-L Da draußen fachsimpeln die Kunden der Finnin, die Lodz vor kurzem besuchte und mochte.
22. R-H Dort drüben rasten die Kunden der Britin, die Basel vor kurzem besuchten und mochten.
R-L Dort drüben rasten die Kunden der Britin, die Basel vor kurzem besuchte und mochte.
I-H Dort drüben faulenzen die Kunden der Britin, die Bern vor kurzem besuchten und mochten.
I-L Dort drüben faulenzen die Kunden der Britin, die Bern vor kurzem besuchte und mochte.
23. R-H Dort vorne schuften die Kunden der Griechin, die Bregenz vor kurzem besuchten und mochten.
R-L Dort vorne schuften die Kunden der Griechin, die Bregenz vor kurzem besuchte und mochte.
I-H Dort vorne arbeiten die Kunden der Griechin, die Linz vor kurzem besuchten und mochten.
I-L Dort vorne arbeiten die Kunden der Griechin, die Linz vor kurzem besuchte und mochte.
24. R-H Dort draußen speisen die Kunden der Chefin, die München vor kurzem besuchten und mochten.
R-L Dort draußen speisen die Kunden der Chefin, die München vor kurzem besuchte und mochte.

I-H Dort draußen frühstücken die Kunden der Chefin, die Fürth vor kurzem besuchten und mochten.
I-L Dort draußen frühstücken die Kunden der Chefin, die Fürth vor kurzem besuchte und mochte.
25. R-H Da hinten plaudern die Nachbarn der Dänin, die Patras vor kurzem besuchten und mochten.
R-L Da hinten plaudern die Nachbarn der Dänin, die Patras vor kurzem besuchte und mochte.
I-H Da hinten fachsimpeln die Nachbarn der Dänin, die Kos vor kurzem besuchten und mochten.
I-L Da hinten fachsimpeln die Nachbarn der Dänin, die Kos vor kurzem besuchte und mochte.
26. R-H Da vorne rasten die Nachbarn der Schwedin, die Lübeck vor kurzem besuchten und mochten.
R-L Da vorne rasten die Nachbarn der Schwedin, die Lübeck vor kurzem besuchte und mochte.
I-H Da vorne faulenzen die Nachbarn der Schwedin, die Trier vor kurzem besuchten und mochten.
I-L Da vorne faulenzen die Nachbarn der Schwedin, die Trier vor kurzem besuchte und mochte.
27. R-H Da drinnen schuften die Nachbarn der Britin, die Kapstadt vor kurzem besuchten und mochten.
R-L Da drinnen schuften die Nachbarn der Britin, die Kapstadt vor kurzem besuchte und mochte.
I-H Da drinnen arbeiten die Nachbarn der Britin, die Paarl vor kurzem besuchten und mochten.
I-L Da drinnen arbeiten die Nachbarn der Britin, die Paarl vor kurzem besuchte und mochte.
28. R-H Dort vorne speisen die Nachbarn der Kundin, die Stuttgart vor kurzem besuchten und mochten.
R-L Dort vorne speisen die Nachbarn der Kundin, die Stuttgart vor kurzem besuchte und mochte.
I-H Dort vorne frühstücken die Nachbarn der Kundin, die Ulm vor kurzem besuchten und mochten.
I-L Dort vorne frühstücken die Nachbarn der Kundin, die Ulm vor kurzem besuchte und mochte.
29. R-H Dort draußen plaudern die Neffen der Britin, die Bregenz vor kurzem besuchten und mochten.
R-L Dort draußen plaudern die Neffen der Britin, die Bregenz vor kurzem besuchte und mochte.
I-H Dort draußen fachsimpeln die Neffen der Britin, die Linz vor kurzem besuchten und mochten.
I-L Dort draußen fachsimpeln die Neffen der Britin, die Linz vor kurzem besuchte und mochte.
30. R-H Dort draußen rasten die Neffen der Finnin, die Nizza vor kurzem besuchten und mochten.
R-L Dort draußen rasten die Neffen der Finnin, die Nizza vor kurzem besuchte und mochte.
I-H Dort draußen faulenzen die Neffen der Finnin, die Cannes vor kurzem besuchten und mochten.
I-L Dort draußen faulenzen die Neffen der Finnin, die Cannes vor kurzem besuchte und mochte.
31. R-H Dort hinten schuften die Neffen der Erbin, die Warschau vor kurzem besuchten und mochten.
R-L Dort hinten schuften die Neffen der Erbin, die Warschau vor kurzem besuchte und mochte.
I-H Dort hinten arbeiten die Neffen der Erbin, die Lodz vor kurzem besuchten und mochten.
I-L Dort hinten arbeiten die Neffen der Erbin, die Lodz vor kurzem besuchte und mochte.
32. R-H Dort hinten speisen die Neffen der Dänin, die Ostrau vor kurzem besuchten und mochten.
R-L Dort hinten speisen die Neffen der Dänin, die Ostrau vor kurzem besuchte und mochte.
I-H Dort hinten frühstücken die Neffen der Dänin, die Prag vor kurzem besuchten und mochten.
I-L Dort hinten frühstücken die Neffen der Dänin, die Prag vor kurzem besuchte und mochte.
33. R-H Da drinnen plaudern die Nichten der Polin, die Brüssel vor kurzem besuchten und mochten.
R-L Da drinnen plaudern die Nichten der Polin, die Brüssel vor kurzem besuchte und mochte.
I-H Da drinnen fachsimpeln die Nichten der Polin, die Gent vor kurzem besuchten und mochten.

I-L Da drinnen fachsimpeln die Nichten der Polin, die Gent vor kurzem besuchte und mochte.
34. R-H Dort hinten rasten die Nichten der Russin, die Basel vor kurzem besuchten und mochten.
R-L Dort hinten rasten die Nichten der Russin, die Basel vor kurzem besuchte und mochte.
I-H Dort hinten faulenzen die Nichten der Russin, die Bern vor kurzem besuchten und mochten.
I-L Dort hinten faulenzen die Nichten der Russin, die Bern vor kurzem besuchte und mochte.
35. R-H Dort draußen schuften die Nichten der Schwedin, die Moskau vor kurzem besuchten und mochten.
R-L Dort draußen schuften die Nichten der Schwedin, die Moskau vor kurzem besuchte und mochte.
I-H Dort draußen arbeiten die Nichten der Schwedin, die Perm vor kurzem besuchten und mochten.
I-L Dort draußen arbeiten die Nichten der Schwedin, die Perm vor kurzem besuchte und mochte.
36. R-H Dort vorne speisen die Nichten der Tschechin, die Stuttgart vor kurzem besuchten und mochten.
R-L Dort vorne speisen die Nichten der Tschechin, die Stuttgart vor kurzem besuchte und mochte.
I-H Dort vorne frühstücken die Nichten der Tschechin, die Ulm vor kurzem besuchten und mochten.
I-L Dort vorne frühstücken die Nichten der Tschechin, die Ulm vor kurzem besuchte und mochte.
37. R-H Da drinnen plaudern die Onkel der Türkin, die Frankfurt vor kurzem besuchten und mochten.
R-L Da drinnen plaudern die Onkel der Türkin, die Frankfurt vor kurzem besuchte und mochte.
I-H Da drinnen fachsimpeln die Onkel der Türkin, die Mainz vor kurzem besuchten und mochten.
I-L Da drinnen fachsimpeln die Onkel der Türkin, die Mainz vor kurzem besuchte und mochte.
38. R-H Da drinnen rasten die Onkel der Chefin, die London vor kurzem besuchten und mochten.
R-L Da drinnen rasten die Onkel der Chefin, die London vor kurzem besuchte und mochte.
I-H Da drinnen faulenzen die Onkel der Chefin, die Leeds vor kurzem besuchten und mochten.
I-L Da drinnen faulenzen die Onkel der Chefin, die Leeds vor kurzem besuchte und mochte.
39. R-H Dort drinnen schuften die Onkel der Russin, die Nizza vor kurzem besuchten und mochten.
R-L Dort drinnen schuften die Onkel der Russin, die Nizza vor kurzem besuchte und mochte.
I-H Dort drinnen arbeiten die Onkel der Russin, die Cannes vor kurzem besuchten und mochten.
I-L Dort drinnen arbeiten die Onkel der Russin, die Cannes vor kurzem besuchte und mochte.
40. R-H Da hinten speisen die Onkel der Kundin, die Salzburg vor kurzem besuchten und mochten.
R-L Da hinten speisen die Onkel der Kundin, die Salzburg vor kurzem besuchte und mochte.
I-H Da hinten frühstücken die Onkel der Kundin, die Wien vor kurzem besuchten und mochten.
I-L Da hinten frühstücken die Onkel der Kundin, die Wien vor kurzem besuchte und mochte.
41. R-H Dort drinnen plaudern die Schwestern der Türkin, die Nizza vor kurzem besuchten und mochten.
R-L Dort drinnen plaudern die Schwestern der Türkin, die Nizza vor kurzem besuchte und mochte.
I-H Dort drinnen fachsimpeln die Schwestern der Türkin, die Cannes vor kurzem besuchten und mochten.
I-L Dort drinnen fachsimpeln die Schwestern der Türkin, die Cannes vor kurzem besuchte und mochte.
42. R-H Da hinten rasten die Schwestern der Tschechin, die Warschau vor kurzem besuchten und mochten.
R-L Da hinten rasten die Schwestern der Tschechin, die Warschau vor kurzem besuchte und mochte.
I-H Da hinten faulenzen die Schwestern der Tschechin, die Lodz vor kurzem besuchten und mochten.
I-L Da hinten faulenzen die Schwestern der Tschechin, die Lodz vor kurzem besuchte und mochte.

43. R-H Da vorne schuften die Schwestern der Erbin, die Ostrau vor kurzem besuchten und mochten.
R-L Da vorne schuften die Schwestern der Erbin, die Ostrau vor kurzem besuchte und mochte.
I-H Da vorne arbeiten die Schwestern der Erbin, die Prag vor kurzem besuchten und mochten.
I-L Da vorne arbeiten die Schwestern der Erbin, die Prag vor kurzem besuchte und mochte.
44. R-H Da vorne speisen die Schwestern der Witwe, die Patras vor kurzem besuchten und mochten.
R-L Da vorne speisen die Schwestern der Witwe, die Patras vor kurzem besuchte und mochte.
I-H Da vorne frühstücken die Schwestern der Witwe, die Kos vor kurzem besuchten und mochten.
I-L Da vorne frühstücken die Schwestern der Witwe, die Kos vor kurzem besuchte und mochte.
45. R-H Da drüben plaudern die Söhne der Chefin, die Warschau vor kurzem besuchten und mochten.
R-L Da drüben plaudern die Söhne der Chefin, die Warschau vor kurzem besuchte und mochte.
I-H Da drüben fachsimpeln die Söhne der Chefin, die Lodz vor kurzem besuchten und mochten.
I-L Da drüben fachsimpeln die Söhne der Chefin, die Lodz vor kurzem besuchte und mochte.
46. R-H Da hinten rasten die Söhne der Finnin, die Lübeck vor kurzem besuchten und mochten.
R-L Da hinten rasten die Söhne der Finnin, die Lübeck vor kurzem besuchte und mochte.
I-H Da hinten faulenzen die Söhne der Finnin, die Trier vor kurzem besuchten und mochten.
I-L Da hinten faulenzen die Söhne der Finnin, die Trier vor kurzem besuchte und mochte.
47. R-H Da drüben schuften die Söhne der Ärztin, die München vor kurzem besuchten und mochten.
R-L Da drüben schuften die Söhne der Ärztin, die München vor kurzem besuchte und mochte.
I-H Da drüben arbeiten die Söhne der Ärztin, die Fürth vor kurzem besuchten und mochten.
I-L Da drüben arbeiten die Söhne der Ärztin, die Fürth vor kurzem besuchte und mochte.
48. R-H Dort drüben speisen die Söhne der Griechin, die Patras vor kurzem besuchten und mochten.
R-L Dort drüben speisen die Söhne der Griechin, die Patras vor kurzem besuchte und mochte.
I-H Dort drüben frühstücken die Söhne der Griechin, die Kos vor kurzem besuchten und mochten.
I-L Dort drüben frühstücken die Söhne der Griechin, die Kos vor kurzem besuchte und mochte.
49. R-H Dort drüben plaudern die Tanten der Erbin, die Kapstadt vor kurzem besuchten und mochten.
R-L Dort drüben plaudern die Tanten der Erbin, die Kapstadt vor kurzem besuchte und mochte.
I-H Dort drüben fachsimpeln die Tanten der Erbin, die Paarl vor kurzem besuchten und mochten.
I-L Dort drüben fachsimpeln die Tanten der Erbin, die Paarl vor kurzem besuchte und mochte.
50. R-H Dort drinnen rasten die Tanten der Lettin, die Basel vor kurzem besuchten und mochten.
R-L Dort drinnen rasten die Tanten der Lettin, die Basel vor kurzem besuchte und mochte.
I-H Dort drinnen faulenzen die Tanten der Lettin, die Bern vor kurzem besuchten und mochten.
I-L Dort drinnen faulenzen die Tanten der Lettin, die Bern vor kurzem besuchte und mochte.
51. R-H Dort drüben schuften die Tanten der Kundin, die Ostrau vor kurzem besuchten und mochten.
R-L Dort drüben schuften die Tanten der Kundin, die Ostrau vor kurzem besuchte und mochte.
I-H Dort drüben arbeiten die Tanten der Kundin, die Prag vor kurzem besuchten und mochten.
I-L Dort drüben arbeiten die Tanten der Kundin, die Prag vor kurzem besuchte und mochte.
52. R-H Dort drüben speisen die Tanten der Schwedin, die Bregenz vor kurzem besuchten und mochten.

R-L Dort drüben speisen die Tanten der Schwedin, die Bregenz vor kurzem besuchte und mochte.
I-H Dort drüben frühstücken die Tanten der Schwedin, die Linz vor kurzem besuchten und mochten.
I-L Dort drüben frühstücken die Tanten der Schwedin, die Linz vor kurzem besuchte und mochte.

53. R-H Dort draußen plaudern die Töchter der Tschechin, die Nizza vor kurzem besuchten und mochten.
R-L Dort draußen plaudern die Töchter der Tschechin, die Nizza vor kurzem besuchte und mochte.
I-H Dort draußen fachsimpeln die Töchter der Tschechin, die Cannes vor kurzem besuchten und mochten.
I-L Dort draußen fachsimpeln die Töchter der Tschechin, die Cannes vor kurzem besuchte und mochte.

54. R-H Dort vorne rasten die Töchter der Ärztin, die Patras vor kurzem besuchten und mochten.
R-L Dort vorne rasten die Töchter der Ärztin, die Patras vor kurzem besuchte und mochte.
I-H Dort vorne faulenzen die Töchter der Ärztin, die Kos vor kurzem besuchten und mochten.
I-L Dort vorne faulenzen die Töchter der Ärztin, die Kos vor kurzem besuchte und mochte.

55. R-H Da drüben schuften die Töchter der Finnin, die London vor kurzem besuchten und mochten.
R-L Da drüben schuften die Töchter der Finnin, die London vor kurzem besuchte und mochte.
I-H Da drüben arbeiten die Töchter der Finnin, die Leeds vor kurzem besuchten und mochten.
I-L Da drüben arbeiten die Töchter der Finnin, die Leeds vor kurzem besuchte und mochte.

56. R-H Dort hinten speisen die Töchter der Witwe, die Salzburg vor kurzem besuchten und mochten.
R-L Dort hinten speisen die Töchter der Witwe, die Salzburg vor kurzem besuchte und mochte.
I-H Dort hinten frühstücken die Töchter der Witwe, die Wien vor kurzem besuchten und mochten.
I-L Dort hinten frühstücken die Töchter der Witwe, die Wien vor kurzem besuchte und mochte.

57. R-H Da drüben plaudern die Vettern der Tschechin, die Ostrau vor kurzem besuchten und mochten.
R-L Da drüben plaudern die Vettern der Tschechin, die Ostrau vor kurzem besuchte und mochte.
I-H Da drüben fachsimpeln die Vettern der Tschechin, die Prag vor kurzem besuchten und mochten.
I-L Da drüben fachsimpeln die Vettern der Tschechin, die Prag vor kurzem besuchte und mochte.

58. R-H Dort hinten rasten die Vettern der Ärztin, die Stuttgart vor kurzem besuchten und mochten.
R-L Dort hinten rasten die Vettern der Ärztin, die Stuttgart vor kurzem besuchte und mochte.
I-H Dort hinten faulenzen die Vettern der Ärztin, die Ulm vor kurzem besuchten und mochten.
I-L Dort hinten faulenzen die Vettern der Ärztin, die Ulm vor kurzem besuchte und mochte.

59. R-H Da drinnen schuften die Vettern der Lettin, die Brüssel vor kurzem besuchten und mochten.
R-L Da drinnen schuften die Vettern der Lettin, die Brüssel vor kurzem besuchte und mochte.
I-H Da drinnen arbeiten die Vettern der Lettin, die Gent vor kurzem besuchten und mochten.
I-L Da drinnen arbeiten die Vettern der Lettin, die Gent vor kurzem besuchte und mochte.

60. R-H Dort drinnen speisen die Vettern der Witwe, die Frankfurt vor kurzem besuchten und mochten.
R-L Dort drinnen speisen die Vettern der Witwe, die Frankfurt vor kurzem besuchte und mochte.
I-H Dort drinnen frühstücken die Vettern der Witwe, die Mainz vor kurzem besuchten und mochten.
I-L Dort drinnen frühstücken die Vettern der Witwe, die Mainz vor kurzem besuchte und mochte.

## B.2 Filler items

*Language filler items: Each of 60 sentences has four permutations. R = Regular metricality. I = Irregular metricality. S = Subject relative clause. O = Object relative clause.*

1. R-S Drinnen verweilt die Afghanin aus Khost, die den Schuster vor kurzem erkannte und grüßte.
   R-O Drinnen verweilt die Afghanin aus Khost, die der Schuster vor kurzem erkannte und grüßte.
   I-S Drinnen wartet die Inderin aus Mau, die den Schuster vor kurzem erkannte und begrüßte.
   I-O Drinnen wartet die Inderin aus Mau, die der Schuster vor kurzem erkannte und begrüßte.
2. R-S Draußen erzählt die Afghanin aus Khost, die den Fischer vor kurzem erkannte und grüßte.
   R-O Draußen erzählt die Afghanin aus Khost, die der Fischer vor kurzem erkannte und grüßte.
   I-S Draußen redet die Inderin aus Mau, die den Fischer vor kurzem sah und begrüßte.
   I-O Draußen redet die Inderin aus Mau, die der Fischer vor kurzem sah und begrüßte.
3. R-S Draußen spaziert die Afghanin aus Khost, die den Jungen vor kurzem erkannte und grüßte.
   R-O Draußen spaziert die Afghanin aus Khost, die der Junge vor kurzem erkannte und grüßte.
   I-S Draußen wandert die Inderin aus Mau, die den Jungen vor kurzem sah und begrüßte.
   I-O Draußen wandert die Inderin aus Mau, die der Junge vor kurzem sah und begrüßte.
4. R-S Drinnen entspannt die Afghanin aus Khost, die den Bauern vor kurzem erkannte und grüßte.
   R-O Drinnen entspannt die Afghanin aus Khost, die der Bauer vor kurzem erkannte und grüßte.
   I-S Drinnen rastet die Inderin aus Mau, die den Bauern vor kurzem erkannte und begrüßte.
   I-O Drinnen rastet die Inderin aus Mau, die der Bauer vor kurzem erkannte und begrüßte.
5. R-S Hinten verweilt die Agentin aus Graz, die den Dänen vor kurzem erkannte und grüßte.
   R-O Hinten verweilt die Agentin aus Graz, die der Däne vor kurzem erkannte und grüßte.
   I-S Hinten wartet die Mieterin aus Linz, die den Dänen vor kurzem sah und begrüßte.
   I-O Hinten wartet die Mieterin aus Linz, die der Däne vor kurzem sah und begrüßte.
6. R-S Drinnen erzählt die Agentin aus Graz, die den Briten vor kurzem erkannte und grüßte.
   R-O Drinnen erzählt die Agentin aus Graz, die der Brite vor kurzem erkannte und grüßte.
   I-S Drinnen redet die Mieterin aus Linz, die den Briten vor kurzem sah und begrüßte.
   I-O Drinnen redet die Mieterin aus Linz, die der Brite vor kurzem sah und begrüßte.
7. R-S Vorne spaziert die Agentin aus Graz, die den Händler vor kurzem erkannte und grüßte.
   R-O Vorne spaziert die Agentin aus Graz, die der Händler vor kurzem erkannte und grüßte.
   I-S Vorne wandert die Mieterin aus Linz, die den Händler vor kurzem erkannte und begrüßte.
   I-O Vorne wandert die Mieterin aus Linz, die der Händler vor kurzem erkannte und begrüßte.
8. R-S Drüben entspannt die Agentin aus Graz, die den Metzger vor kurzem erkannte und grüßte.
   R-O Drüben entspannt die Agentin aus Graz, die der Metzger vor kurzem erkannte und grüßte.
   I-S Drüben rastet die Mieterin aus Linz, die den Metzger vor kurzem erkannte und begrüßte.
   I-O Drüben rastet die Mieterin aus Linz, die der Metzger vor kurzem erkannte und begrüßte.
9. R-S Drinnen verweilt die Artistin aus Prag, die den Tischler vor kurzem erkannte und grüßte.

R-O Drinnen verweilt die Artistin aus Prag, die der Tischler vor kurzem erkannte und grüßte.
I-S Drinnen wartet die Malerin aus Brünn, die den Tischler vor kurzem sah und begrüßte.
I-O Drinnen wartet die Malerin aus Brünn, die der Tischler vor kurzem sah und begrüßte.

10. R-S Drinnen erzählt die Artistin aus Prag, die den Dänen vor kurzem erkannte und grüßte.
R-O Drinnen erzählt die Artistin aus Prag, die der Däne vor kurzem erkannte und grüßte.
I-S Drinnen redet die Malerin aus Brünn, die den Dänen vor kurzem sah und begrüßte.
I-O Drinnen redet die Malerin aus Brünn, die der Däne vor kurzem sah und begrüßte.

11. R-S Hinten spaziert die Artistin aus Prag, die den Fischer vor kurzem erkannte und grüßte.
R-O Hinten spaziert die Artistin aus Prag, die der Fischer vor kurzem erkannte und grüßte.
I-S Hinten wandert die Malerin aus Brünn, die den Fischer vor kurzem sah und begrüßte.
I-O Hinten wandert die Malerin aus Brünn, die der Fischer vor kurzem sah und begrüßte.

12. R-S Vorne entspannt die Artistin aus Prag, die den Schweden vor kurzem erkannte und grüßte.
R-O Vorne entspannt die Artistin aus Prag, die der Schwede vor kurzem erkannte und grüßte.
I-S Vorne rastet die Malerin aus Brünn, die den Schweden vor kurzem erkannte und begrüßte.
I-O Vorne rastet die Malerin aus Brünn, die der Schwede vor kurzem erkannte und begrüßte.

13. R-S Drinnen verweilt die Athletin aus Burg, die den Türken vor kurzem erkannte und grüßte.
R-O Drinnen verweilt die Athletin aus Burg, die der Türke vor kurzem erkannte und grüßte.
I-S Drinnen wartet die Trainerin aus Leer, die den Türken vor kurzem sah und begrüßte.
I-O Drinnen wartet die Trainerin aus Leer, die der Türke vor kurzem sah und begrüßte.

14. R-S Draußen erzählt die Athletin aus Burg, die den Bauern vor kurzem erkannte und grüßte.
R-O Draußen erzählt die Athletin aus Burg, die der Bauer vor kurzem erkannte und grüßte.
I-S Draußen redet die Trainerin aus Leer, die den Bauern vor kurzem erkannte und begrüßte.
I-O Draußen redet die Trainerin aus Leer, die der Bauer vor kurzem erkannte und begrüßte.

15. R-S Drüben spaziert die Athletin aus Burg, die den Metzger vor kurzem erkannte und grüßte.
R-O Drüben spaziert die Athletin aus Burg, die der Metzger vor kurzem erkannte und grüßte.
I-S Drüben wandert die Trainerin aus Leer, die den Metzger vor kurzem sah und begrüßte.
I-O Drüben wandert die Trainerin aus Leer, die der Metzger vor kurzem sah und begrüßte.

16. R-S Drüben entspannt die Athletin aus Burg, die den Schuster vor kurzem erkannte und grüßte.
R-O Drüben entspannt die Athletin aus Burg, die der Schuster vor kurzem erkannte und grüßte.
I-S Drüben rastet die Trainerin aus Leer, die den Schuster vor kurzem sah und begrüßte.
I-O Drüben rastet die Trainerin aus Leer, die der Schuster vor kurzem sah und begrüßte.

17. R-S Draußen verweilt die Baronin aus Rom, die den Dichter vor kurzem erkannte und grüßte.
R-O Draußen verweilt die Baronin aus Rom, die der Dichter vor kurzem erkannte und grüßte.
I-S Draußen wartet die Sängerin aus Lille, die den Dichter vor kurzem sah und begrüßte.
I-O Draußen wartet die Sängerin aus Lille, die der Dichter vor kurzem sah und begrüßte.

18. R-S Drüben erzählt die Baronin aus Rom, die den Schweden vor kurzem erkannte und grüßte.
R-O Drüben erzählt die Baronin aus Rom, die der Schwede vor kurzem erkannte und grüßte.

| | | |
|---|---|---|
| | I-S | Drüben redet die Sängerin aus Lille, die den Schweden vor kurzem erkannte und begrüßte. |
| | I-O | Drüben redet die Sängerin aus Lille, die der Schwede vor kurzem erkannte und begrüßte. |
| 19. | R-S | Vorne spaziert die Baronin aus Rom, die den Griechen vor kurzem erkannte und grüßte. |
| | R-O | Vorne spaziert die Baronin aus Rom, die der Grieche vor kurzem erkannte und grüßte. |
| | I-S | Vorne wandert die Sängerin aus Lille, die den Griechen vor kurzem erkannte und begrüßte. |
| | I-O | Vorne wandert die Sängerin aus Lille, die der Grieche vor kurzem erkannte und begrüßte. |
| 20. | R-S | Hinten entspannt die Baronin aus Rom, die den Dänen vor kurzem erkannte und grüßte. |
| | R-O | Hinten entspannt die Baronin aus Rom, die der Däne vor kurzem erkannte und grüßte. |
| | I-S | Hinten rastet die Sängerin aus Lille, die den Dänen vor kurzem sah und begrüßte. |
| | I-O | Hinten rastet die Sängerin aus Lille, die der Däne vor kurzem sah und begrüßte. |
| 21. | R-S | Draußen verweilt die Beamtin aus Bonn, die den Dichter vor kurzem erkannte und grüßte. |
| | R-O | Draußen verweilt die Beamtin aus Bonn, die der Dichter vor kurzem erkannte und grüßte. |
| | I-S | Draußen wartet die Richterin aus Mainz, die den Dichter vor kurzem sah und begrüßte. |
| | I-O | Draußen wartet die Richterin aus Mainz, die der Dichter vor kurzem sah und begrüßte. |
| 22. | R-S | Hinten erzählt die Beamtin aus Bonn, die den Jungen vor kurzem erkannte und grüßte. |
| | R-O | Hinten erzählt die Beamtin aus Bonn, die der Junge vor kurzem erkannte und grüßte. |
| | I-S | Hinten redet die Richterin aus Mainz, die den Jungen vor kurzem sah und begrüßte. |
| | I-O | Hinten redet die Richterin aus Mainz, die der Junge vor kurzem sah und begrüßte. |
| 23. | R-S | Hinten spaziert die Beamtin aus Bonn, die den Briten vor kurzem erkannte und grüßte. |
| | R-O | Hinten spaziert die Beamtin aus Bonn, die der Brite vor kurzem erkannte und grüßte. |
| | I-S | Hinten wandert die Richterin aus Mainz, die den Briten vor kurzem sah und begrüßte. |
| | I-O | Hinten wandert die Richterin aus Mainz, die der Brite vor kurzem sah und begrüßte. |
| 24. | R-S | Vorne entspannt die Beamtin aus Bonn, die den Fischer vor kurzem erkannte und grüßte. |
| | R-O | Vorne entspannt die Beamtin aus Bonn, die der Fischer vor kurzem erkannte und grüßte. |
| | I-S | Vorne rastet die Richterin aus Mainz, die den Fischer vor kurzem erkannte und begrüßte. |
| | I-O | Vorne rastet die Richterin aus Mainz, die der Fischer vor kurzem erkannte und begrüßte. |
| 25. | R-S | Hinten verweilt die Floristin aus Köln, die den Pfarrer vor kurzem erkannte und grüßte. |
| | R-O | Hinten verweilt die Floristin aus Köln, die der Pfarrer vor kurzem erkannte und grüßte. |
| | I-S | Hinten wartet die Gärtnerin aus Hamm, die den Pfarrer vor kurzem erkannte und begrüßte. |
| | I-O | Hinten wartet die Gärtnerin aus Hamm, die der Pfarrer vor kurzem erkannte und begrüßte. |
| 26. | R-S | Hinten erzählt die Floristin aus Köln, die den Bauern vor kurzem erkannte und grüßte. |
| | R-O | Hinten erzählt die Floristin aus Köln, die der Bauer vor kurzem erkannte und grüßte. |
| | I-S | Hinten redet die Gärtnerin aus Hamm, die den Bauern vor kurzem sah und begrüßte. |
| | I-O | Hinten redet die Gärtnerin aus Hamm, die der Bauer vor kurzem sah und begrüßte. |
| 27. | R-S | Drüben spaziert die Floristin aus Köln, die den Tischler vor kurzem erkannte und grüßte. |
| | R-O | Drüben spaziert die Floristin aus Köln, die der Tischler vor kurzem erkannte und grüßte. |
| | I-S | Drüben wandert die Gärtnerin aus Hamm, die den Tischler vor kurzem erkannte und begrüßte. |

| | | |
|---|---|---|
| | I-O | Drüben wandert die Gärtnerin aus Hamm, die der Tischler vor kurzem erkannte und begrüßte. |
| 28. | R-S | Drinnen entspannt die Floristin aus Köln, die den Türken vor kurzem erkannte und grüßte. |
| | R-O | Drinnen entspannt die Floristin aus Köln, die der Türke vor kurzem erkannte und grüßte. |
| | I-S | Drinnen rastet die Gärtnerin aus Hamm, die den Türken vor kurzem sah und begrüßte. |
| | I-O | Drinnen rastet die Gärtnerin aus Hamm, die der Türke vor kurzem sah und begrüßte. |
| 29. | R-S | Drüben verweilt die Französin aus Brest, die den Schweden vor kurzem erkannte und grüßte. |
| | R-O | Drüben verweilt die Französin aus Brest, die der Schwede vor kurzem erkannte und grüßte. |
| | I-S | Drüben wartet die Nigrerin aus Say, die den Schweden vor kurzem sah und begrüßte. |
| | I-O | Drüben wartet die Nigrerin aus Say, die der Schwede vor kurzem sah und begrüßte. |
| 30. | R-S | Drüben erzählt die Französin aus Brest, die den Bäcker vor kurzem erkannte und grüßte. |
| | R-O | Drüben erzählt die Französin aus Brest, die der Bäcker vor kurzem erkannte und grüßte. |
| | I-S | Drüben redet die Nigrerin aus Say, die den Bäcker vor kurzem sah und begrüßte. |
| | I-O | Drüben redet die Nigrerin aus Say, die der Bäcker vor kurzem sah und begrüßte. |
| 31. | R-S | Vorne spaziert die Französin aus Brest, die den Pfarrer vor kurzem erkannte und grüßte. |
| | R-O | Vorne spaziert die Französin aus Brest, die der Pfarrer vor kurzem erkannte und grüßte. |
| | I-S | Vorne wandert die Nigrerin aus Say, die den Pfarrer vor kurzem erkannte und begrüßte. |
| | I-O | Vorne wandert die Nigrerin aus Say, die der Pfarrer vor kurzem erkannte und begrüßte. |
| 32. | R-S | Drüben entspannt die Französin aus Brest, die den Griechen vor kurzem erkannte und grüßte. |
| | R-O | Drüben entspannt die Französin aus Brest, die der Grieche vor kurzem erkannte und grüßte. |
| | I-S | Drüben rastet die Nigrerin aus Say, die den Griechen vor kurzem sah und begrüßte. |
| | I-O | Drüben rastet die Nigrerin aus Say, die der Grieche vor kurzem sah und begrüßte. |
| 33. | R-S | Drinnen verweilt die Friseurin aus Hof, die den Fischer vor kurzem erkannte und grüßte. |
| | R-O | Drinnen verweilt die Friseurin aus Hof, die der Fischer vor kurzem erkannte und grüßte. |
| | I-S | Drinnen wartet die Kellnerin aus Fürth, die den Fischer vor kurzem erkannte und begrüßte. |
| | I-O | Drinnen wartet die Kellnerin aus Fürth, die der Fischer vor kurzem erkannte und begrüßte. |
| 34. | R-S | Vorne erzählt die Friseurin aus Hof, die den Schuster vor kurzem erkannte und grüßte. |
| | R-O | Vorne erzählt die Friseurin aus Hof, die der Schuster vor kurzem erkannte und grüßte. |
| | I-S | Vorne redet die Kellnerin aus Fürth, die den Schuster vor kurzem erkannte und begrüßte. |
| | I-O | Vorne redet die Kellnerin aus Fürth, die der Schuster vor kurzem erkannte und begrüßte. |
| 35. | R-S | Hinten spaziert die Friseurin aus Hof, die den Türken vor kurzem erkannte und grüßte. |
| | R-O | Hinten spaziert die Friseurin aus Hof, die der Türke vor kurzem erkannte und grüßte. |
| | I-S | Hinten wandert die Kellnerin aus Fürth, die den Türken vor kurzem sah und begrüßte. |
| | I-O | Hinten wandert die Kellnerin aus Fürth, die der Türke vor kurzem sah und begrüßte. |
| 36. | R-S | Hinten entspannt die Friseurin aus Hof, die den Jungen vor kurzem erkannte und grüßte. |
| | R-O | Hinten entspannt die Friseurin aus Hof, die der Junge vor kurzem erkannte und grüßte. |
| | I-S | Hinten rastet die Kellnerin aus Fürth, die den Jungen vor kurzem sah und begrüßte. |
| | I-O | Hinten rastet die Kellnerin aus Fürth, die der Junge vor kurzem sah und begrüßte. |

37. R-S Drüben verweilt die Kollegin aus Kiel, die den Bauern vor kurzem erkannte und grüßte.
R-O Drüben verweilt die Kollegin aus Kiel, die der Bauer vor kurzem erkannte und grüßte.
I-S Drüben wartet die Rentnerin aus Burg, die den Bauern vor kurzem erkannte und begrüßte.
I-O Drüben wartet die Rentnerin aus Burg, die der Bauer vor kurzem erkannte und begrüßte.
38. R-S Vorne erzählt die Kollegin aus Kiel, die den Metzger vor kurzem erkannte und grüßte.
R-O Vorne erzählt die Kollegin aus Kiel, die der Metzger vor kurzem erkannte und grüßte.
I-S Vorne redet die Rentnerin aus Burg, die den Metzger vor kurzem sah und begrüßte.
I-O Vorne redet die Rentnerin aus Burg, die der Metzger vor kurzem sah und begrüßte.
39. R-S Vorne spaziert die Kollegin aus Kiel, die den Griechen vor kurzem erkannte und grüßte.
R-O Vorne spaziert die Kollegin aus Kiel, die der Grieche vor kurzem erkannte und grüßte.
I-S Vorne wandert die Rentnerin aus Burg, die den Griechen vor kurzem sah und begrüßte.
I-O Vorne wandert die Rentnerin aus Burg, die der Grieche vor kurzem sah und begrüßte.
40. R-S Draußen entspannt die Kollegin aus Kiel, die den Pfarrer vor kurzem erkannte und grüßte.
R-O Draußen entspannt die Kollegin aus Kiel, die der Pfarrer vor kurzem erkannte und grüßte.
I-S Draußen rastet die Rentnerin aus Burg, die den Pfarrer vor kurzem sah und begrüßte.
I-O Draußen rastet die Rentnerin aus Burg, die der Pfarrer vor kurzem sah und begrüßte.
41. R-S Drinnen verweilt die Kroatin aus Split, die den Dichter vor kurzem erkannte und grüßte.
R-O Drinnen verweilt die Kroatin aus Split, die der Dichter vor kurzem erkannte und grüßte.
I-S Drinnen wartet die Spanierin aus Elx, die den Dichter vor kurzem erkannte und begrüßte.
I-O Drinnen wartet die Spanierin aus Elx, die der Dichter vor kurzem erkannte und begrüßte.
42. R-S Hinten erzählt die Kroatin aus Split, die den Schweden vor kurzem erkannte und grüßte.
R-O Hinten erzählt die Kroatin aus Split, die der Schwede vor kurzem erkannte und grüßte.
I-S Hinten redet die Spanierin aus Elx, die den Schweden vor kurzem erkannte und begrüßte.
I-O Hinten redet die Spanierin aus Elx, die der Schwede vor kurzem erkannte und begrüßte.
43. R-S Draußen spaziert die Kroatin aus Split, die den Händler vor kurzem erkannte und grüßte.
R-O Draußen spaziert die Kroatin aus Split, die der Händler vor kurzem erkannte und grüßte.
I-S Draußen wandert die Spanierin aus Elx, die den Händler vor kurzem sah und begrüßte.
I-O Draußen wandert die Spanierin aus Elx, die der Händler vor kurzem sah und begrüßte.
44. R-S Hinten entspannt die Kroatin aus Split, die den Metzger vor kurzem erkannte und grüßte.
R-O Hinten entspannt die Kroatin aus Split, die der Metzger vor kurzem erkannte und grüßte.
I-S Hinten rastet die Spanierin aus Elx, die den Metzger vor kurzem erkannte und begrüßte.
I-O Hinten rastet die Spanierin aus Elx, die der Metzger vor kurzem erkannte und begrüßte.
45. R-S Drinnen verweilt die Rumänin aus Dej, die den Händler vor kurzem erkannte und grüßte.
R-O Drinnen verweilt die Rumänin aus Dej, die der Händler vor kurzem erkannte und grüßte.
I-S Drinnen wartet die Schweizerin aus Genf, die den Händler vor kurzem erkannte und begrüßte.
I-O Drinnen wartet die Schweizerin aus Genf, die der Händler vor kurzem erkannte und begrüßte.
46. R-S Drüben erzählt die Rumänin aus Dej, die den Tischler vor kurzem erkannte und grüßte.

R-O Drüben erzählt die Rumänin aus Dej, die der Tischler vor kurzem erkannte und grüßte.
I-S Drüben redet die Schweizerin aus Genf, die den Tischler vor kurzem erkannte und begrüßte.
I-O Drüben redet die Schweizerin aus Genf, die der Tischler vor kurzem erkannte und begrüßte.
47. R-S Draußen spaziert die Rumänin aus Dej, die den Dichter vor kurzem erkannte und grüßte.
R-O Draußen spaziert die Rumänin aus Dej, die der Dichter vor kurzem erkannte und grüßte.
I-S Draußen wandert die Schweizerin aus Genf, die den Dichter vor kurzem erkannte und begrüßte.
I-O Draußen wandert die Schweizerin aus Genf, die der Dichter vor kurzem erkannte und begrüßte.
48. R-S Vorne entspannt die Rumänin aus Dej, die den Schuster vor kurzem erkannte und grüßte.
R-O Vorne entspannt die Rumänin aus Dej, die der Schuster vor kurzem erkannte und grüßte.
I-S Vorne rastet die Schweizerin aus Genf, die den Schuster vor kurzem erkannte und begrüßte.
I-O Vorne rastet die Schweizerin aus Genf, die der Schuster vor kurzem erkannte und begrüßte.
49. R-S Drinnen verweilt die Slowakin aus Svit, die den Händler vor kurzem erkannte und grüßte.
R-O Drinnen verweilt die Slowakin aus Svit, die der Händler vor kurzem erkannte und grüßte.
I-S Drinnen wartet die Ungarin aus Pecs, die den Händler vor kurzem sah und begrüßte.
I-O Drinnen wartet die Ungarin aus Pecs, die der Händler vor kurzem sah und begrüßte.
50. R-S Draußen erzählt die Slowakin aus Svit, die den Dänen vor kurzem erkannte und grüßte.
R-O Draußen erzählt die Slowakin aus Svit, die der Däne vor kurzem erkannte und grüßte.
I-S Draußen redet die Ungarin aus Pecs, die den Dänen vor kurzem sah und begrüßte.
I-O Draußen redet die Ungarin aus Pecs, die der Däne vor kurzem sah und begrüßte.
51. R-S Vorne spaziert die Slowakin aus Svit, die den Bäcker vor kurzem erkannte und grüßte.
R-O Vorne spaziert die Slowakin aus Svit, die der Bäcker vor kurzem erkannte und grüßte.
I-S Vorne wandert die Ungarin aus Pecs, die den Bäcker vor kurzem erkannte und begrüßte.
I-O Vorne wandert die Ungarin aus Pecs, die der Bäcker vor kurzem erkannte und begrüßte.
52. R-S Hinten entspannt die Slowakin aus Svit, die den Türken vor kurzem erkannte und grüßte.
R-O Hinten entspannt die Slowakin aus Svit, die der Türke vor kurzem erkannte und grüßte.
I-S Hinten rastet die Ungarin aus Pecs, die den Türken vor kurzem sah und begrüßte.
I-O Hinten rastet die Ungarin aus Pecs, die der Türke vor kurzem sah und begrüßte.
53. R-S Drinnen verweilt die Statistin aus Minsk, die den Briten vor kurzem erkannte und grüßte.
R-O Drinnen verweilt die Statistin aus Minsk, die der Brite vor kurzem erkannte und grüßte.
I-S Drinnen wartet die Bankerin aus Perm, die den Briten vor kurzem erkannte und begrüßte.
I-O Drinnen wartet die Bankerin aus Perm, die der Brite vor kurzem erkannte und begrüßte.
54. R-S Vorne erzählt die Statistin aus Minsk, die den Griechen vor kurzem erkannte und grüßte.
R-O Vorne erzählt die Statistin aus Minsk, die der Grieche vor kurzem erkannte und grüßte.
I-S Vorne redet die Bankerin aus Perm, die den Griechen vor kurzem erkannte und begrüßte.
I-O Vorne redet die Bankerin aus Perm, die der Grieche vor kurzem erkannte und begrüßte.
55. R-S Vorne spaziert die Statistin aus Minsk, die den Pfarrer vor kurzem erkannte und grüßte.
R-O Vorne spaziert die Statistin aus Minsk, die der Pfarrer vor kurzem erkannte und grüßte.

I-S Vorne wandert die Bankerin aus Perm, die den Pfarrer vor kurzem erkannte und begrüßte.
I-O Vorne wandert die Bankerin aus Perm, die der Pfarrer vor kurzem erkannte und begrüßte.

56. R-S Draußen entspannt die Statistin aus Minsk, die den Bäcker vor kurzem erkannte und grüßte.
R-O Draußen entspannt die Statistin aus Minsk, die der Bäcker vor kurzem erkannte und grüßte.
I-S Draußen rastet die Bankerin aus Perm, die den Bäcker vor kurzem erkannte und begrüßte.
I-O Draußen rastet die Bankerin aus Perm, die der Bäcker vor kurzem erkannte und begrüßte.

57. R-S Drüben verweilt die Studentin aus Trier, die den Jungen vor kurzem erkannte und grüßte.
R-O Drüben verweilt die Studentin aus Trier, die der Junge vor kurzem erkannte und grüßte.
I-S Drüben wartet die Schülerin aus Ulm, die den Jungen vor kurzem erkannte und begrüßte.
I-O Drüben wartet die Schülerin aus Ulm, die der Junge vor kurzem erkannte und begrüßte.

58. R-S Drüben erzählt die Studentin aus Trier, die den Bäcker vor kurzem erkannte und grüßte.
R-O Drüben erzählt die Studentin aus Trier, die der Bäcker vor kurzem erkannte und grüßte.
I-S Drüben redet die Schülerin aus Ulm, die den Bäcker vor kurzem sah und begrüßte.
I-O Drüben redet die Schülerin aus Ulm, die der Bäcker vor kurzem sah und begrüßte.

59. R-S Draußen spaziert die Studentin aus Trier, die den Tischler vor kurzem erkannte und grüßte.
R-O Draußen spaziert die Studentin aus Trier, die der Tischler vor kurzem erkannte und grüßte.
I-S Draußen wandert die Schülerin aus Ulm, die den Tischler vor kurzem erkannte und begrüßte.
I-O Draußen wandert die Schülerin aus Ulm, die der Tischler vor kurzem erkannte und begrüßte.

60. R-S Draußen entspannt die Studentin aus Trier, die den Briten vor kurzem erkannte und grüßte.
R-O Draußen entspannt die Studentin aus Trier, die der Brite vor kurzem erkannte und grüßte.
I-S Draußen rastet die Schülerin aus Ulm, die den Briten vor kurzem erkannte und begrüßte.
I-O Draußen rastet die Schülerin aus Ulm, die der Brite vor kurzem erkannte und begrüßte.

# Appendix C

# Additional participant information

*Table C.1: Additional participant information. Experiments in parentheses indicate that data was collected but excluded to due experimental criteria. VWM = Verbal Working Memory. TPD = Time and Pitch Discrimination.*

| Participant | Experiment | Sex | Age | Musicianship | Age at musical training begin | Musical training total years | Practice hours per week | VWM | TPD |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1a, 1b | female | 20 | musician | 7 | 13 | 1 | -0.19 | 0.07 |
| 2 | 1a, 1b | male | 21 | musician | 7 | 14 | 14 | -0.73 | 0.63 |
| 3 | 1a (1b) | female | 23 | musician | 10 | 13 | 2 | -0.83 | 0.60 |
| 4 | 1a, 1b | male | 21 | musician | 3 | 18 | 10 | 0.13 | -0.13 |
| 5 | 1a, 1b | female | 21 | musician | 5 | 16 | 5 | 0.20 | -0.01 |
| 6 | 1a, 1b | male | 21 | musician | 6 | 15 | 4 | 1.15 | 0.65 |

Continued on next page

Table C.1 – continued from previous page

| Participant | Experiment | Sex | Age | Musicianship | Age at musical training begin | Musical training total years | Practice hours per week | VWM | TPD |
|---|---|---|---|---|---|---|---|---|---|
| 7 | 1a, 1b | female | 26 | non-musician | – | – | – | 1.06 | 0.10 |
| 8 | 1a, 1b | male | 21 | musician | 11 | 10 | 8 | 0.26 | 0.26 |
| 9 | 1a, 1b | female | 20 | musician | 5 | 15 | 1,5 | -0.34 | -1.28 |
| 10 | 1a, 1b | male | 21 | musician | 7 | 14 | 6 | 1.29 | 1.44 |
| 11 | 1a (1b) | female | 22 | musician | 7 | 15 | 6 | -0.45 | -0.40 |
| 12 | 1a, 1b | male | 27 | non-musician | – | – | – | 0.23 | -0.96 |
| 13 | 1a (1b) | female | 20 | musician | 6 | 14 | 5 | 1.58 | 0.40 |
| 14 | 1a, 1b | male | 22 | musician | 9 | 13 | 3 | 1.44 | 0.19 |
| 15 | 1a, 1b | female | 21 | musician | 7 | 14 | 4 | -0.36 | -0.42 |
| 16 | 1a, 1b | male | 24 | non-musician | – | – | – | -1.50 | -1.95 |
| 17 | 1a (1b) | female | 24 | non-musician | – | – | – | -0.52 | -1.32 |
| 18 | 1a, 1b | male | 27 | non-musician | – | – | – | 1.29 | -0.85 |
| 19 | 1a, 1b | female | 29 | non-musician | – | – | – | 0.54 | -0.17 |
| 20 | 1a, 1b | male | 26 | non-musician | – | – | – | 0.29 | -0.16 |
| 21 | 1a, 1b | male | 30 | non-musician | – | .. | – | -0.84 | -0.56 |
| 22 | 1a (1b) | female | 31 | non-musician | – | – | – | -0.27 | -2.31 |
| 23 | 1a (1b) | male | 28 | non-musician | – | – | – | -1.43 | -0.20 |
| 24 | 1a, 1b | female | 30 | non-musician | – | – | – | -1.20 | -1.60 |
| 25 | 1a, 1b | male | 22 | musician | 11 | 11 | 10 | 1.69 | 1.62 |

Continued on next page

Table C.1 – continued from previous page

| Participant | Experiment | Sex | Age | Musicianship | Age at musical training begin | Musical training total years | Practice hours per week | VWM | TPD |
|---|---|---|---|---|---|---|---|---|---|
| 26 | 1a, 1b | female | 23 | non-musician | – | – | – | -1.11 | -1.82 |
| 27 | 1a, 1b | male | 25 | non-musician | – | – | – | 1.52 | -1.37 |
| 28 | 1a, 1b | female | 20 | musician | 6 | 14 | 12 | -0.79 | 1.24 |
| 29 | 1a, 1b | male | 22 | non-musician | – | – | – | -0.10 | -0.27 |
| 30 | 1a, 2b, 3a, 3c | male | 27 | musician | 24 | 3 | 2,5 | 1.19 | -0.28 |
| 31 | 1a, 2b, 3a, 3c | male | 24 | musician | 16 | 8 | 3,5 | -0.37 | 1.18 |
| 32 | 1a, 2b, 3a, 3c | male | 27 | musician | 23 | 4 | 6 | -0.87 | 0.67 |
| 33 | 1a, 2b, 3a, 3c | male | 28 | musician | 18 | 10 | 1 | -1.38 | -1.13 |
| 34 | 1a (2b, 3a, 3c) | male | 27 | musician | 18 | 9 | 4 | 0.16 | 1.15 |
| 35 | 1a, 2b, 3a, 3c | female | 25 | musician | 20 | 5 | 2 | 1.37 | -0.24 |
| 36 | 1a, 2b, 3a, 3c | female | 24 | musician | 12 | 12 | 2 | -1.18 | -1.09 |
| 37 | 1a, 2b, 3a, 3c | female | 26 | musician | 11 | 15 | 4 | 1.75 | 0.33 |
| 38 | 1a, 2b, 3a, 3c | female | 25 | musician | 12 | 13 | 2 | -1.60 | -1.24 |
| 39 | 1a, 2b, 3a, 3c | male | 25 | musician | 16 | 9 | 10,5 | -0.20 | 0.79 |
| 40 | 1a, 2b, 3a, 3c | female | 26 | musician | 8 | 18 | 3,5 | 0.72 | -0.96 |
| 41 | 1a, 2b, 3a, 3c | female | 25 | musician | 16 | 9 | 2 | -0.85 | 1.60 |
| 42 | 1a (2b, 3a, 3c) | male | 25 | musician | 18 | 7 | 10 | 0.16 | -1.24 |
| 43 | 1a, 2b, 3a, 3c | female | 22 | musician | 14 | 8 | 1,5 | 0.16 | -0.28 |
| 44 | 1a, 2b, 3a, 3c | female | 21 | musician | 8 | 13 | 1,5 | 1.09 | 1.76 |

Continued on next page

Table C.1 – continued from previous page

| Participant | Experiment | Sex | Age | Musicianship | Age at musical training begin | Musical training total years | Practice hours per week | VWM | TPD |
|---|---|---|---|---|---|---|---|---|---|
| 45 | 1a, 2b, 3a, 3c | male | 26 | musician | 18 | 8 | 5 | -0.56 | 0.35 |
| 46 | 1a (2b, 3a, 3c) | female | 27 | musician | 7 | 20 | 2 | 0.44 | 0.17 |
| 47 | 1a, 2b, 3a, 3c | male | 22 | musician | 10 | 12 | 5 | -1.86 | 1.00 |
| 48 | 1a, 2b, 3a, 3c | male | 28 | musician | 11 | 17 | 3,75 | 0.10 | 1.36 |
| 49 | 1a, 2b, 3a, 3c | female | 21 | musician | 9 | 12 | 22,5 | 0.71 | 1.21 |
| 50 | 1a, 2b, 3a, 3c | female | 19 | musician | 11 | 8 | 2 | 0.91 | 0.48 |
| 51 | 1a, 2b, 3a, 3c | female | 28 | musician | 13 | 15 | 2 | -1.01 | -0.52 |
| 52 | 1a, 2b, 3a, 3c | male | 22 | musician | 9 | 13 | 15 | 0.72 | 1.23 |
| 53 | 1a (2b, 3a, 3c) | male | 24 | musician | 18 | 6 | 12,5 | -0.25 | 0.99 |
| 54 | 1a (2b, 3a, 3c) | female | 26 | musician | 10 | 16 | 5,5 | -0.20 | 0.05 |
| 55 | 1a, 2b, 3a, 3c | female | 21 | musician | 16 | 5 | 4,5 | 0.23 | 1.18 |
| 56 | 1a, 2b, 3a, 3c | male | 25 | musician | 8 | 17 | 6 | -0.43 | 0.06 |
| 57 | 1a, 2b, 3a, 3c | male | 27 | musician | 14 | 13 | 7 | 0.12 | -0.21 |
| 58 | 1a, 3b | male | 28 | non-musician | – | – | – | 0.61 | 0.67 |
| 59 | 1a, 3b | male | 26 | musician | 13 | 13 | 2 | 1.33 | 1.62 |
| 60 | 1a, 3b | male | 29 | non-musician | – | – | – | -2.36 | -0.18 |
| 61 | 1a, 3b | male | 23 | non-musician | – | – | – | -1.07 | 0.16 |
| 62 | 1a, 3b | female | 25 | musician | 12 | 13 | 10 | -1.22 | 0.30 |
| 63 | 1a, 3b | male | 25 | non-musician | – | – | – | 0.91 | 0.46 |

Continued on next page

Table C.1 – continued from previous page

| Participant | Experiment | Sex | Age | Musicianship | Age at musical training begin | Musical training total years | Practice hours per week | VWM | TPD |
|---|---|---|---|---|---|---|---|---|---|
| 64 | 1a, 3b | male | 24 | non-musician | – | – | – | 0.18 | -0.59 |
| 65 | 1a, 3b | female | 21 | musician | 6 | 15 | 3 | -0.87 | 1.14 |
| 66 | 1a, 3b | female | 23 | musician | 14 | 9 | 10 | 0.83 | 1.39 |
| 67 | 1a, 3b | female | 25 | musician | 10 | 15 | 5 | 1.02 | 0.10 |
| 68 | 1a, 3b | female | 22 | musician | 11 | 11 | 5 | -0.49 | -0.22 |
| 69 | 1a, 3b | female | 21 | musician | 6 | 15 | 2 | -2.32 | -0.31 |
| 70 | 1a, 3b | female | 22 | non-musician | – | – | – | 0.34 | -0.05 |
| 71 | 1a, 3b | female | 25 | musician | 7 | 18 | 3 | 0.49 | 0.96 |
| 72 | 1a, 3b | female | 28 | non-musician | – | – | – | 0.66 | -1.00 |
| 73 | 1a, 3b | female | 25 | non-musician | – | – | – | -0.30 | -0.46 |
| 74 | 1a, 3b | male | 25 | non-musician | – | – | – | -0.96 | -1.62 |
| 75 | 1a, 3b | male | 34 | non-musician | – | – | – | 0.36 | 1.20 |
| 76 | 1a, 3b | female | 21 | musician | 7 | 14 | 6 | -1.40 | 1.94 |
| 77 | 1a, 3b | male | 23 | non-musician | – | – | – | 0.35 | -1.41 |
| 78 | 1a, 3b | female | 23 | musician | 10 | 13 | 3 | -1.38 | 0.66 |
| 79 | 1a, 3b | female | 27 | non-musician | – | – | – | 0.06 | -0.92 |
| 80 | 1a, 3b | male | 22 | musician | 8 | 14 | 1,5 | 2.02 | -0.21 |
| 81 | 1a, 3b | female | 26 | non-musician | – | – | – | -1.28 | 0.10 |
| 82 | 1a, 3b | female | 31 | non-musician | – | – | – | 1.53 | 0.28 |

Continued on next page

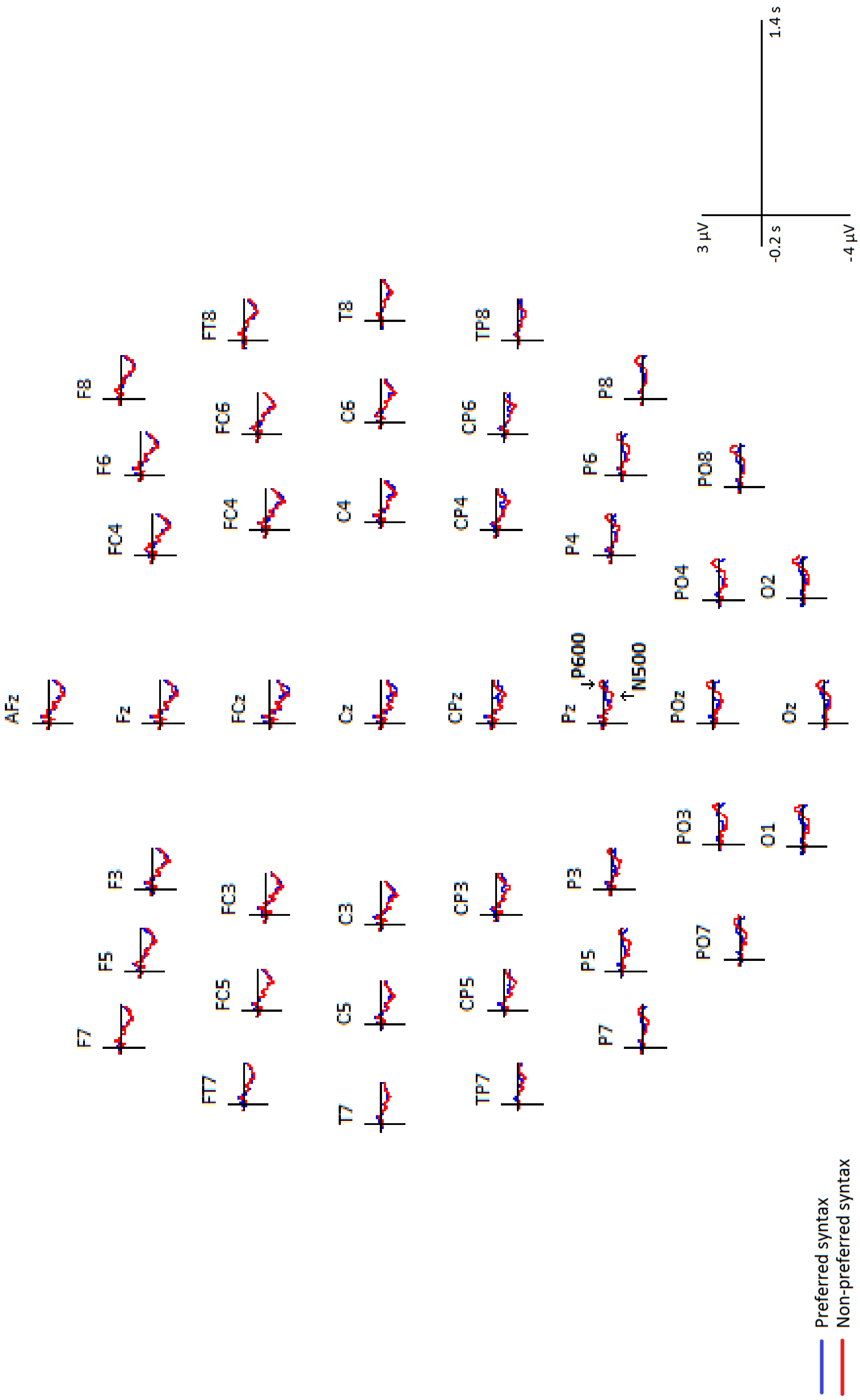
Table C.1 – continued from previous page

| Participant | Experiment | Sex | Age | Musicianship | Age at musical training begin | Musical training total years | Practice hours per week | VWM | TPD |
|---|---|---|---|---|---|---|---|---|---|
| 83 | 1a, 3b | female | 24 | non-musician | – | – | – | -0.26 | -0.38 |
| 84 | 1a, 3b | male | 24 | non-musician | – | – | – | 1.23 | 0.43 |
| 85 | 1a, 3b | male | 26 | non-musician | – | – | – | 0.93 | -1.21 |
| 86 | 1a, 3b | male | 29 | non-musician | – | – | – | -0.02 | -2.67 |
| 87 | (1a, 1b) | female | 22 | non-musician | – | – | – | | |
| 88 | (1a, 3b) | male | 29 | non-musician | – | – | – | | |

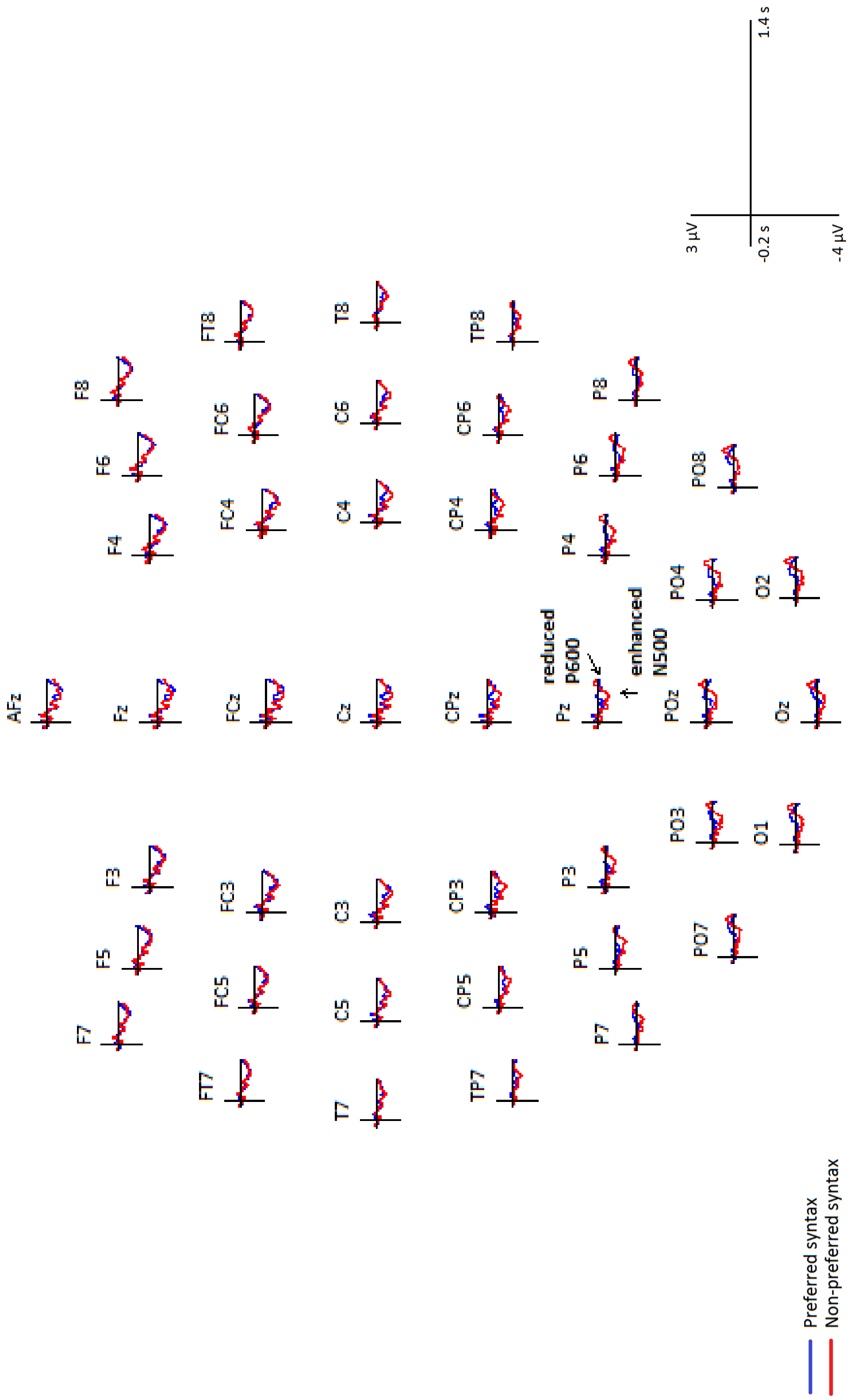
## Appendix D

# ERP headplots

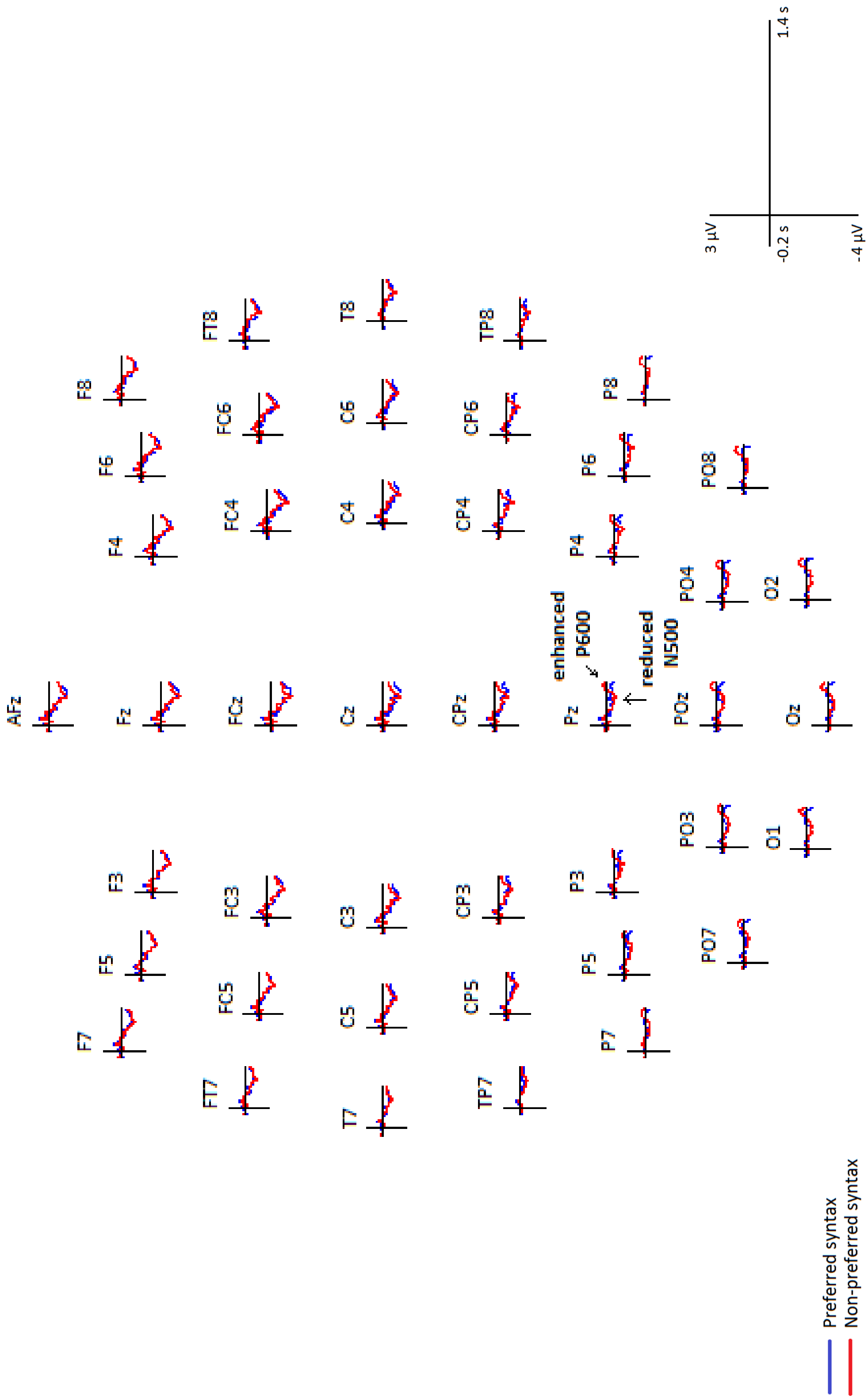
## D.1 Music main syntax comparison



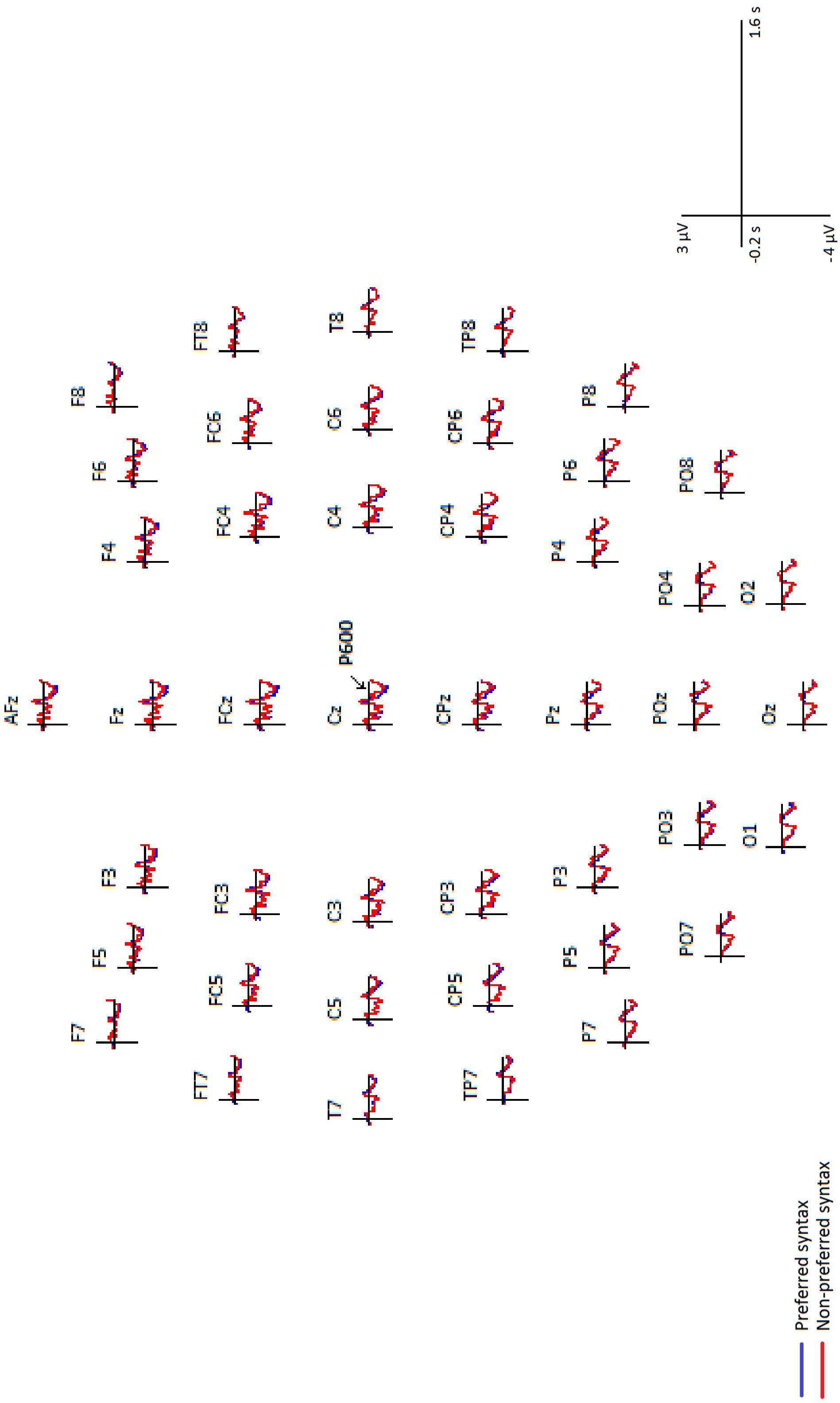
## D.2 Music syntax, regular metricality



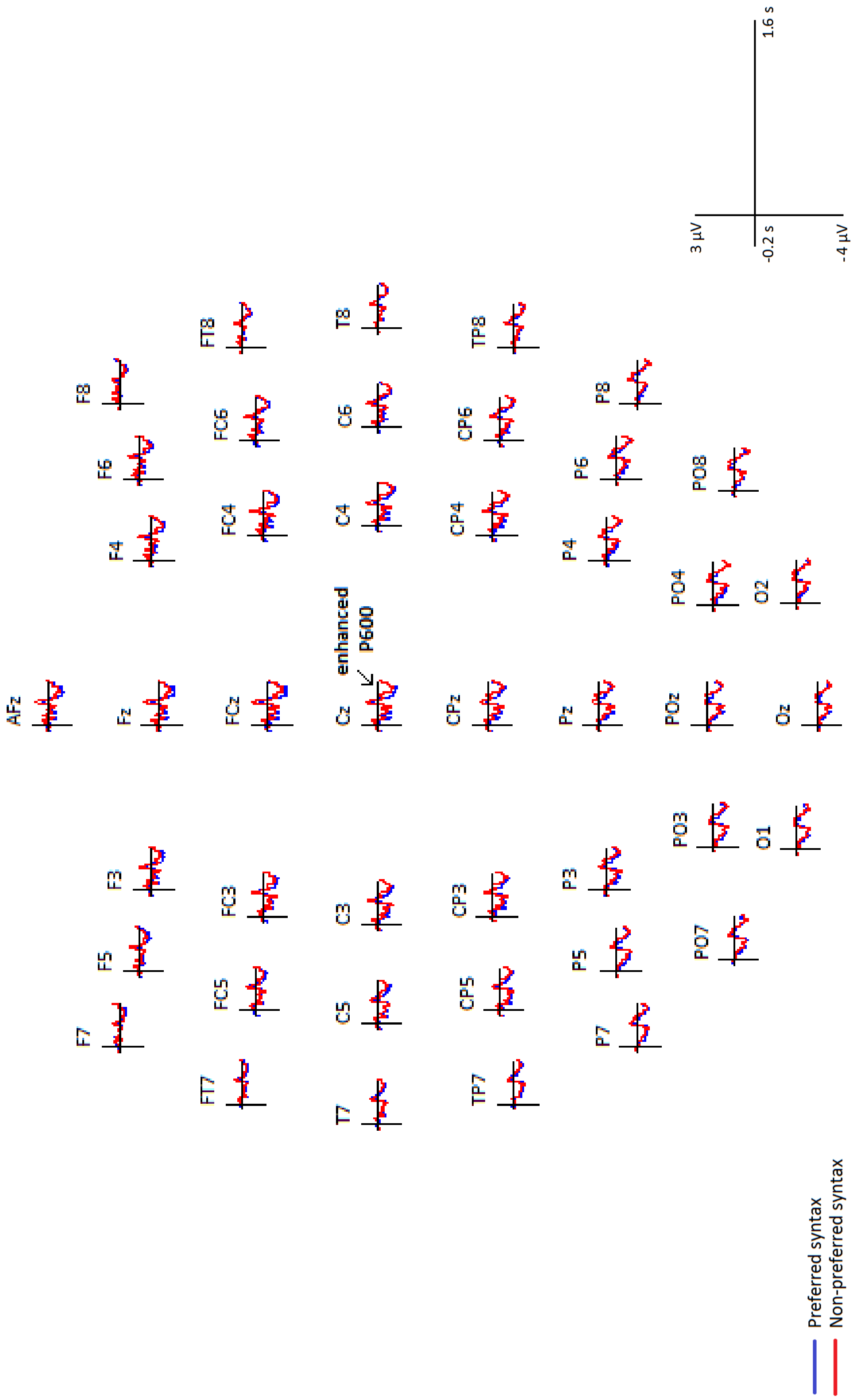
## D.3 Music syntax, irregular metricality



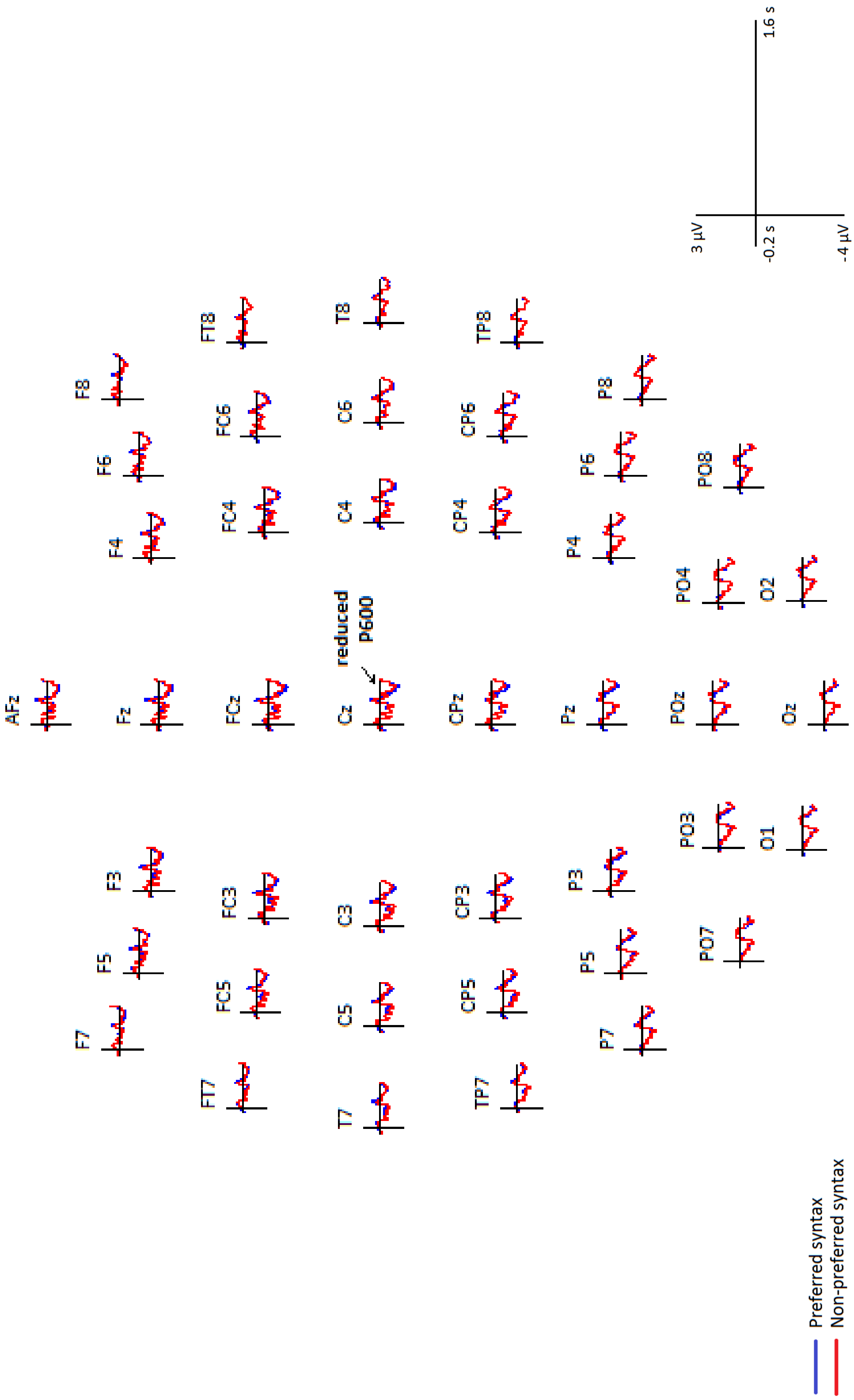
## D.4 Language syntax, main comparison



## D.5 Language syntax, regular metricality



## D.6 Language syntax, irregular metricality



# References


Abercrombie, D. (2006). *Elements of general phonetics* (Vol. 203). Edinburgh: Edinburgh University Press.

Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex*, *21*(7), 1507–1518. doi: 10.1093/cercor/bhq198

Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency following responses to vowel sounds. *Hearing Research*, *245*, 35–47. doi: 10.1016/j.heares.2008.08.004

Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., ... Gazzaley, A. (2013). Video game training enhances cognitive control in older adults. *Nature*, *501*(7465), 97–101. doi: 10.1038/nature12486

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, *83*(2), 111–130. doi: 10.1016/S0022-0965(02)00124-8

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica*, *66*(1–2), 46–63. doi: 10.1159/000208930

Astheimer, L. B., & Sanders, L. D. (2009). Listeners modulate temporally selective attention during natural speech processing. *Hearing Research*, *80*(1), 23–34. doi: 10.1016/j.biopsycho.2008.01.015

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic Bulletin & Review*, *22*(2), 366–377. doi: 10.3758/s13423-014-0699-x

Augurzky, P. (2006). *Attaching relative clauses in german: The role of implicit and explicit prosody in sentence processing* (Unpublished doctoral dissertation). MPI Series in Human Cognitive and Brain Sciences, 77, Leipzig.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. doi: 10.1016/S1364-6613(00)01538-2

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), R136–R140. doi: 10.1016/j.cub.2009.12.014

Baddeley, A., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89. doi: 10.1016/S0079-7421(08)60452-1

Baddeley, A., & Wilson, B. (1988). Comprehension and working memory: A single case neuropsychological study. *Journal of Memory and Language*, *27*, 479–498. doi: 10.1016/0749-596X(88)90021-6

Baker, M. C. (2001). Syntax. In M. Aronoff & J. Rees-Miller (Eds.), *The handbook of linguistics* (pp. 265–294). Maiden, MA.: Blackwell Publishers.

Bastiaansen, M., Mazaheri, A., & Jensen, O. (2012). Beyond erps: Oscillatory neuronal dynamics. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford handbook of event-related potential components* (pp. 31–49). Oxford, UK: Oxford University Press.

Böcker, K. B., Bastiaansen, M., Vroomen, J., Brunia, C. H., & Gelder, B. (1999). An erp correlate of metrical stress in spoken word recognition. *Cognitive Brain Research*, *36*(6), 706–720. doi: 10.1111/1469-8986.3660706

Berger, H. (1929). Über das elektroenkephalogramm des menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, *87*, 527–570.

Besson, M., & Faïta, F. (1995). An event-related potential (erp) study of musical expectancy: Comparisons of musicians with non-musicians. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(6), 1278–1296. doi: 10.1037/0096-1523.21.6.1278

Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other non-linguiustic contexts. *Psychophysiology*, *24*(1), 14.

Bever, T. G., & Chiarello, R. J. (1974). Cerebral dominance in musicians and nonmusicians. *Science*, *185*(4150), 537–539. doi: 10.1126/science.185.4150.537

Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research*, *1355*, 112–125. doi: 10.1016/j.brainres.2010.07.100

Bigand, E., Delbé, C., Poulin-Charronnat, B., Leman, M., & Tillmann, B. (2014). Empirical evidence for musical syntax processing? computer simulations reveal the contribution of auditory short-term memory. *Frontiers in Systems Neuroscience*, *8*(94). doi: 10.3389/fnsys.2014.00094

Bigand, E., Madurell, F., Tillmann, B., & Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 184–197. doi: 10.1037/0096-1523.25.1.184

Bigand, E., & Poulin-Charronnat, B. (2006). Are we 'experienced listeners'? a review of the musical capacities that do not depend on formal musical training. *Cognition*, *100*(1), 100–130. doi: 10.1016/j.cognition.2005.11.007

Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., & Ward, B. D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nature Neuroscience*, *7*(3), 295–301. doi: 10.1038/nn1198

Block, R. A. (2003). Psychological timing without a timer: The roles of attention and memory. In H. Helfrich (Ed.), *Time and mind ii* (pp. 41–60). Göttingen: Hogrefe & Huber.

Bod, R. (2002). Music and artificial intelligence: Second international conference, icmai 2002 edinburgh, scotland, uk, september 12–14, 2002 proceedings. In C. Anagnostopoulou, M. Ferrand, & A. Smaill (Eds.), (pp. 5–17). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/3-540-45722-4_3

Bolton, T. L. (1894). Rhythm. *The American Journal of Psychology*, *6*, 145–238. doi: 10.2307/1410948

Boltz, M. G. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics*, *53*, 585–600. doi: 10.3758/BF03211736

Bornkessel, I. D., Fiebach, C. J., & Friederici, A. D. (2004). On the cost of syntactic ambiguity in human language comprehension: and individual differences approach. *Cognitive Brain Research*, *21*(1), 11–21. doi: 10.1016/j.cogbrainres.2004.05.007

Broca, P. P. (1861). Remarques sur le siège de la faculté du langage articulé, suivi d'une observation d'aphémie (perte de parole). *Bulletin de la Société Anatomique*, *6*, 330–357.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, *14*(4), 362–366. doi: 10.1111/1467-9280.24441

Brod, G., & Opitz, B. (2012). Does it really matter? separating the effects of musical training on syntax acquisition. *Frontiers in Psychology*, *3*(543), 1–8. doi: 10.3389/fpsyg.2012.00543

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the p600 in language comprehension. *Brain Research*, *1446*, 127–143. doi: 10.1016/j.brainres.2012.01.055

Buiatti, M., Peña, M., & Dehaene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *Neuroimage*, *44*(2), 509–519. doi: 10.1016/j.neuroimage.2008.09.015

Buonomano, D. V. (2007). The biology of time across different scales. *Nature Chemical Biology*, *3*, 594–597. doi: 10.1038/nchembio1007-594

Buszáki, G. (2006). *Rhythms of the brain*. New York: Oxford University Press.

Caplan, D., & Waters, G. (2013). Memory mechanisms supporting syntactic comprehension. *Psychonomic bulletin & review*, *20*(2), 243–268. doi: 10.3758/s13423-012-0369-9

Carrus, E., Pearce, M. T., & Bhattacharya, J. (2013). Melodic pitch expectation interacts with neural responses to syntactic but not semantic violations. *Cortex*, *49*(8), 2186–2200. doi: 10.1016/j.cortex.2012.08.024

Cason, N., & Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia*, *50*(11), 2652–2658. doi: 10.1016/j.neuropsychologia.2012.07.018

Chan, A. S., Ho, Y.-C., & Cheung, M.-C. (1998). Music training improves verbal memory. *Nature*, *396*(6707), 128. doi: 10.1038/24075

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive psychology*, *4*(1), 55–81. doi: 10.1016/0010-0285(73)90004-2

Chomsky, N. (1957). *Syntactic structures*. Berlin: Walter de Gruyter.

Christensen, T. (Ed.). (2006). *The cambridge history of western music theory*. Cambridge University Press.

Church, R. M., & Broadbent, H. A. (1990). Alternative representations of time, number, and rate. *Cognition*, *37*(1–2), 55–81. doi: 10.1016/0010-0277(90)90018-F

Collette, F., & der Linden, M. V. (2002). Brain imaging of the central executive component of working memory. *Neuroscience & Biobehavioral Reviews*, *26*(2), 105–125. doi: 10.1016/S0149-7634(01)00063-X

Collins, T., Tillmann, B., Barrett, F. S., Delbé, C., & Janata, P. (2014). A combined model of sensory and cognitive representations underlying tonal expectations in music: From audio signals to behavior. *Psychological Review*, *121*(1), 33–65. doi: 10.1037/a0034695

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547–552. doi: 10.1016/j.tics.2003.10.005

Corr, P. J. (2008). Reinforcement sensitivity theory (rst): Introduction. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 1–43). Cambridge: Cambridge University Press.

Corr, P. J., Revelle, W., Wilt, J., & Rosenthal, A. (2010). General models of individual differences in cognition: The commentaries. In A. Gruszka, G. Matthews, & B. Szymura (Eds.), *Handbook of individual differences in cognition: Attention, memory, and executive control* (pp. 63–70). New York: Springer.

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in english. *Journal of Phonetics*, *26*(2), 145–171. doi: 10.1006/jpho.1998.0070

Dalla Bella, S., Farrugia, N., Benoit, C.-E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2016). *Baasta: A battery for the assessment of auditory sensorimotor and timing abilities.* Submitted.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. doi: 10.1016/S0022-5371(80)90312-6

Darwin, C. J., & Donovan, A. (1980). Spoken language generation and understanding: Proceedings of the nato advanced study institute held at bonas, france, june 26 – july 7, 1979. In J. C. Simon (Ed.), (pp. 77–85). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-009-9091-3_4

Demorest, S. M., & Osterhout, L. (2012). Erp responses to cross-cultural melodic expectancy violations. *Annals of the New York Academy of Sciences*, *1252*(1), 152–157. doi: 10.1111/j.1749-6632.2012.06464.x

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158—164. doi: 10.1038/nn.4186

Ding, N., & Simon, J. (2009). Neural representations of complex temporal modulations in the human auditory cortex. *Journal of Neurophysiology*, *102*(5), 2731–2743. doi: 10.1152/jn.00523.2009

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, *85*(2), 761–768. doi: 10.1016/j.neuroimage.2013.06.035

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences*, *112*(45), E6233-E6242. doi: 10.1073/pnas.1508431112

Donner, T. H., & Siegel, M. (2011). A framework for local cortical oscillation patterns. *Trends in Cognitive Sciences*, *15*(5), 191–199. doi: 10.1016/j.tics.2011.03.007

Eisenberg, P. (1991). Syllabische struktur und wortakzent. prinzipien der prosodik deutscher wörter. *Zeitschrift für Sprachwissenschaft*, *10*(1), 37—64. doi: 10.1515/zfsw.1991.10.1.37

Eisler, H., Eisler, A. D., & Hellström, . (2008). Psychophysical issues in the study of time perception. In S. Grondin (Ed.), *Psychology of time* (pp. 75–110). Bingley, U.K: Emerald Group.

Ellis, R. J., & Jones, M. R. (2010). Rhythmic context modulates foreperiod effects. *Attention, Perception, & Psychophysics*, *72*(8), 2274–2288. doi: 10.3758/APP.72.8.2274

Ellis, R. J., Norton, A. C., Overy, K., Winner, E., Alsop, D. C., & Schlaug, G. (2012). Differentiating maturational and training influences on fmri activation during music processing. *Neuroimage*, *60*(3), 1902–1912. doi: 10.1016/j.neuroimage.2012.01.138

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*, *128*(3), 309–331. doi: 10.1037/0096-3445.128.3.309

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. doi: 10.1037/0033-295X.100.3.363

Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard Business Review*, *85*, 114–121. Retrieved from `https://hbr.org`

Ericsson, K. A., & Smith, J. (1991). Prospects and limits of the empirical study of expertise: An introduction. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 1–39). Cambridge, U.K.: Cambridge University Press.

Evers, S., Dannert, J., Rödding, D., Rötter, G., & Ringelstein, E.-B. (1999). The cerebral haemodynamics of music perception: A transcranial doppler sonography study. *Brain*, *122*(1), 75–85. doi: 10.1093/brain/122.1.75

Fabb, N. (2001). Linguistics and literature. In M. Aronoff & J. Rees-Miller (Eds.), *The handbook of linguistics* (pp. 446–465). Maiden, MA.: Blackwell Publishers.

Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic Bulletin & Review*, *2*(4), 460–493. doi: 10.3758/BF03210983

Featherstone, C. R., Morrison, C. M., Waterman, M. G., & MacGregor, L. J. (2013). Semantics, syntax, or neither? a case for resolution in the interpretation of n500 and p600 responses to harmonic incongruities. *PloS ONE*, *8*(11), e76600. doi: 10.1371/journal.pone.007660

Featherstone, C. R., Waterman, M. G., & Morrison, C. M. (2011). Norming the odd: Creation, norming, and validation of a stimulus set for the study of incongruities across music and language. *Behavior Research Methods*, *44*(1), 81–94. doi: 10.3758/s13428-011-0137-1

Fedorenko, E., Patel, A. D., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory & Cognition*, *31*(1), 1–9. doi: 10.3758/MC.37.1.1

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: Insights from erps and fmri. *Journal of Psycholinguistic Research*, *30*(3), 321–338. doi: 10.1023/A:1010447102554

Field, A. (2005). *Discovering statistics using spss* (2nd ed.). Sage.

Fitzroy, A., & Sanders, L. D. (2013). Musical expertise modulates early processing of syntactic violations in language. *Frontiers in Auditory Cognitive Neuroscience*, *3*(603), 1–15. doi: 10.3389/fpsyg.2012.00603

Fitzroy, A., & Sanders, L. D. (2015). Musical meter modulates the allocation of attention across time. *Journal of Cognitive Neuroscience*, *27*(12), 2339–2351. doi: 10.1162/jocn_a_00862

Fiveash, A., & Pammer, K. (2014). Music and language: Do they draw on similar syntactic working memory resources? *Psychology of Music*, *42*(2), 190–209. doi: 10.1177/0305735612463949

Fodor, J. D. (2002). Prosodic disambiguation in silent reading. In *Proceedings of the north east linguistic society.*

Fraisse, P. (1957). *Psychologie du temps [psychology of time]*. Paris: Presses Universitaires de France.

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronic sequence. *Journal of the Acoustical Society of America*, *98*, 2524–2531. doi: 10.1121/1.413218

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*(3), 259–281. doi: 10.1006/brln.1995.1048

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78–84. doi: 10.1016/S1364-6613(00)01839-8

Friederici, A. D., & Frisch, S. (2000). Verb argument structure processing: the role of verb-specific and argument–specific information. *Journal of memory and language*, *43*(3), 476–507. doi: 10.1006/jmla.2000.2709

Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research*, *31*(1), 45–63. doi: 10.1023/A:1014376204525

Friederici, A. D., Kotz, S. A., Scott, S. K., & Obleser, J. (2009). Disentangling syntax and intelligibility in auditory language comprehension. *Human Brain Mapping*, *31*(3), 448–457. doi: 10.1002/hbm.20878

Friederici, A. D., Kotz, S. A., Werheid, K., Hein, G., & von Cramon, D. Y. (2003). Syntactic comprehension in parkinson's disease: Investigating early automatic and late integrational processes using event-related brain potentials. *Neuropsychology*, *17*(1), 133–142. doi: 10.1037/0894-4105.17.1.133

Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., & Donchin, E. (2000). Syntactic parsing preferences and their on-line revisions: a spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research*, *11*(2), 305–323. doi:

10.1016/S0926-6410(00)00065-3

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*(3), 183–192. doi: 10.1016/0926-6410(93)90026-2

Friederici, A. D., Steinhauer, K., Mecklinger, A., & Meyer, M. (1998). Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses. *Biological Psychology*, *47*(3), 193-221. doi: 10.1016/S0301-0511(97)00033-1

Friederici, A. D., von Cramon, D. Y., & Kotz, S. A. (1999). Language related brain potentials in patients with cortical and subcortical left hemisphere lesions. *Brain*, *122*(6), 1033–1047. doi: 10.1093/brain/122.6.1033

Friederici, A. D., Wang, Y., Herrmann, C. S., Maess, B., & Oertel, U. (2000). Localization of early syntactic processes in frontal and temporal cortical areas: A magnetoencephalographic study. *Human Brain Mapping*, *11*(1), 1–11. doi: 10.1002/1097-0193(200009)11:1<1::AID-HBM10>3.0.CO;2-B

Frisch, S., Kotz, S. A., von Cramon, D. Y., & Friederici, A. D. (2003). Why the p600 is not just a p300: the role of the basal ganglia. *Clinical Neurophysiology*, *114*(2), 336–340. doi: 10.1016/S1388-2457(02)00366-8

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The p600 as an indicator of syntactic ambiguity. *Cognition*, *85*(3), B83 - B92. doi: 10.1016/S0010-0277(02)00126-9

Féry, C. (1997). Uni und studis: Die besten wörter des deutschen. *Linguistische Berichte*, *172*, 461–489.

Fundi, D. (2008). *Edward abbey: Freedom & wilderness i.* Retrieved from `http://www.podcastcafe.org/radiofreefundi/MP3/Edward_Abbey_1.mp3`

Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception–training increases hierarchical metre processing. *European Journal of Neuroscience*, *32*(11), 1979–1985. doi: 10.1111/j.1460-9568.2010.07462.x

Geiser, E., Ziegler, E., Jancke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex*, *45*(1), 93–102. doi: 10.1016/j.cortex.2007.09.010

Gibbon, J. (1977). Scalar expectancy theory and weber's law in animal timing. *Psychological Review*, *84*(3), 279–325. doi: 10.1037/0033-295X.84.3.279

Gibson, E. (2000). The dependency locality theory: a distance-based theory of linguistic complexity. In *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT Press.

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517. doi:

10.1038/nn.3063

Gordon, R. L., Magne, C. L., , & Large, E. W. (2011). Eeg correlates of song prosody: A new look at the relationship between linguistic and musical rhythm. *Frontiers in Psychology*, *2*(352). doi: 10.3389/fpsyg.2011.00352

Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & McAuley, J. D. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science*, *18*(4), 635–644. doi: 10.1111/desc.12230

Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: converging perspectives. *Laboratory Phonology*, *4*(1), 67–92. doi: 10.1515/lp-2013-0004

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the p600. *Language and Cognitive Processes*, *25*(2), 149–188. doi: 10.1080/01690960902965951

Graffi, G. (2001). *200 years of syntax: A critical survey* (Vol. 98). John Benjamins Publishing.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906. doi: 10.1162/jocn.2007.19.5.893

Grahn, J. A., & Brett, M. (2009). Impairment of beat-based rhythm discrimination in parkinson's disease. *Cortex*, *45*(1), 54–61. doi: 10.1016/j.cortex.2008.01.005

Grahn, J. A., & Rowe, J. B. (2013). Finding and feeling the musical beat: Striatal dissociations between detection and prediction of regularity. *Cerebral Cortex*, *23*(4), 913.

Grassi, M., & Soranzo, A. (2009). Mlp: A matlab toolbox for rapid and reliable auditory threshold estimation. *Behavior Research Methods*, *41*(1), 20–28. doi: 10.3758/BRM.41.1.20

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech – a syllable-centric perspective. *Journal of Phonetics*, *31*(3–4), 465–485. doi: 10.1016/j.wocn.2003.09.005

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, *72*(3), 561–582. doi: 10.3758/APP.72.3.561

Grube, M., & Griffiths, T. D. (2009). Metricality-enhanced temporal encoding and the subjective perception of rhythmic sequences. *Cortex*, *45*(1), 72–79. doi: 10.1016/j.cortex.2008.01.006

Gunter, T. C., Friederici, A. D., & Hahne, A. (1999). Brain responses during sentence reading: Visual input affects central processes. *NeuroReport*, *10*(15), 3175–3178. Retrieved from http://journals.lww.com/neuroreport/Abstract/1999/10190/Brain_responses_during_sentence_reading__visual.9.aspx

Gunter, T. C., Wagner, S., & Friederici, A. D. (2003). Working memory and lexical ambiguity resolution as revealed by erps: A difficult case for activation theories. *Journal of Cognitive Neuroscience*, *15*(5), 643–657. doi: 10.1162/jocn.2003.15.5.643

Hagoort, P., Wassenaar, M., & Brown, C. M. (2003). Syntax-related erp-effects in dutch. *Cognitive Brain Research*, *16*(1), 38–50. doi: 10.1016/S0926-6410(02)00208-2

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11*(2), 194–205. doi: 10.1162/089892999563328

Hambrick, D. Z., & Meinz, E. J. (2011). Limits on the predictive power of domain-specific experience and knowledge in skilled performance. *Current Directions in Psychological Science*, *20*(5), 275–279. doi: 10.1177/0963721411422061

Handy, T. C. (2005). Basic principles of erp quantification. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook* (pp. 33–56). Cambridge MA: MIT press.

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in percieving musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 956–974. doi: 10.1037/0096-1523.30.5.956

Harrison, M. (1995). *Contemporary music theory: Level one*. Milwaukee: Hal Leonard.

Hausdorff, J. M., Doniger, G. M., Springer, S., Yogev, G., Simon, E. S., & Giladi, N. (2006). A common cognitive profile in elderly fallers and in patients with parkinson's disease: the prominence of impaired executive function and attention. *Experimental Aging Research*, *32*(4), 411–429. doi: 10.1080/03610730600875817

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.

Hemforth, B., Konieczny, L., & Scheepers, C. (2000). Relative clauses and coordinations. In B. Hemforth & L. Konieczny (Eds.), *German sentence processing* (pp. 161–186). Dordrecht: Kluwer.

Henry, M., & Hermann, B. (2014). Low-frequency neural oscillations support dynamic attending in temporal context. *Timing and Time Perception*, *2*, 62–86. doi: 10.1163/22134468-00002011

Herholz, S. C., & Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron*, *76*(3), 486–502. doi: 10.1016/j.neuron.2012.10.011

Hevner, K. (1935). The affective character of the major and minor modes in music. *The American Journal of Psychology*, *47*(1), 103–118. doi: 10.2307/1416710

Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology*, *17*(3), 439–450. doi: 10.1037/0894-4105.17.3.439

Hocking, I. (2003). *Resources and parsing* (Unpublished doctoral dissertation). University of Exeter.

Hulme, C., & Melby-Lervåg, M. (2012). Current evidence does not support the claims made for cogmed working memory training. *Journal of Applied Research in Memory and Cognition*, *1*(3), 197–200. doi: 10.1016/j.jarmac.2012.06.006

Humphreys, M. S., & Revelle, W. (1984). Personality, motivation, and performance: A theory of the relationship between individual differences and information processing. *Psychological Review*, *91*(2), 153–184. doi: 10.1037/0033-295X.91.2.153

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science*, *15*(5), 356–360. doi: 10.1111/j.0956-7976.2004.00683.x

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences*, *108*(25), 10081–10086. doi: 10.1073/pnas.1103228108

James, C. E., Dupuis-Lozeron, E., & Hauert, C. A. (2012). Appraisal of musical syntax violations by primary school children: Effects of age and practice. *Swiss Journal of Psychology*, *71*(3), 161–168. doi: 10.1024/1421-0185/a000084

Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *NeuroImage*, *47*(2), 735–744. doi: 10.1016/j.neuroimage.2009.04.090

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, *83*(5), 323–355. doi: 10.1037/0033-295X.83.5.323

Jones, M. R. (2008). Musical time. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 81–92). Oxford, UK: Oxford University Press.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*(3), 459–491. doi: 10.1037/0033-295X.96.3.459

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, *13*(4), 313–319. doi: 10.1111/1467-9280.00458

Jongsma, M. L. A., Desain, P., & Honing, H. (2004). Rhythmic context influences the auditory evoked potentials of musicians and non-musicians. *Biological Psychology*, *66*(2), 129–152. doi: 10.1016/j.biopsycho.2003.10.002

Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, *15*(1), 98–100. doi: 10.1162/089892903321107855

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (1991). *Principles of neural science*. Connecticut: Appleton & Lange.

Kappenman, E. S., & Luck, S. J. (2012). Erp components: The ups and downs of brainwave recordings. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford handbook of event-related potential components* (pp. 3–30). Oxford, UK: Oxford University Press.

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: Encoding time in neural network states. *Neuron*, *53*(3), 427-–438. doi: 10.1016/j.neuron.2007.01.006

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, *15*(2), 207–212. doi: 10.1016/j.conb.2005.03.005

Koelsch, S. (2011a). Towards a neural basis of processing musical semantics. *Physics of Life Reviews*, *8*(2), 89–105. doi: 10.1016/j.plrev.2011.04.004

Koelsch, S. (2011b). Transitional zones of meaning and semantics in music and language: Reply to comments on “towards a neural basis of processing musical semantics”. *Physics of Life Reviews*, *8*(2), 125–128. doi: 10.1016/j.plrev.2011.05.011

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fmri study. *NeuroImage*, *25*(4), 1068–1076. doi: 10.1016/j.neuroimage.2004.12.050

Koelsch, S., Gunter, T. C., Friederici, A. D., & Schroeger, E. (2000). Brain indices of music processing: “non-musicians” are musical. *Journal of Cognitive Neuroscience*, *12*(3), 520–541. doi: 10.1162/089892900562183

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An erp study. *Journal of Cognitive Neuroscience*, *17*(10), 1565–1577. doi: 10.1162/089892905774597290

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Interaction between syntax processing in language and in music: An erp study. *Psychophysiology*, *44*(3), 476–490. doi: 10.1111/j.1469-8986.2007.00517.x

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences*, *110*(38), 15443–15448. doi: 10.1073/pnas.1300272110

Koelsch, S., Schmidt, B. H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*, *39*, 657–663. doi: 10.1017.S0048577202010508

Koelsch, S., Schroeger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *NeuroReport*, *10*(6), 1309–1313. Retrieved from `http://journals.lww.com/neuroreport`

Kohler, K. J. (2009). Rhythm in speech and language. *Phonetica*, *66*(1–2), 29–45. doi: 10.1159/000208929

Kotz, S. A., Frisch, S., von Cramon, D. Y., & Friederici, A. D. (2003). Syntactic language processing: Erp lesion data on the role of the basal ganglia. *Journal of the International Neuropsychological Society*, *9*(7), 1053–1060. doi: 10.1017/S1355617703970093

Kotz, S. A., & Gunter, T. C. (2015). Can rhythmic auditory cuing remediate language-related deficits in parkinson's disease? *Annals of the New York Academy of Sciences*, *1337*(1), 62–68. doi: 10.1111/nyas.12657

Kotz, S. A., Gunter, T. C., & Wonneberger, S. (2005). The basal ganglia are receptive to rhythmic compensation during auditory syntactic processing: Erp patient data. *Brain and Language*, *95*(1), 70–71. doi: 10.1016/j.bandl.2005.07.039

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Sciences*, *14*(9), 392–399. doi: 10.1016/j.tics.2010.06.005

Kotz, S. A., Schwartze, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex*, *45*(8), 982–990. doi: 10.1016/j.cortex.2009.02.010

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*(8), 599–605. doi: 10.1038/nrn2882

Kraus, N., Skoe, E., Parbery-Clark, A., & Ashley, R. (2009). Experience-induced malleability in neural encoding of pitch, timbre and timing: implications for language and music. *Annals of the New York Academy of Sciences*, *1169*, 543–557. doi: 10.1111/j.1749-6632.2009.04549.x

Kraus, N., Strait, D. L., & Parbery-Clark, A. (2012). Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. *Annals of the New York Academy of Sciences*, *1252*, 100–107. doi: 10.1111/j.1749-6632.2012.06463.x

Kuck, H., Grossbach, M., Bangert, M., & Altenmüller, E. (2003). Brain processing of meter and rhythm in music: Electrophysiological evidence for a common network. *Annals of the New York Academy of Sciences*, *999*, 244–253. doi: 10.1196/annals.1284.035

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49. doi: 10.1016/j.brainres.2006.12.063

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. doi: 10.1016/S1364-6613(00)01560-6

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the n400 component of the event-related brain potential (erp). *Annual Review of*

*Psychology*, *62*, 621–647. doi: 10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. Retrieved from http://www.sciencemag.org

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. doi: 10.1525/mp.2009.26.4.377

Large, E. W. (2008). Resonating to musical rhythm: theory and experiment. In S. Grondin (Ed.), *The psychology of time* (pp. 189–231). United Kingdom: Emerald.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How we track time-varying events. *Psychological Review*, *106*(1), 119–159. doi: 10.1037/0033-295X.106.1.119

Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, *6*(2–3), 177–208. doi: 10.1080/09540099408915723

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*(1), 1–37. doi: 10.1207/s15516709cog2601_1

Large, E. W., & Snyder, J. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, *1169*(1), 46–57. doi: 10.1111/j.1749-6632.2009.04550.x

Lee, D. (Ed.). (1974). *Plato: The republic* (2nd ed.). London, UK: Penguin Books.

Leonge, V. (2012). *Prosodic rhythm in the speech amplitude envelope: Amplitude modulation phase hierarchies (amphs) and amph models* (Unpublished doctoral dissertation). Cambridge, United Kingdom.

Leonge, V., & Goswami, U. (2005). Impaired extraction of speech rhythm from temporal modulation patterns in speech in developmental dyslexia. *Frontiers in human neuroscience*, *8*(96). doi: 10.3389/fnhum.2014.00096

Lerdahl, F. (2001). *Tonal pitch space*. Oxford University Press.

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

Levitin, D. J., & Menon, V. (2005). The neural locus of temporal structure and expectancies in music: Evidence from functional neuroimaging at 3 tesla. *Music Perception*, *22*(3), 563–575. doi: 10.1016/j.neuroimage.2003.08.016

Levy, G., Jacobs, D. M., Tang, M. X., Côté, L. J., Louis, E. D., Alfaro, B., … Marder, K. (2002). Memory and executive function impairment predict dementia in parkinson's disease. *Movement disorders*, *17*(6), 1221–1226. doi: 10.1002/mds.10280

Liberman, M. (1975). *The intonational system of english* (Doctoral dissertation, MIT, Cambridge, MA). Retrieved from http://hdl.handle.net/1721.1/27376

Liberman, M., & Prince, A. S. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, *8*(2), 249–336. Retrieved from http://www.mitpressjournals.org/loi/ling

Lidji, P., Palmer, C., Peretz, I., & Morningstar, M. (2011). Listeners feel the beat: entrainment to english and french speech rhythms. *Psychonomic Bulletin & Review*, *18*, 1035–1041. doi: 10.3758/s13423-011-0163-0

Lieberman, P., Friedman, J., & Feldman, L. S. (1990). Syntax comprehension deficits in parkinson's disease. *The Journal of Nervous and Mental Disease*, *178*(6), 360–365. Retrieved from http://journals.lww.com/jonmd

Lieberman, P., Kako, E., Friedman, J., Tajchman, G., Feldman, L. S., & Jiminez, E. B. (1992). Speech production, syntax comprehension, and cognitive deficits in parkinson's disease. *Brain and Language*, *43*(2), 169–189. doi: 10.1016/0093-934X(92)90127-Z

London, J. (2012). *Hearing in time: Psychological aspects of musical meter*. New York: Oxford University.

Lotze, M., Scheler, G., Tan, H. R., Braun, C., & Birbaumer, N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage*, *20*, 1817–1829. doi: 10.1016/j.neuroimage.2003.07.018

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, *54*(6), 1001–1010. doi: 10.1016/j.neuron.2007.06.004

Lyons, R. G. (2004). *Understanding digital signal processing* (2nd ed.). Upper Saddle River, NJ: Prentice Hall Professional Technical Reference.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in broca's area: An meg study. *Nature Neuroscience*, *4*(5), 540–545. doi: 10.1038/87502

Magne, C., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., & Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken french: Behavioral and electrophysiological evidence. *Cerebral Cortex*, *17*(11), 2659–2668. doi: 10.1093/cercor/bhl174

Magne, C., Gordon, R. L., & Midha, S. (2010). Influence of metrical expectancy on reading words: An erp study. In *Proceedings of fifth international conference on speech prosody, chicago, il.* Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.302.6043&rep=rep1&type=pdf

Makeig, S., & Onton, J. (2012). Erp features and eeg dynamics: An ica perspective. In

S. J. Luck & E. S. Kappenman (Eds.), *Oxford handbook of event-related potential components* (pp. 51–86). Oxford, UK: Oxford University Press.

Makuuchi, M., & Friederici, A. D. (2013). Hierarchical functional connectivity between the core language system and the working memory system. *Cortex*, *49*(9), 2416—2423. doi: 10.1016/j.cortex.2013.01.007

Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, *23*(2), 294–305. doi: 10.1162/jocn.2010.21413

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of eeg- and meg-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. doi: 10.1016/j.jneumeth.2007.03.024

*Martin luther king, jr.: I have a dream.* (n.d.). Retrieved from `http://www.americanrhetoric.com/speeches/mlkihaveadream.htm`

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. doi: 10.1037/0096-3445.135.3.348

Menninghaus, W., Bohrn, I. C., Knoop, C. A., Kotz, S. A., Schlotz, W., & Jacobs, A. M. (2015). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition*, *143*, 48–60. doi: 10.1016/j.cognition.2015.05.026

Meter. (1986). In D. M. Randel (Ed.), *The new harvard dictionary of music.* Cambridge, MA: The Belknap Press of Harvard University Press.

Meyer, L. (2012). *The working memory of argument-verb dependencies: Spatiotemporal brain dynamics during sentence processing* (Unpublished doctoral dissertation). MPI Series in Human Cognitive and Brain Sciences, 145, Leipzig.

Meyer, L., Obleser, J., & Friederici, A. D. (2013). Left parietal alpha enhancement during working memory-intensive sentence processing. *Cortex*, *49*(3), 711–721. doi: 10.1016/j.cortex.2012.03.006

Miendlarzewska, E. A., & Trost, W. J. (2013). How musical training affects cognitive development: rhythm, reward and other modulating variables. *Frontiers in Neuroscience*, *7*(279), 1–18. doi: 10.3389/fnins.2013.00279

Miyake, A., & Shah, P. (1999). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 442–481). Cambridge, UK: Cambridge University Press.

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011).

Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, *22*(11), 1425–1433. doi: 10.1177/0956797611416999

Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J., & Wolf, M. (2013). Links between early rhythm skills, musical training, and phonological awareness. *Reading and Writing*, *26*(5), 739–769. doi: 10.1007/s11145-012-9389-0

Mottier, G. (1951). Über untersuchungen der sprache lesegestörter kinder. *Folia Phoniatrica et Logopaedica*, *3*(3), 170–177. doi: 10.1159/000262507

Naik, G. R., & Kumar, D. K. (2011). An overview of independent component analysis and its applications. *Informatica*, *35*(1), 63–81. Retrieved from `http://www.informatica.si`

Nakano, M., Ohishi, Y., Kameoka, H., Mukai, R., & Kashino, K. (2012, March). Bayesian nonparametric music parser. In *Acoustics, speech and signal processing (icassp), 2012 ieee international conference on* (p. 461-464). doi: 10.1109/ICASSP.2012.6287916

Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., & Jonides, J. (2013). A meta-analysis of executive components of working memory. *Cerebral Cortex*, *23*(2), 264–282. doi: 10.1093/cercor/bhs007

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris Publications.

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes - evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *3*(2), 151–165. doi: 10.1162/jocn.1991.3.2.151

Nieminen, S., Istók, E., Brattico, E., & Tervaniemi, M. (2012). The development of the aesthetic experience of music: Preference, emotions, and beauty. *Musicae Scientiae*, *16*(3), 372–391. doi: 10.1177/1029864912450454

Nieuwland, M. S., & Berkum, J. J. A. V. (2006). Individual differences and contextual bias in pronoun resolution: Evidence from erps. *Brain Research*, *1118*(1), 155–167. doi: 10.1016/j.brainres.2006.08.022

Nisbett, R. E., Aronson, J., an W Dickens, C. B., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: new findings and theoretical developments. *American Psychologist*, *67*(2), 130. doi: 10.1037/a0026699

Nittono, H., Bito, T., Hayashi, M., Sakata, S., & Hori, T. (2000). Event-related potentials elicited by wrong terminal notes: effects of temporal disruption. *Biological Psychology*, *52*(1), 1–16. doi: 10.1016/S0301-0511(99)00042-3

Nozaradan, S. (2014). Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*(1658). doi: 10.1098/rstb.2013.0393

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience*, *31*(28), 10234–10240. doi: 10.1523/JNEUROSCI.0411-11.2011

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. doi: 10.1523/JNEUROSCI.3203-12.2012

Näätänen, R., & Alho, P. P. T. R. K. (2007). The mismatch negativity (mmn) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544–2590. doi: 10.1016/j.clinph.2007.04.026

Näätänen, R., Lehtokoski, A., Lennest, M., Luuki, A., Alliki, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*, 432–434. doi: 10.1038/385432a0

Näätänen, R., & Picton, T. (1987). The n1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, *24*, 375–425. doi: 10.1111/j.1469-8986.1987.tb00311.x

Obermeier, C., Kotz, S. A., Jessen, S., Raettig, T., von Koppenfels, M., & Menninghaus, W. (2016). Aesthetic appreciation of poetry correlates with ease of processing in event-related potentials. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(2), 362–373. doi: 10.3758/s13415-015-0396-x

Oechslin, M., Ville, D. V. D., Lazeyras, F., Hauert, C.-A., & James, C. E. (2013). Degree of musical expertise modulates higher order brain functioning. *Cerebral Cortex*, *23*(9), 2213–2224. doi: 10.1093/cercor/bhs206

Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., ... Imabayashi, E. (2001). Functional anatomy of musical perception in musicians. *Cerebral Cortex*, *11*(8), 754–760. doi: 10.1093/cercor/11.8.754

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). Fieldtrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*(156869). doi: 10.1155/2011/156869

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806. doi: 10.1016/0749-596X(92)90039-Z

Overy, K., Nicolson, R. I., Fawcett, A. J., & Clarke, E. F. (2001). Dyslexia and music: Measuring musical timing skills. *Dyslexia*, *9*(1), 18–36. doi: 10.1002/dys.233

Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, *22*(12), 2728–2744. doi: 10.1162/jocn.2009.21393

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal* *of Experimental Psychology: Human Perception and Performance*, *16*(4), 728–741. doi: 10.1037/0096-1523.16.4.728

Pantev, C. (1995). Evoked and induced gamma-band activity of the human cortex. *Brain Topography*, *7*(4), 321–330. doi: 10.1007/BF01195258

Pantev, C., & Herholz, S. C. (2011). Plasticity of the human auditory cortex related to musical training. *Neuroscience and Biobehavioral Reviews*, *35*(10), 2140–2154. doi: 10.1016/j.neubiorev.2011.06.010

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, *392*, 811–814. doi: 10.1038/33918

Parbery-Clark, A., Tierney, A., Strait, D. L., & Kraus, N. (2012). Musicians have fine-tuned neural distinction of speech syllables. *Journal of Neuroscience*, *219*, 111–119. doi: 10.1016/j.neuroscience.2012.05.042

Parncutt, R. (2012). Major-minor tonality, schenkerian prolongation, and emotion: A commentary on huron and davis. *Empirical Musicology Review*, *7*(3–4). Retrieved from `http://emusicology.org`

Pastor, M. A., Artieda, J., Jahanshahi, M., & Obeso, J. A. (1992). Time estimation and reproduction is abnormal in parkinson's disease. *Brain*, *115*(1), 211–255. doi: 10.1093/brain/115.1.211

Patel, A. D. (2003). Language, music, syntax, and the brain. *Nature Neuroscience*, *6*(7), 674–681. doi: 10.1038/nn1082

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? the opera hypothesis. *Frontiers in Psychology*, *2*. doi: 10.3389/fpsyg.2011.00142

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event related potential study. *Journal of Cognitive Neuroscience*, *10*(6), 717–733. doi: 10.1162/089892998563121

Patel, A. D., Iversen, J. R., Wassenaar, M., & Hagoort, P. (2008). Musical syntactic processing in agrammatic broca's aphasia. *Aphasiology*, *22*(7), 776–789. doi: 10.1080/02687030701803804

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage*, *50*(1), 302–313. doi: 10.1016/j.neuroimage.2009.12.019

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. *Annals of the New York Academy of Sciences*, *999*, 58–75. doi: 10.1196/annals.1284.006

Perfors, A., & Ong, J. H. (2012). Musicians are better at learning non-native sound contrasts even in non-tonal languages. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Pro-*

*ceedings of the 34th annual conference of the cognitive science society* (pp. 839–844). Austin, TX: Cognitive Science Society.

Perruchet, P., & Poulin-Charronnat, B. (2013). Challenging prior evidence for a shared syntactic processor for language and music. *Psychonomic Bulletin & Review*, *20*(2), 310–317. doi: 10.3758/s13423-012-0344-5

Piston, W. (1941). *Harmony*. New York: W. W. Norton.

Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(3), 564–573. doi: 10.1037/0096-1523.16.3.564

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech communication*, *41*(1), 245–255. doi: 10.1016/S0167-6393(02)00107-3

Polanía, R., Nitsche, M. A., Korman, C., Batsikadze, G., & Paulus, W. (2012). The importance of timing in segregated theta phase-coupling for cognitive performance. *Current Biology*, *22*(14), 1314–1318. doi: 10.1016/j.cub.2012.05.021

Polich, J. (2007). Updating p300: An integrative theory of p3a and p3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Port, R. F. (2003). Meter and speech. *Journal of Phonetics*, *31*(3–4), 599–611. doi: 10.1016/j.wocn.2003.08.001

Poulin-Charronnat, B., Bigand, E., & Koelsch, S. (2006). Processing of musical syntax tonic versus subdominant: an event-related potential study. *Journal of Cognitive Neuroscience*, *18*(9), 1545–1554. doi: 10.1162/jocn.2006.18.9.1545

Prince, A. S. (1983). Relating to the grid. *Linguistic Inquiry*, *14*(1), 19–100. Retrieved from `http://www.mitpressjournals.org/loi/ling`

Prince, A. S., & Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell. (Revision of 1993 technical report, Rutgers University Center for Cognitive Science)

Rammsayer, T., & Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception: An Interdisciplinary Journal*, *24*(1), 37–48. doi: 10.1525/mp.2006.24.1.37

Ramsay, M. C., & Reynolds, C. R. (1995). Separate digits tests: A brief history, a literature review, and a reexamination of the factor structure of the test of memory and learning (tomal). *Neuropsychology Review*, *5*(3), 151–171. doi: 10.1007/BF02214760

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., ... Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, *142*(2), 359. doi: 10.1037/a0029082

Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. *Journal of Cognitive Neuroscience*, *13*(2), 241–255. doi: 10.1162/089892901564298

Reilly, E. L. (2005). Eeg recording and operation of the apparatus. In E. Niedermeyer & F. L. da Silva (Eds.), *Electroencephalography: Basic principles, clinical applications, and related fields* (5th ed., pp. 139–159). Philadelphia: Lippencott Williams and Wilkins.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review*, *12*(6), 969–992. doi: 10.3758/BF03206433

Revelle, W. (1993). Individual differences in personality and motivation: 'non-cognitive' determinants of cognitive performance. In A. Baddely & L. Weiskranz (Eds.), *Attention: Selection, awareness, and control: A tribute to donald broadbent* (pp. 347–373). New York: Clarendon Press/Oxford University Press.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, *5*(1), 35–53. doi: 10.1080/17459737.2011.573676

Roncaglia-Denissen, M. P., Schmidt-Kassow, M., & Kotz, S. A. (2013). Speech rhythm facilitates syntactic ambiguity resolution: Erp evidence. *PLoS One*, *8*(2), e56000. doi: 10.1371/journal.pone.0056000

Rothermich, K., & Kotz, S. A. (2013). Predictions in speech comprehension: fmri evidence on the meter-semantic interface. *NeuroImage*, *70*, 89–100. doi: 10.1016/j.neuroimage.2012.12.013

Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2012). Rhythm's gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia*, *50*(2), 232–244. doi: 10.1016/j.neuropsychologia.2011.10.025

Rothermich, K., Schmidt-Kassow, M., Schwartze, M., & Kotz, S. A. (2010). Event-related potential responses to metric violations: rules versus meaning. *NeuroReport*, *21*(8), 580–584. doi: 10.1097/WNR.0b013e32833a7da7

Sammler, D., Koelsch, S., Ball, T., Brandt, A., Grigutsch, M., ..., H.-J. H., & Schulze-Bonhage, A. (2013). Co-localizing linguistic and musical syntax with intracranial eeg. *NeuroImage*, *64*, 134–146. doi: 10.1016/j.neuroimage.2012.09.035

Sammler, D., Koelsch, S., & Friederici, A. D. (2011). Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex*, *47*(6), 659–673. doi: 10.1016/j.cortex.2010.04.007

Santee, J. L., & Egeth, H. E. (1982). Do reaction time and accuracy measure the same aspects of letter recognition? *Journal of Experimental Psychology: Human Perception and Performance*, *8*(4), 489–501. doi: 10.1037/0096-1523.8.4.489

Sanz, M., Laka, I., & Tanenhaus, M. K. (Eds.). (2013). *Language down the garden path: The cognitive and biological basis for linguistic structurese*. Oxford University Press. doi: 10.1093/acprof:oso/9780199677139.001.0001

Schlaug, G. (2001). The brain of musicians. *Annals of the New York Academy of Sciences*, *930*(1), 281–299. doi: 10.1111/j.1749-6632.2001.tb05739.x

Schlewesky, M., Fanselow, G., Kliegl, R., & Krems, J. (2000). The subject preference in the processing of locally ambiguous wh-questions in german. In B. Hemforth & L. Konieczny (Eds.), *German sentence processing* (pp. 65–93). Dordrecht: Kluwer.

Schmidt-Kassow, M., & Kotz, S. A. (2008). Entrainment of syntactic processing? erp-responses to predictable time intervals during syntactic reanalysis. *Brain Research*, *1226*, 144–155. doi: 10.1016/j.brainres.2008.06.017

Schmidt-Kassow, M., & Kotz, S. A. (2009a). Attention and perceptual regularity in speech. *NeuroReport*, *20*(18), 1643–1647. doi: 10.1097/WNR.0b013e328333b0c6

Schmidt-Kassow, M., & Kotz, S. A. (2009b). Event-related brain potentials suggest a late interaction of meter and syntax in the p600. *Journal of Cognitive Neuroscience*, *21*(9), 1693–1708. doi: 10.1162/jocn.2008.21153

Schmidt-Kassow, M., Rothermich, K., Schwartze, M., & Kotz, S. A. (2011). Attention and perceptual regularity in speech. *NeuroImage*, *54*(1), 568–576. doi: 10.1016/j.neuroimage.2010.07.062

Schmuckler, M. A., & Boltz, M. G. (1994). Harmonic and rhythmic influences on musical expectancy. *Perception & Psychophysics*, *56*(3), 313–325. doi: 10.3758/BF03209765

Schön, D., & Besson, M. (2005). Visually induced auditory expectancy in music reading: a behavioural and electrophysiological study. *Journal of Cognitive Neuroscience*, *17*(4), 694–705. doi: 10.1162/0898929053467532

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. doi: 10.1016/j.tins.2008.09.012

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2011). Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biological Psychology*, *87*(1), 146–151. doi: 10.1016/j.biopsycho.2011.02.021

Seger, C. A., Spiering, B. J., Sares, A. G., Quraini, S. I., Alpeter, C., David, J., & Thaut, M. H. (2013). Corticostriatal contributions to musical expectancy perception. *Journal of Cognitive Neuroscience*, *25*(7), 1062–1077. doi: 10.1162/jocn_a_00371

Selkirk, E. O. (1980). The role of prosodic categories in english word stress. *Linguistic Inquiry*, *11*(3), 563–605. Retrieved from `http://www.mitpressjournals.org/loi/ling`

Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Selkirk, E. O. (2011). The syntax-phonology interface. In J. Goldsmith, J. Riggle, & A. Yu (Eds.), *The handbook of phonological theory.* (2nd ed., pp. 435–484). Oxford: Blackwell.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*(1), 4–27. doi: 10.1037/0096-3445.125.1.4

Sharbrough, F., Chatrian, G.-E., Lesser, R. P., Lüders, H., Nuwer, M., & Picton, T. W. (1991). American electroencephalographic society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *8*(2), 200–202.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*(4), 628. doi: 10.1037/a0027473

Shook, A., Marian, V., Bartolotti, J., & Schroeder, S. R. (2013). Musical experience influences statistical learning of a novel language. *The American Journal of Psychology*, *126*(1), 95–104. doi: 10.3758/BF03209765

Singer, W. (2013). Cortical dynamics revisited. *Trends in Cognitive Sciences*, *17*(12), 616–626. doi: 10.1016/j.tics.2013.09.006

Slevc, L. R., & Okada, B. M. (2015). Processing structure in language and music: A case for shared reliance on cognitive control. *Psychonomic Bulletin & Review*, *22*(3), 637–652. doi: 10.3758/s13423-014-0712-4

Slevc, L. R., Rosenberg, J. C., & Patel, A. D. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review*, *16*(2), 374–381. doi: 10.3758/16.2.374

Sloboda, J. (1991). Musical expertise. In A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 153–171). Cambridge, UK: Cambridge University Press.

Snyder, J., & Large, E. W. (2004). Tempo dependence of middle-and long-latency auditory responses: power and phase modulation of the eeg at multiple time-scales. *Clinical Neurophysiology*, *115*(8), 1885–1895. doi: 10.1016/j.clinph.2004.03.024

Snyder, J., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, *24*(1), 117–126. doi: 10.1016/j.cogbrainres.2004.12.014

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2011). Perception of speech in noise: Neural correlates. *Journal of Cognitive Neuroscience*, *23*(9), 2268-2279. doi: 10.1162/jocn.2010.21556

Spearman, C. (1904). General intelligence, objectively determined and measured. *American*

*Journal of Psychology*, *15*(2), 201–292. doi: 10.2307/1412107

Spencer, K. M. (2005). Averaging, detection, and classification of single-trial erps. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook* (pp. 209–228). Cambridge MA: MIT press.

Spencer-Smith, M., & Klingberg, T. (2015). Benefits of a working memory training program for inattention in daily life: a systematic review and meta-analysis. *PloS one*, *10*(3), e0119522. doi: 10.1371/journal.pone.0119522

Steinbeis, N., & Koelsch, S. (2008). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex*, *18*(5), 1169–1178. doi: 10.1093/cercor/bhm149

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (elan) in syntax studies. *Brain and Language*, *120*(2), 135–162. doi: 10.1016/j.bandl.2011.07.001

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction time experiments. *American Scientist*, *57*, 421-–457. Retrieved from `http://www.americanscientist.org/`

Strait, D. L., O'Connell, S., Parbery-Clark, A., & Kraus, N. (2014). Musicians' enhanced neural differentiation of speech sounds arises early in life: Developmental evidence from ages three to thirty. *Cerebral Cortex*, *24*(9), 2512–2521. doi: 10.1093/cercor/bht103

Strait, D. L., Parbery-Clark, A., Hittner, E., & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and Language*, *123*(3), 191–201. doi: 10.1016/j.bandl.2012.09.001

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. doi: 10.1037/h0054651

Sussman, E., Ritter, W., & Vaughan, H. G. (1998). Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Research*, *789*(1), 130–138. doi: 10.1016/S0006-8993(97)01443-1

Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, *136*(1), 64–81. doi: 10.1037/0096-3445.136.1.64

Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: A case for the preeminence of temporal processing. *Annals of the New York Academy of Sciences*, *682*(6), 27–47. doi: 10.1111/j.1749-6632.1993.tb22957.x

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience*, *31*(10), 3805-–3812. doi: 10.1523/JNEUROSCI.5561-10.2011

Temperley, D., & Marvin, E. W. (2008). Pitch-class distribution and the identification of key. *Music Perception*, *25*(2), 193–212. doi: 10.1525/MP.2008.25.3.193

Teplan, M. (2002). Fundamentals of eeg measurement. *Measurement Science Review*, *2*(2). Retrieved from `http://www.measurement.sk`

Tervaniemi, M., Sannemann, C., Noyranen, M., Salonen, J., & Pihko, E. (2011). Importance of the left auditory areas in chord discrimination in music experts as demonstrated by meg. *European Journal of Neuroscience*, *34*(3), 517–523. doi: 10.1111/j.1460-9568.2011.07765.x

Tewes, U. (1994). *Hamburg: Wehsler-intelligenztest für erwachsene*. Bern: Huber: Prentice Hall Professional Technical Reference.

Thaut, M. H., Peterson, D. A., & McIntosh, G. C. (2005). Temporal entrainment of cognitive functions. *Annals of the New York Academy of Sciences*, *1060*(1), 243–254. doi: 10.1196/annals.1360.017

Thut, G., Schyns, P. G., & Gross, J. (2011). Entrainment of perceptually relevant brain oscillations by non-invasive rhythmic stimulation of the human brain. *Frontiers in Psychology*, *2*(170). doi: 10.3389/fpsyg.2011.00170

Tierney, A., & Kraus, N. (2015). Neural entrainment to the rhythmic structure of music. *Journal of Cognitive Neuroscience*, *27*(2), 400–408. doi: 10.1162/jocn_a_00704

Tierney, A., Krizman, J., Skoe, E., Johnston, K., & Kraus, N. (2013). High school music classes enhance the neural processing of speech. *Frontiers in Psychology*, *4*, 855. doi: 10.3389/fpsyg.2013.00855

Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A. D., & von Cramon, D. Y. (2006). Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex. *NeuroImage*, *31*, 1771–1782. doi: 10.1016/j.neuroimage.2006.02.028

Traxler, M. J. (2007). Working memory contributions to relative clause attachment processing: A hierarchical linear modeling analysis. *Memory & Cognition*, *35*(5), 1107–1121. doi: 10.3758/BF03193482

Treisman, M., Faulkner, A., Naish, P. L. N., & Brogan, D. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock". *Psychological Monographs:General and Applied*, *77*(13), 1–31. doi: 10.1037/h0093864

Treisman, M., Faulkner, A., Naish, P. L. N., & Brogan, D. (1990). The internal clock: Evidence for a temporal oscillation underlying time perception with some estimates of its characteristic frequency. *Perception*, *19*(6), 705–743. doi: 10.1068/p190705

Tuuri, K., & Eerola, T. (2012). Formulating a revised taxonomy for modes of listening. *Journal of New Music Research*, *41*(2), 137–152. doi: 10.1080/09298215.2011.614951

Tyler, L. K., Marslen-Wilson, W. D., Randall, B., Wright, P., Devereux, B. J., Zhuang, J., . . . Stamatakis, E. A. (2011). Left inferior frontal cortex and syntax: function, structure and behaviour in patients with left hemisphere damage. *Brain*, *134*(2), 415–431. doi: 10.1093/brain/awq369

Tzounopoulos, T., & Kraus, N. (2009). Learning to encode timing: Mechanisms of plasticity in the auditory brainstem. *Neuron*, *62*(4), 463–469. doi: 10.1016/j.neuron.2009.05.002

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26. doi: 10.1016/j.cogpsych.2014.01.003

Van Dyke, J. A., Johns, C. L., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition*, *131*(3), 373–403. doi: 10.1016/j.cognition.2014.01.007

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequencyinevent-related brainpotentials. *Memory & Cognition*, *18*(4), 380–393. doi: 10.3758/BF03197127

van Ede, F., de Lange, F. P., & Maris, E. (2012). Attentional cues affect accuracy and reaction time via different cognitive and neural processes. *The Journal of Neuroscience*, *32*(30), 10408-–10412. doi: 10.1523/JNEUROSCI.1337-12.2012

Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, *2*(4), 229–239. doi: 10.1038/35067550

von Salis-Seewis, J. G. (ca. 1800). *Ermunterung, 1–4.* Retrieved 25 November 2013 from http://gedichte.xbib.de/Salis-Seewis_gedicht_19.+Ermunterung.htm.

Vos, S. H., Gunter, T. C., Schriefers, H., & Friederici, A. D. (2001). Syntactic parsing and working memory: The effects of syntactic complexity, reading span, and concurrent load. *Language and Cognitive Processes*, *16*(1), 65–103. doi: 10.1080/01690960042000085

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music – brain responses to rhythmic incongruity. *Cortex*, *45*(1), 80–92. doi: 10.1016/j.cortex.2008.05.014

Vuust, P., Pallesen, K. J., Bailey, C., van Zuijen, T. L., Gjedde, A., Roepstorff, A., & Østergaard, L. (2005). To musicians, the message is in the meter: Pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage*, *24*(2), 560–564. doi: 10.1016/j.neuroimage.2004.08.039

Vuust, P., Wallentin, M., Mouridsen, K., Østergaard, L., & Roepstorff, A. (2011). Tapping polyrhythms in music activates language areas. *Neuroscience Letters*, *494*(3), 211–

216. doi: 10.1016/j.neulet.2011.03.015

Wagner, M. (2010). Prosody and recursion in coordinate structures and beyond. *Natural Language and Linguistic Theory*, *28*(1), 183–237. doi: 10.1007/s11049-009-9086-0

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The musical ear test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, *20*(3), 188–196. doi: 10.1016/j.lindif.2010.02.004

Wan, C. Y., & Schlaug, G. (2010). Music making as a tool for promoting brain plasticity across the life span. *The Neuroscientist*, *16*(5), 566–577. doi: 10.1177/1073858410377805

Wearden, J. H. (2003). Applying the scalar timing model to human time psychology: Progress and challenges. In H. Helfrich (Ed.), *Time and mind ii* (pp. 21–39). Göttingen: Hogrefe & Huber.

Wechsler, D. (1997). *Wais-iii: Wechsler adult intelligence scale*. San Antonio, TX: Psychological Corporation.

White-Schwoch, T., Carr, K. W., Anderson, S., Strait, D. L., & Kraus, N. (2013). Older adults benefit from music training early in life: Biological evidence for long-term training-driven plasticity. *The Journal of Neuroscience*, *33*(45), 17667–17674. doi: 10.1523/JNEUROSCI.2560-13.2013

White-Schwoch, T., & Kraus, N. (2013). Physiologic discrimination of stop consonants relates to phonological ability in pre-readers: A biomarker for subsequent reading ability? *Frontiers in Human Neuroscience*, *7*, 899. doi: 10.3389/fnhum.2013.00899

Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data – a practical approach. *Journal of Neuroscience Methods*, *250*, 34 - 46. doi: http://dx.doi.org/10.1016/j.jneumeth.2014.08.002

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420–422. doi: 10.1038/nn1872

Wright, C., & Simms, B. (2006). *Music in western civilization*. Belmont, CA: Thompson Schirmer.

Yarkoni, T., & Braver, T. S. (2010). Cognitive neuroscience approaches to individual differences in working memory and executive control: Conceptual and methodological issues. In A. Gruszka, G. Matthews, & B. Szymura (Eds.), *Handbook of individual differences in cognition: Attention, memory, and executive control* (p. 87—107). New York: Springer.

Ystad1, S., Magne, C., Farner, S., Pallone, G., Aramaki, M., Besson, M., & Kronland-Martinet, R. (2007). Electrophysiological study of algorithmically processed metric/rhythmic variations in language and music. *EURASIP Journal on Audio, Speech,* *and Music Processing*, *2007*(030194). doi: 10.1155/2007/30194

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, *11*(10), 946–953. doi: 10.1093/cercor/11.10.946

Zendel, B. R., & Alain, C. (2012). Musicians experience less age-related decline in central auditory processing. *Psychology and Aging*, *27*(2), 410–417. doi: 10.1037/a0024816

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| Ag-AgCl | silver-silver chloride |
| ANOVA | analysis of variance |
| bpm | beats per minute |
| dB | decibels |
| EEG | electroencephalography |
| EOG | electrooculography |
| ERP | event-related potential |
| fMRI | functional magnetic resonance imaging |
| ICA | independent component analysis |
| M | mean |
| MEG | magnetoencephalography |
| ms | milliseconds |
| $\mu$V | micro Volts |
| N | number (of participants) |
| ROI | region of interest |
| s | seconds |
| SD | standard deviation |
| SE | standard error |
| SNR | signal-to-noise ratio |